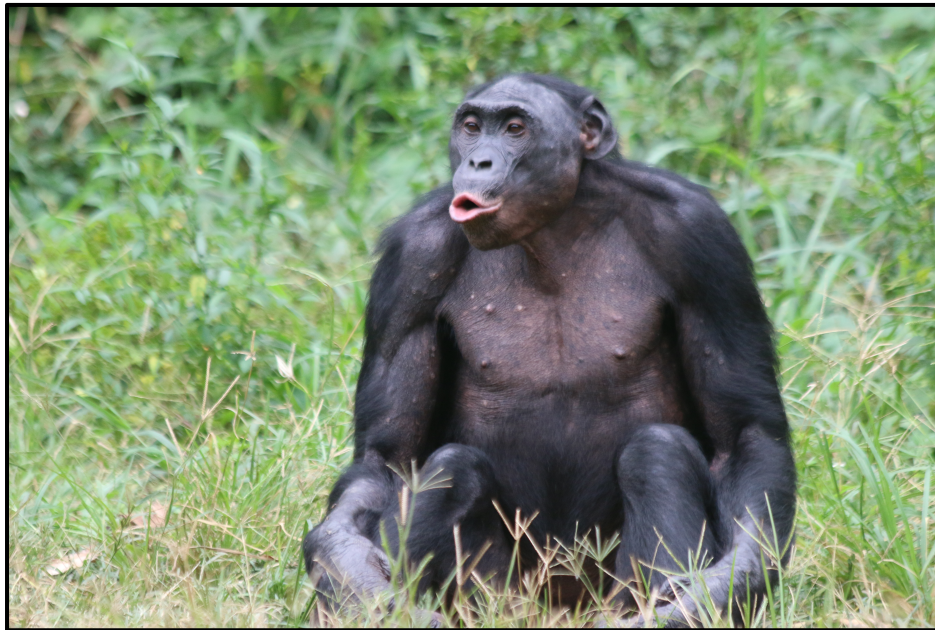THIS THESIS

is presented to

UNIVERSITÉ de LYON/SAINT-ETIENNE

in cotutelle with

UNIVERSITY OF ST. ANDREWS

for the purpose of a

DOCTORAL DEGREE

Publicly defended on 14 October 2016
by SUMIR KEENAN

IDENTITY INFORMATION IN BONOBO VOCAL COMMUNICATION : FROM SENDER TO RECEIVER

Thesis supervisors:
Florence LEVRÉRO, Nicolas MATHEVON & Klaus ZUBERBÜHLER

Jury Composition:
Isabelle CHARRIER : Associate Researcher, CNRS, Université Paris Sud   *Reviewer*
Christine DREA : Professor, Duke University   *Reviewer*
David REBY : Reader, University of Sussex   *Examiner*
Florence LEVRÉRO : Associate Professor, Université de Lyon/Saint-Etienne   *Thesis supervisor*
Nicolas MATHEVON : Professor, Université de Lyon/Saint-Etienne   *Thesis supervisor*
Klaus ZUBERBÜHLER : Professor, University of Neuchâtel   *Thesis supervisor*

Identity information in bonobo vocal communication : From sender to receiver

L' information "identité individuelle" dans la communication vocale du bonobo: de l'émetteur au récepteur

Sumir Keenan

Cover photo courtesy of Jeroen Stevens.

PUBLICATIONS

At the time of thesis submission the following manuscript has been published :

Keenan, S., Mathevon, N., Stevens, J. M. G., Guéry, J. P., Zuberbühler, K., & Levréro, F. (2016). Enduring voice recognition in bonobos. *Scientific Reports*, *6*, 22046.

At the time of thesis submission the following manuscript has been submitted :

Keenan, S., Mathevon, N., Nicolè, F., Stevens, J. M. G., Zuberbühler, K., & Levréro, F. Individual vocal signatures across the bonobos' graded repertoire. (*Submitted to Animal Behaviour*).

At the time of thesis submission the following manuscript is in preparation for submission :

Keenan, S., Stevens, J. M. G., Mathevon, N., Staes, N., Zuberbühler, K., & Levréro, F. Vocal convergence in bonobos - genetic and social influences. (*In preparation*).

1. Candidate's declarations:

I, Sumir Keenan, hereby certify that this thesis, which is approximately 55,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in December 2012 and as a candidate for the degree of Doctorate in Philosophy in Psychology in December 2013; the higher study for which this is a record was carried out in the University of St Andrews between 2012 and 2016.

Date 25/05/2017   signature of candidate

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctorate of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 25/05/2017   signature of supervisor

3. Permission for publication: (*to be signed by both candidate and supervisor*)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY
a)   No embargo on print copy
b)   Embargo on all or part of print copy for a period of *2 years* (maximum five) on the following ground(s):
   - Publication would be commercially damaging to the researcher, or to the supervisor, or the University
   - **Publication would preclude future publication**
   - Publication would be in breach of laws or ethics
c)   Permanent or longer term embargo on all or part of print copy for a period of … years (the request will be referred to the Pro-Provost and permission will be granted only in exceptional circumstances).

ELECTRONIC COPY
a)   No embargo on electronic copy
b)   Embargo on all or part of electronic copy for a period of *2 years* (maximum five) on the following ground(s):
   - Publication would be commercially damaging to the researcher, or to the supervisor, or the University
   - **Publication would preclude future publication**
   - Publication would be in breach of law or ethics
c)   Permanent or longer term embargo on all or part of electronic copy for a period of … years (the request will be referred to the Pro-Provost and permission will be granted only in exceptional circumstances).

ABSTRACT AND TITLE EMBARGOES

An embargo on the full text copy of your thesis in the electronic and printed formats will be granted automatically in the first instance. This embargo includes the abstract and title except that the title will be used in the graduation booklet.

If you have selected an embargo option indicate below if you wish to allow the thesis abstract and/or title to be published. If you do not complete the section below the title and abstract will remain embargoed along with the text of the thesis.

   a)   I agree to the title and abstract being published   **YES**
   b)   I require an embargo on abstract   **NO**
   c)   I require an embargo on title   **NO**

Date 25/05/17   Signature of candidate   signature of supervisor

Please note initial embargos can be requested for a maximum of five years. An embargo on a thesis submitted to the Faculty of Science or Medicine is rarely granted for more than two years in the first instance, without good justification. The Library will not lift an embargo before confirming with the student and supervisor that they do not intend to request a continuation. In the absence of an agreed response from both student and supervisor, the Head of School will be consulted. Please note that the total period of an embargo, including any continuation, is not expected to exceed ten years. Where part of a thesis is to be embargoed, please specify the part and the reason.

to Sven for many delicious dinners and great conversations. I'm also so grateful that my time in Belgium lead me to such wonderful people, **Nicky**, **Lilith**, **Kim,** and **An**, I am so glad I am able to call you friends. Thank you to **Alide**, **Ricardo**, **Eva,** and **Joris** for opening their home to me during my stay in Apeldoorn and including me in family dinners! Thank you to **Gladez** for making Vallee des Singes so much fun, for letting me have the bottom bunk and mostly for being the best and biggest DD! Thank you, Guillaume, for the years of hospitality in St. Etienne, for the dance shows, afternoon glasses of wine and sports in the living room! Thank you to everyone in the ENES lab – **Clementine**, **Avelyne**, **Alexis**, **Ingrid**, **Marie, Emilie, Laura,** and **Nicolas Boyer** – for the research chats and support and of course the occasional beer and sushi! I would also like to extend my warmest gratitude to **Colette Bouchet,** who spent many, many hours organising and managing my travel requests and liaising with the university administration on my behalf. And thank you to **Nora**, **Ben,** and **Ardea** for making me an honorary member of your wonderful little family. Thank you to my St. Andrews officemates and friends, **Lotty** and **Brittany**, for the venting and colouring! **Kate**, thank you for the many late nights of talking, the fantastic research conversations, advice and for being a great friend. And thank you **Lizzie** for being my constant cheerleader!

I am so lucky to have the most spectacular family – **Alessandra**, **Rachel**, **Jarah**, **Sarah**, **Susan**, **Ricky,** and **Mom**. They are the kindest, warmest, funniest, most loving people I know, and I am so grateful for each and every one of them. Thank you for all your support and understanding over the last four years.

There are two people in particular who have supported me so thoroughly during these rewarding, but very challenging, four years. **Aidan** thank you so very much from the bottom of my heart. Your patience, sincerity and love allow me to face every new day with confidence and joy. **Dad**, your support during this PhD has been nothing short of BRILLIANT! Thank you for the Monday phone calls, for the genuine interest in my research, the constant and helpful advice, the shoulder to lean on, and most importantly for "Oh, Statistics", a song that will likely never get out of my head!

TABLE OF CONTENTS

INTRODUCTION

I. INTRODUCTION

A. Information in communication

Communication is an inherently social event that serves to transmit information between individuals and reduce uncertainty in an unpredictable world (Seyfarth et al., 2010). The information conveyed in communicative signals can be as diverse as the intentional exchange of information in human language or as instinctive as a mongoose scent marking his territory. The diversity and complexity of information conveyed through communicative channels varies greatly, and much research has been devoted to discovering how animals communicate with one another, what information they transmit and what are the cognitive mechanisms driving their interactions.

Biological communication is studied by a wide and diverse range of disciplines, including, but not limited to, evolutionary biologists, linguists, neurologists, and comparative psychologists. In the simplest sense, and by an overall consensus, communication can be described as involving two actors. A *sender* produces a signal, which passes through the environment and is detected by a *receiver,* who then uses it to guide his/her actions or response (Bradbury & Vehrencamp, 2011). Outside of this basic definition, there is a general lack of consistency or agreement across disciplines on many of the tenants of communication, with the concept of information featuring prominently in many debates (Dall et al., 2005).

Definitions of information are far ranging and depend heavily on the field defining it. Many researchers inherently accept the concept of information within animal communication, and instead focus on whether said information is encoded and decoded in signals via language-like properties, such as semantics and reference (see Townsend & Manser, 2013; Zuberbuhler, 2003, 2005 for overviews on mammals and primates respectively). At the same time, others have argued that

information is a linguistically biased concept that should be abandoned in favour of studying how call structures function to stimulate or affect a receivers' nervous system and thus influence their response (Owren et al., 2010; Owren & Rendall, 2001; Rendall et al., 2009). And yet others still have suggested that the concept should be included, but cautiously (Scarantino, 2010) or only when embedded within an overarching adaptive perspective of communication (Scott-Philips, 2008).

It is overwhelmingly clear that if one is going to accept the concept of information, it is absolutely necessary to provide a comprehensible, working definition. As such, for this thesis communication is considered to consist of signals, and adopts the framework described by Maynard-Smith & Harper (2003, pp. 3) that: signals are defined as "any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved … It follows that the signal must carry information – about the state or future actions of the signaller or about the external world – that is of interest to the receiver."

In addition to this statement, this thesis adopts the classic concept that treats *information as a reduction in uncertainty* (Bergman & Sheehan, 2013; Dall et al., 2005; Scott-Philips, 2008; Seyfarth et al., 2010; Shannon & Weaver, 1949)*.* More precisely, "Information as a reduction of uncertainty in the recipient is useful because it connects communication to learning theory and to research on the mechanisms by which animals associate signals (or cues) with each other or with the outcomes of specific behaviours." (Seyfarth et al. 2010, pp. 4).

Information can be communicated through many different modalities, for example: the bold colouring of the strawberry poison frog (Siddiqi et al., 2004), the dance of the honey-bee (Dornhaus & Chittka, 2004), the scent marking of wolves (Sillero-Zubiri & Macdonald, 1998), or the gestures of apes (Hobaiter & Bynre 2014). While each modality has conditions in which it is the most effective

communication channel, the ubiquitous use of the auditory channel, in a variety of contexts across many species, has made vocal communication the focus of much research.

Vocalisations have the capacity to transmit a breadth of information about the social and ecological environment to listeners (Seyfarth et al., 2010); however, in direct social interactions the identity information encoded in vocalisations has a fundamental role. This thesis investigates multiple aspects of identity information in the vocal repertoire of the bonobo (*Pan paniscus*), a forest dwelling ape species that lives in a complex and flexible social environment. I first examine the identity information contained in a sender's signal, through a detailed investigation of five common call types within the bonobo vocal repertoire (**manuscript 1**). Following this, I investigate if social familiarity and relatedness can affect the acoustic properties that encode identity information in these five call types (**manuscript 2**). Lastly, I investigate individual recognition from the receiver's side by using playback experiments to test if bonobos can use vocal identity information (individual vocal signatures) to recognise familiar individuals (**manuscript 3**).

Before proceeding it is important to clarify a few terms that will be used interchangeably. Firstly, 'signaller', 'sender', 'caller', and 'vocaliser' all refer to the individual who is emitting a vocalisation. On the other hand, 'receiver' and 'listener' refer to the individual who extracts and potentially acts upon information from the emitted vocalisation. Lastly, the 'identity information' encoded in a signal can be referred to as such and additionally as 'individuality', 'identity signature', 'individual signature', or in the case of vocalisations, 'individual vocal signature'.

B. Vocal communication – the auditory modality

1. Sound production

In general terms, vocalisations are produced as air is moved over specialized structures in an animal's body that vibrate and generate sound waves (Bradbury & Vehrencamp, 2011). There is a great variety in the anatomy of sound production in the animal kingdom, and to provide an exhaustive review would not be appropriate here. As this thesis investigated vocal communication in the bonobo, I will thus provide a description of the general features and mechanisms of sound production in primates.
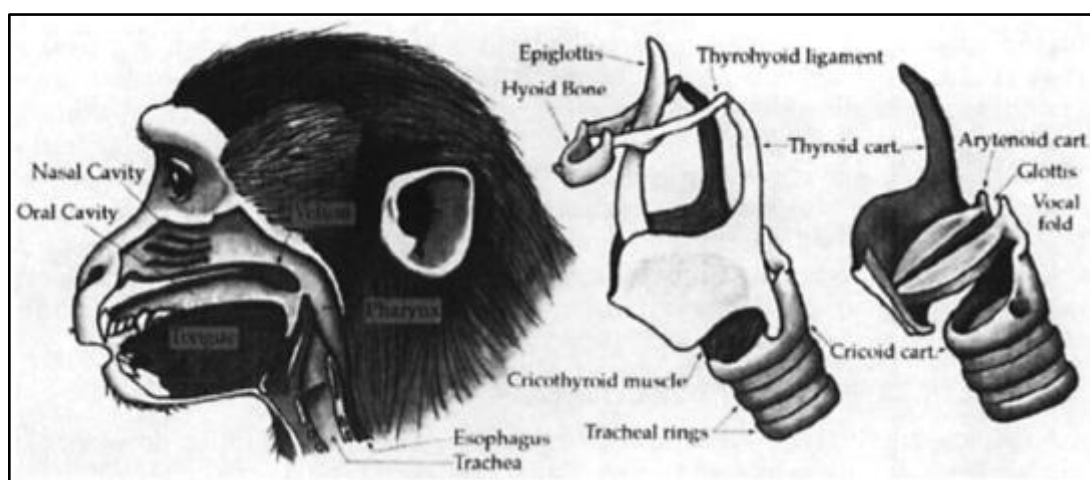
Figure 1. View of primate vocal tract (left) taken from Fitch & Hauser (1995), including images of the larynx (center) and the larynx with the thyroid cartilage removed to show internal vocal fold structures (right).

In all primates, including humans, vocalisations are produced when air from the lungs is pushed through the vocal folds of the larynx, causing them to vibrate and create sound waves, which are then filtered by the vocal tract above the larynx (Figure 1). Specifically, when air is expelled through the larynx, the vocal folds oscillate in a periodic fashion creating the fundamental frequency of the sound (Fitch & Hauser, 1995, 1998; Ganzafar & Rendall, 2008) (Figure 1 & 2). This

fundamental frequency (F₀) (sometimes referred to as pitch) is one of the strongest perceptual features of a call (Ganzafar & Rendall, 2008). The oscillation of the vocal chords also creates harmonics, which are overtones in integer multiples of the fundamental frequency (Fitch & Hauser, 1995; Ganzafar & Rendall, 2008) (Figure 2). These harmonics, while not as perceptual salient as the fundamental, give the sound a 'pure' tonal quality when the harmonics are weak and a 'buzzy' tone when the harmonics are strong (Fitch & Hauser, 1995). Manipulation of the vocal fold muscles and variation in air pressure and airflow are responsible for additional features of a sound wave, such as a duration and amplitude shape (Taylor & Reby, 2010). Additionally, certain irregularities in sound production can disrupt the periodicity of the wave and create nonlinear phenomena or 'noise' in a vocalisation, such as is observed in screams (Fitch & Hauser, 1995).
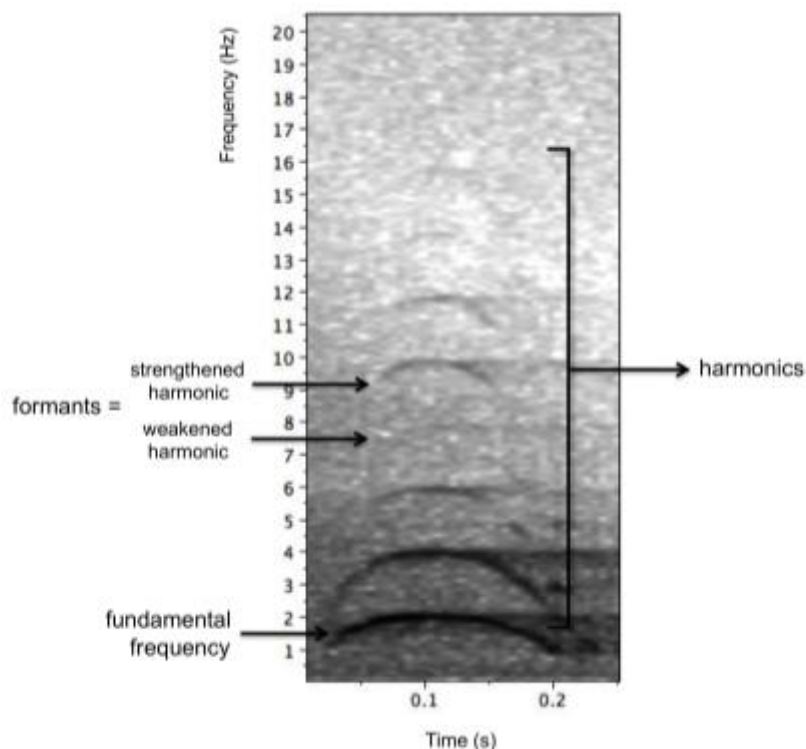
Figure 2. Spectrogram image of a bonobo vocalization with labels indicating fundamental, harmonics and formants.

As the sound wave created by the vocal folds moves outward it is filtered through the superlaryngeal vocal tract, which modifies sounds depending on its length and

shape (Figure 1). The energy at some frequencies can be weakened (attenuated) or strengthened by the resonant properties of the vocal tract walls creating formants, which are augmented harmonics (Fitch & Hauser, 1995; Ganzafar & Rendall, 2008) (Figure 2). Lastly, the movement of the articulators (tongue, lips, epiglottis) can further alter the sound as it leaves the body.

In order to learn where information about a signaller is encoded during sound production, many studies have investigated correlations between the acoustic properties of human and non-human primate vocalisations and features such as body-size, sex, age, and identity (see Ey et al., 2007 for a review in primates; Taylor & Reby (2010) for a more general review). Three main hypotheses have resulted, and often directly depend on which trait is being investigated: 1) that information about the signaller is dictated by the source of sound production (the larynx), 2) by the filter of sound production (vocal tract), or 3) by both the source and the filter. In non-human primates, identity information has been found to be encoded by both the source features, such as call duration, fundamental frequency, and fundamental contour (rhesus macaques – Rendell et al., 1998; baboons – Rendell et al., 2005; chimpanzees – Riede et al., 2004), and the filter features, such as formant frequencies (baboons – Owren, 1997; Rendall, 2003; rhesus macaques – Fitch, 1997; Rendell et al., 1998). It is therefore arguable that individual vocal signatures are created during sound production at both the source and the filter levels.

Once a vocalisation leaves the vocal tract and is emitted by a signaller it must pass through the environment before it can reach a receiver. As such, the environment can play a significant role in how information is propagated.

2. The environment

Vocalisations are vital for communication in environments where other modalities are limited or non-functional, and are particularly advantageous for long-distance communication, as sound can reach far greater distances than sight or smell. For many species, long distance communication is critical for group cohesion, mate attraction and to deter rivals, and vocalisations have been found to propagate from hundreds of metres in primates (Mitani & Stuht, 1998; Spillman et al., 2010; Wich & Nunn, 2002; Zimmerman, 1995), to several kilometres in elephants (Garstang et al., 1995; McComb et al., 2003) and lions (Grinnell & McComb, 2001; Pfefferle et al., 2007), to hundreds of kilometres in whales (Croll et al., 2002; Payne & Webb, 1971; Tyack & Clark, 2000). However, the environment can have strong impact on sound propagation – in general, sound travels better in open plains than in forests, in deep water than shallow, and certain frequencies are affected by heat and humidity (Bradbury & Vehrencamp, 2011; Catchpole & Slater, 2008; Forrest, 1994; Larom et al., 1997; Mercado & Fraser, 1999). Therefore, the vocal behaviour and acoustic repertoires of many species are specifically adapted to transmit information in their native environments, despite challenging communicative contexts, such as in dense forests (Brown et al., 1995; Waser & Brown, 1986; Wich & Nunn, 2002), underwater (Bass & Clark, 2003; Tyack & Clark, 2000), in the dark (Jones & Holderied, 2007; Schnitzler et al., 2003) or in fluctuating atmospheric conditions (Larom et al., 1997).

The dense forest canopy and undergrowth of the bonobo's natural habitat can degrade sounds through reverberation off foliage, filter out certain sound frequencies, or simply absorb the sound wave reducing its propagation distance (Catchpole & Slater, 2008; Morton, 1975; Richards & Wiley, 1980). However, the thick foliage can have an even more drastic effect on visibility and can reduce clear sight to only a few meters; while the bonobo loud call is thought to propagate through this environment up to a distance of 500 meters (Hohmann & Fruth, 1994), making it a vastly more effective and flexible communication channel.

C. The sender – Identity encoded in acoustic signals

Much of the research on identity signalling has focused on the receiver and how individualised features enable recognition, while considerably less research has focused on the sender and the information encoded in the signal (Tibbetts & Dale, 2007). As such, this thesis endeavours to create a balance between the sender and receiver perspectives by researching both the coding of information in vocal signals (manuscript 1 & 2) and the decoding of the identity information that facilitates individual recognition (manuscript 3). In the following section I will discuss the first half of the communicative equation – the sender.

1. Information encoded in acoustic signals

a. Variation in information

Vocalisations are an integral aspect of the social lives of many species and can transmit a wealth of information to listeners about the environment and the individual calling (Seyfarth & Cheney, 2003). Individuals can gain information from vocalisations when a signal is consistently used in a specific context, which enables listeners to associate the signal with the event and thus allows receivers to make predictions about a signaller's behaviour or an upcoming event (Seyfarth et al., 2010). In recent years, much research has been devoted to how and if vocalisations function by referring to specific environmental events. Evidence from a number of divergent species, such as chickadees (Baker & Becker, 2002; Templeton et al., 2005), chickens (Evans et al., 1993), great tits (Suzuki, 2013; 2015), meerkats (Manser, 2001), prairie dogs (Kiriazis & Slobodchikoff, 2006), Siberian jays (Greisser, 2008), and many primate species (for review see – Zuberbuhler, 2003) have shown that alarm calls not only inform listeners to the presence of a predator, but can encode more specific information, such as the type of predator, size of predator and/or urgency of the threat. In addition to this,

some species have shown that their vocalisations or vocal sequences can functionally refer to specific food types, e.g. chimpanzees (Slocombe & Zuberbuhler 2005, 2006), fowl (Evans & Evans, 2007), ravens (Bugnyar et al., 2001), marmosets (Kitzmann & Caine, 2009), and bonobos (Clay & Zuberbuhler 2009, 2011a). (It should be mentioned that there has been considerable debate regarding the concept of functionally referential signals and its validity within the study of non-human animals (see Scarantino & Clay, 2015 and Wheeler & Fischer, 2012, 2015 for arguments), however as it is not the focus of this thesis the arguments will not be expounded here.)

For most species, the majority of vocal communication serves to mediate social interactions by providing information to listeners about a signaller's current state, forthcoming action or stable characteristics (Bergman & Sheehan, 2013; Bradbury & Verhrencamp, 2011; Catchpole & Slater, 2008; Maynard Smith & Harper, 2003; Searcy & Nowicki, 2010). Research has shown that vocalisations can provide listeners with information as to the signaller's intended action or desired consequence, such as interest in mating, aggressive action, reconciliation, preferred moving direction, and desire to be groomed or given food/item (e.g. baboons – Cheney et al., 1995; Semple, et al., 2002; elephants – Poole et al., 1988; gibbons - Cowlinshaw, 1996; macaques – Bauers, 1993; sparrows – Searcy et al., 2014). Vocalisations can inform on unstable states, such as an individual's location, fertility, fighting ability, or reproductive quality (e.g. baboon – Fischer et al., 2004; bison – Wyman, 2012; deer – Briefer, et al., 2010; Reby & McComb, 2003; skuas – Janicke et al., 2008), as well as on more stable individual characteristics, such as, sex, age, size, dominance status, or territory (e.g. baboons – Bergman et al., 2003; deer – Reby & McComb, 2003; horses – Lemasson et al., 2009; hyenas – Mathevon et al., 2010; rhesus macaques – Fitch, 1997).

b. Identity information

For social living species, the identity information encoded in vocalisations is critical in communicative interactions. Individual signatures are any trait or traits that vary between individuals but are stable and reliable within an individual over time, and thus allow for consistent discrimination between individuals. (Beecher, 1989; Bergman & Sheenhan, 2013; Tibbetts & Dale, 2007). For vocalisations, individuality is created by individual differences in vocal anatomy, whether at the source or the filter or both, that cause audible differences in sound production (as described in section 1 – Sound production). When the variability of the acoustic features of species-specific call types is greater between individuals than from within a single individual, the calls are considered to be individually distinct (Beecher, 1989).

A huge variety of species have demonstrated individual vocal signatures, including: bats (Voight-Heucke et al., 2010), hyenas (Mathevon et al., 2010), seals (Insley, 2000; Charrier et al., 2003), deer (Vannoni & McElligott, 2007), elephants (Soltis et al., 2005), marmots (Matrosova et al., 2011), and a number of bird (e.g. Aubin & Jouventin, 2002; Catchpole & Slater, 2008; Charrier et al., 2001; Dentressangle et al., 2012; Kondo et al., 2010; Vignal et al., 2004) and non-human primate species (eg. Chapman & Weary, 1990; Fischer et al., 2001; Levréro & Mathevon, 2013; Price et al., 2009; Wich et al., 2003).

Identity information facilitates individual recognition which allows differentiation between social partners in many contexts, such as in dominance hierarchies, territorial defence, mating and parent-offspring identification, and group cohesion and coordination, all of which are vital for individual, group, and even species fitness and survival (Sheehan & Tibbits, 2009; Tibbetts & Dale, 2007). Identity signatures also allow individuals to accumulate social knowledge about conspecifics following social interactions and associate this knowledge with the stable features of an individual's voice (Bergman & Sheehan, 2013). This

association between social knowledge and identity signatures can greatly increase an animal's fitness, as it facilitates adaptively constructive social decisions based on current context and previous experience.

2. Influences on acoustic signals

a. Evolution of individuality in vocal repertoires

While an individual's unique vocal anatomy can create individual vocal signatures, the external world can also have a significant impact on vocalisations, whether by temporary changes in acoustic structure or long-term evolutionary changes on the whole of a species vocal repertoire. Environmental conditions are a prime example of this, as they can greatly impact sound propagation and affect the reliability of information transmission. For decades, much research has explored how the environment affects sound propagation (Forrest, 1994; Marten & Marler, 1977; Mercado & Fraser, 1999; Morton, 1975; Mouterde et al., 2014; Neuweiler, 1983) and how species' vocal repertoires have adapted to suit their habitats. While there is an overwhelming consensus that the selective forces imposed by a species' environmental habitat have the potential to shape its overall vocal repertoire, there is less agreement on the importance of the environment's impact on individual vocal signatures specifically.

Some researchers have argued that it is the social, and not physical, environment that plays the most important role in shaping the identity information in vocalisations. Pollard & Blumstein (2011) found that in sciurid rodents (squirrels, marmots and prairie dogs specifically) individuals living in large social groups had more signature information encoded in their calls than individuals living in smaller social groups. In addition to this, Bouchet *et al*. (2013) found that vocal variation and individual distinctiveness increased in tandem with the complexity of the social systems of three separate monkey species. These evidences are bolstered by the number of large group living species whose mother-offspring recognition is

highly dependent on strong vocal signatures in the mother, offspring or both (Briefer et al., 2012; Charrier et al., 2001, 2003, 2010; Knornschild & von Helversen, 2008; Mathevon et al., 2003).

The majority of research has approached this topic by investigating identity information in the repertoire of a single species, and multiple theories have emerged as to which calls within a species repertoire should contain the strongest identity information, with two primary hypotheses – the 'distance communication hypothesis' and the 'social function hypothesis'. The distance communication hypothesis suggests that vocalisations used for long-distance communication and/or in low visibility environments should have adapted particularly strong individual vocal signatures to counteract signal degradation, and evidence supporting this has been found in a variety of species (Chapman & Weary, 1990; Lemasson, et al., 2009; McComb et al., 2003; Mitani, et al., 1996; Wich et al., 2003). On the other hand, the 'social function hypothesis' suggests that call types used in direct social interactions should have the strongest identity information, and has found support in species that are generally spatially cohesive (Bouchet et al., 2012; Charrier et al., 2001; Lemasson & Hausberger, 2011; Price et al., 2009; Snowdon et al., 1997).

b. Social influence on acoustic structure

While much research has been devoted to investigating how and why evolution favoured identity signatures in the vocalisations of many species, other research has focused on what impacts an individual's voice during his/her lifetime. Some research has found that the vocalisations of genetically related individuals are more acoustically similar than those of unrelated individuals (e.g. Blumstein et al., 2013; Deecke et al., 2010; Levrero et al., 2015). Vocal similarity between related individuals has been suggested to provide a mechanism for kin recognition, enabling increased cooperative behaviour and altruism between related individuals (Silk, 2002). As vocal signatures are created by vocal anatomy, which

is dictated by an individual's genetics, these findings are not unexpected. However, for species, such as bonobos, who live in social groups characterised by a mix of related and unrelated individuals, recognition, affiliation, and cooperation between individuals extends beyond relatives.

In recent years, a number of studies have focused on how sociality, in particular social learning, affects vocal communication. Evidence of vocal production learning, in which vocal signals are changed as a direct result social experience (Janik & Slater, 2000), has been found in an extensive number of bird species (Beecher & Brenowitz, 2005; Brown & Farabaugh, 1997; Catchpole & Slater, 2008) and a limited number of mammal species (Janik & Slater, 1997), namely bats (Boughman, 1998), cetaceans (Foote et al., 2006), elephants (Poole, et al., 2005) and pinnepeds (Reichmuth & Casey; 2014). A prime example of this is the song of some bird (Catchpole & Slater, 2008) and whale (Janik, 2014; Rendell & Whitehead, 2001) species, which change on a seasonal basis due to social experience.

While some species show changes in vocal structure over time, others demonstrate extreme flexibility in their vocalisations. Species such as parrots (Bradbury & Balsby, 2016) seals (Reichmuth & Casey; 2014) and dolphins (King et al., 2013) are known to be able to imitate a range of sounds and even immediately incorporate them into their individual repertoire. In addition to this, many of these same species have demonstrated group vocal dialects (Henry et al., 2015), which occur when the acoustic features of species specific calls converge between group members when compared to non-group members (e.g. bats – Boughman, 1997; cetaceans – Deecke et al., 2000; birds – Podos & Warren, 2007). While the precise function of these dialects is debated, and may vary between species, they are a clear indication that an individual's social life, along with their anatomy, can have a profound effect on acoustic structure.

Despite years of intensive research, the repertoires of non-human primate species have rarely demonstrated a high degree of vocal flexibility either during development or throughout adulthood. Early studies with chimpanzees attempted to teach them novel sounds, namely human words, were marked by failure, despite the chimpanzees' apparent understanding of many of the words they were incapable of reproducing them (Fitch, 2010; Hayes, 1951; Kellogg & Kellogg, 1933; Premack & Premack, 1972). Further studies with squirrel monkeys and macaques found that the monkeys developed species typical vocal repertoires despite sensory (in the form of deafening) or social deprivation (in the form of cross-fostering) (Hammerschmidt et al., 2001; Newman & Symmes, 1982; Owren et al., 1992; Winter et al., 1973). Therefore, non-human primates' historic inability to imitate sound, and tendency to develop normal vocal repertoires despite deficient situations, led to the general consensus that their vocal repertoires are innate and fixed, and any social communicative learning would be in the form of learning when to use specific vocal signals (contextual usage learning) and not in learning or modifying their acoustic features (vocal production learning) (Egnor & Hauser, 2004; Janik & Slater, 2000).

However, these ideas have been challenged by increasing evidence that non-human primate vocalisations are subject to acoustic flexibility and vocal production learning in specific social situations, albeit limited in comparison to some bird and cetacean species (Hammerschmidt & Fischer, 2008; Snowdon, 2009). Several studies have now demonstrated that, in some species, individuals converge the acoustic features of their calls with closely affiliated group members or during vocal chorusing with other individuals (e.g. Diana monkeys – Candiotti et al., 2012; Campbell's monkeys – Lemasson et al., 2011; Chimpanzees – Mitani & Brandt 1994; Mitani & Gros-Louis, 1998; Japanese macaques – Suguira, 1998). In addition to this, many species have shown evidence of vocal dialects, suggesting that the vocal features of individuals within an entire group can converge (e.g. chimpanzees – Crockford et al., 2004; macaques – Fischer et al., 1998; Tanaka et al., 2006; marmosets – De La Torre & Snowden, 2009), and a

couple studies have actively observed this vocal convergence, and resulting group dialect, after the merging of different populations (Elowson & Snowdon, 1994; Watson et al., 2015).

These types of changes in the acoustic structure of individual's vocalisations can potentially change the information being communicated or add additional information to a call. In either case, it is vital to explore and understand if the acoustic features of vocalisations that encode identity information, and which many species rely on for individual recognition, can be modified by the social environment.

D. The receiver – Identity information for individual recognition

As we have explored identity information from the sender perspective, the following section will explore identity information from the receiver side, through an investigation of individual recognition.

1. Individual recognition

During social interactions individuals are often classified by conspecifics on the basis of a number of characteristics, such as sex, rank, group membership, kinship, or familiarity (Mateo, 2004; Tibbits & Dale, 2007). This type of categorisation allows for quick social decision-making and differential treatment of conspecifics often with only limited social information. However, for species that have complex social networks characterised by repeated interaction with familiar individuals, such as bonobos, this type of categorisation may need to be supplemented with specific social knowledge of an individual.

Individual recognition occurs when a signaller is identified according to their unique traits by a receiver – a receiver who has associated these traits with specific social knowledge of the signaller, which can be utilised in further social interactions (Bradbury & Vehrencamp, 2011; Tibbetts & Dale, 2007). In general, individual recognition occurs in many species across multiple modalities and has been demonstrated in species as diverse as wasps (visual recognition – Sheehan & Tibbetts, 2011) and lemurs (scent recognition – Palagi & Dapporto, 2006). However, the transmission capabilities of vocalisations across long distances and in a number of complex environments make individual vocal signatures a crucial recognition tool for a number of species (Shapiro, 2010; Taylor & Reby, 2010).

Individual vocal recognition is commonly used by many species for territorial defence, where individuals react either more or less aggressively depending on whether they recognise their neighbours, the 'dear enemy' concept (Breifer et al., 2010; Catchpole & Slater, 2008; Temeles, 1994). Individual vocal recognition has shown to be vital for parent-offspring recognition in species where offspring are regularly separated from the parent, such as in seals (Charrier et al., 2001) and penguins (Aubin & Jouventin, 2002) that use vocalisations to locate their offspring in heavily populated colonies. Within non-human primates' individual vocal recognition can function in a broad array of contexts, including territorial or resource defence (Briseno-Jaramillo et al., 2014; Mitani, 1990), dominance maintenance (Slocombe et al., 2010), parent-offspring recognition (Levrero & Mathevon, 2013) and to mediate social interactions (Bouchet et al., 2012), and, as in other orders, the most important of which will depend on a species' particular social structure and habitat (Shapiro, 2010). Many primate species, including bonobos, live in groups where familiar individuals are often separated from one another through the nature of their environment, group structure or feeding dynamics. Consequently, individual vocal recognition is particularly valuable in these species to maintain contact between group members (i.e. group cohesion) and to coordinate group movements (Leliveld, et al., 2011; Mitani et al., 1996;

Rasoloharijaona et al., 2006; Rendall et al., 1996; Santorelli et al., 2013; Waser, 1977)

2. Long-term vocal recognition

Beyond the effects that the environment can have on signaller and the acoustic properties of vocalisations that code identity information, the external world can also impact the receivers' ability to decode identity information. Much research has been devoted to how distance or environmental factors can impact recognition ability and many have demonstrated that individuals can still recognise familiar conspecifics even after considerable signal degradation (e.g. Charlton et al., 2012; Mouterde et al., 2014; Rek & Osiejuk, 2011). Other studies have experimentally investigated individual recognition after manual distortion of specific features in a vocalisation (eg. Charrier et al., 2003; Jouventin et al., 1999; Searby & Jouventin, 2003; Vignal et al., 2008), and found recognition is possible when certain features are distorted but not others.

A limited number of studies have investigated the effect that long-term separation between individuals has on vocal recognition. These studies have demonstrated that some species are capable of remembering former social partners after being separated for a number of months (hooded warbers – Godard, 1991) or years (fur seals – Insley, 2000; Campbell's monkeys – Lemasson & Hausberger, 2003; cotton-top tamarins – Matthews & Snowdon, 2011; ravens – Boeckle & Bugnyar, 2012), and in dolphins for at least a remarkable 20 years (Bruck, 2013).

The multifaceted social lives and impressive cognitive abilities of apes make it reasonable to hypothesise that this taxonomic family have the capacity for long-term recognition of past social partners. This is particularly true for bonobos who display similar fusion-fission social structures as two other mammal species, dolphins (Bruck, 2013) and elephants (McComb et al., 2000), that have shown long-term vocal recognition. Despite this, to date there have been no studies

investigating study long-term vocal recognition in any ape species. As such, the final thesis of this manuscript provides unique insight into a facet of the communicative capability of our closest living relatives.

E. The model – *Pan paniscus*

Figure 3. A mixed sex and age group of bonobos at Planckendael. From left to right: Lingoye, Nayoki (infant in background), Busira, Djanoa (adult in background), Lina and Louisoko (in foreground, bottom-right).

In 1933 Harold Coolidge described a rare species of chimpanzee that was known to live on the south bank of the Congo River (Coolidge, 1933). Despite his arguments, and by many others in the years following, that this 'pygmy chimpanzee' was in fact a separate species entirely, the bonobo continued to be misclassified and subsequently understudied when compared to the common chimpanzee (*Pan trogolodytes*) (Kano, 1992). In the last few decades this has fortunately changed, however the previous years of scientific neglect have left researchers with a number of unanswered questions. We now know that the bonobo and the chimpanzee are together human's closest living relatives, both

sharing 99.4% of our genetic composition (Prüfer et al., 2012; Wildman et al., 2003). As such, study of the bonobo can help to understand this remarkably unique species, which can improve and focus conservation efforts, and provide an invaluable comparative tool to explore how we as humans evolved.

In the following section I will provide some general information on the bonobo, providing more detail when relevant for this thesis, and conclude with a brief discussion on the importance of the bonobo as a study species.

1. Morphology, habitat, diet and life history

Bonobos are the smallest of the four great ape species, largely resembling chimpanzees but with a few key, discernible differences, most notably of which is their slender build. They are the least sexually dimorphic ape species in height, with males measuring an average of 1190 mm and females 1170 mm (Coolidge & Shea, 1982) and are slightly more dimorphic in weight, with males averaging around 44 kg (45.0 ± 8.4 kg – Jungers & Susman, 1984; 43.43 ± 6.4 kg – Parish, 1996) and females around 34 kg (33.2 ± 4.2 kg – Jungers & Susman, 1984; 35.85 ± 5.9 kg – Parish, 1996). Like chimpanzees they have dark hair covering most of their bodies, but unlike chimpanzees have almost a completely black face with the exception of pale lips (Figure 3).

Bonobos are native to only to the Democratic Republic of Congo, in equatorial forests bordered by the rivers, Congo, Lualuba, Kasai and Sankuru (Idani et al., 2008; Kano, 1992). The bonobo habitat is primarily dense tropical rainforests, mainly characterised by primary dry and swamp forests, as well as some secondary forests, or disturbed forests, around human villages (Hashimoto et al., 1998; Kano, 1992), and southern populations may also live in forest/savanna mosaic habitats (Myers-Thompson, 2002). The phenology of the various plants in the bonobo home range is diverse and complex, with different plant species fruiting and flowering at various times throughout the year (Idani et al., 1994;

Kano, 1992; Pennec et al., 2016; White, 1998). As such, the equatorial forest provides bonobos with steady and abundant sources of food, which arguably has had a profound impact on their social structure (White, 1996, 1998). Bonobos consume a huge variety of plant material consisting of fruits, leaves, and terrestrial herbaceous vegetation, but also seeds, shoots, bark, roots, mushrooms and honey (Bermejo, 1995; Idani, 1994; Kano, 1992; Serckx et al., 2015; White, 1998). Bonobos are also known to consume animal protein in the form of small mammals, insects and eggs (Bermejo, 1995; Kano, 1992), and recent studies have found that some bonobo populations also hunt monkeys and duikers (Fruth & Hohmann, 2002; Hohmann & Fruth, 2008; Surbeck & Hohmann, 2008).

Bonobos, like all ape species, have an extended period of maturation. Infants gain physical independence from their mothers from between six months to a year after birth, and even then, rarely range further than a few metres from their mother for their first couple years (Kano, 1992; Kuroda, 1989). Weaning age is variable and depends on the interbirth interval of the mother, however bonobos are generally weaned between four and five years old (de Lathouwers & Van Elsacker, 2006; Kuroda, 1989). Sexual maturity usually occurs earlier in females, between the ages of 6–10, while males typically begin to show a sharp increase in urinary testosterone around the age of eight (Behringer et al., 2014). Females disperse from their natal group at sexual maturity, while males generally remain with their mothers for life (Furuchi, 1989; Kano, 1992).

2. Social structure

The social life of the bonobo is one of the most unique and interesting among primates and substantially differentiates them from the closely related chimpanzee. Bonobos live in multi-female, multi-male groups that are characterised by female dominance and fission-fusion dynamics. Females are the backbone of bonobo society and maintain significant relationships with both sexes (Furuchi, 2011; Parish, 1996; Stevens et al., 2006; White 1996). Females form

strong affiliative relationships with each other regardless of relatedness, and provide regular coalitionary support to one another (Furuchi, 1989; Parish, 1994; Stevens et al., 2006; Tokuyama & Furuichi, 2016). The strong bonds that occur between unrelated female bonobos are unique among primates (White, 1996) and rare in animal kingdom overall. However, bonobos also cultivate and sustain close mixed sex relationships between related and unrelated individuals, and philopatric males maintain strong bonds with their mothers throughout adult life (Furuchi, 1989, 1997; Hohmann, et al., 1999; Stevens et al., 2006). In general, both in captivity and in the wild, male-male relationships are found to be comparatively weak (Furuchi and Ihobe, 1994; Kano, 1992; Stevens et al., 2006). This is in direct contrast to the profoundly complex and strong affiliative and agnostic relationships between males in chimpanzee social groups (Newton-Fischer, 2002).

Dominance within bonobo society is dynamic, and there is mixed evidence as to whether hierarchies exist in a strictly linear fashion or if it is more dependent on context and available support (Paoli et al. 2006; Paoli & Palagi, 2008; Stevens et al., 2005, 2008; Vervaecke et al., 2000a). It is clear, however, that females occupy the highest ranking positions (Paoli & Palagi, 2008; Stevens et al., 2008; Surbeck & Hohmann, 2013). Additionally, the close network of female alliances allows them to form all-female coalitions, in which aggression is generally directed toward male group members (Vervaecke et al., 2000b; Tokuyama & Furuichi, 2016). This coalitionary support is arguably what allows females to maintain consistent dominance over males and preferential access to resources (Vervaecke et al., 2000b; White, 1996). Males can move up the dominance hierarchy if their mothers are high ranking within the group and, although perhaps less reliably, via close social bonding with unrelated females (Furuchi, 1997; Hohmann et al., 1999; Hohmann & Fruth, 2003; White, 1996). Therefore, aside from alpha positions, which are always occupied by females, dominance is mixed across the sexes and males can dominate females in certain contexts (Stevens et al., 2007; Surbeck & Hohmann, 2013).

Bonobo society is comprised of large communities of tens of individuals with loose and often overlapping home ranges (Hashimoto et al., 1998). Within a community, groups of individuals regularly separate from one another, so-called fission, into small cohesive parties or subgroups (anywhere from 5-25 individuals reported, excluding infants) to forage and feed (Kano, 1982; White, 1998; Idani, 1991; Mulavwa et al., 2008). Unlike chimpanzees, who also demonstrate fission-fusion dynamics, bonobo parties consist of individuals of both sexes and all age classes. The composition of parties can change throughout the day as individuals join and leave, with fissions normally being accompanied by extensive vocal exchanges between parties (White, 1998) and dispersed foraging parties regularly communicate with other community members via long-distance vocalisations (Hohmann & Fruth, 1994, 2002; Furuchi, 2011).

When compared to chimpanzees, bonobos are considered to be relatively tolerant and considerably less aggressive, including toward individuals of other communities (Boesch at al., 2002; Wilson et al., 2014). Inter-community encounters almost always begin with an exchange of vocalisations and are generally followed by complete avoidance or temporary social interaction (Furuchi, 2011). Social interactions between two communities are often comprised of mild forms of aggression and signs of stress, such as increased directed vocal and gestural displays between individuals, but rarely involve physical aggression and fatal inter-group encounters have never been conclusively recorded (Hohmann & Fruth, 1994, 2002; Furuchi, 2011, Wilson et al., 2014). In a few cases, affiliative interactions have been observed between communities, with the groups even temporarily feeding or nesting together (Furuchi, 2011).

Of all the unique features of bonobos perhaps the most widely known aspect of bonobo society is their sexual behaviour. Bonobos are known to use sexual interaction extensively beyond reproductive purposes, to moderate social relationships between individuals by diffusing tension, increasing bonding and

facilitating reconciliation (Idani, 1991; Fruth & Hohmann, 2006; Furuchi, 1989; White, 1996). In addition to this, bonobos engage in sexual behaviours, such as mounting, genital-genital rubbing and rump-rump contact in almost all age-sex combinations (Hashimoto, 1997; Hohmann & Fruth, 2000).

3. Communication

Despite the increasing interest in bonobos over the last few decades, the majority of studies have focused on their social structure, socio-ecology and cognition. However, in recent years a limited number of studies have focused on bonobo communication. Communication in bonobos is complex and often comprised of multimodal and context dependent signalling (Clay et al., 2015; Genty et al., 2014, 2015; Pollick & de Waal, 2007; Pollick et al., 2008). While facial and olfactory communication have received relatively little attention in bonobos, there have been a number of studies investigating their gestures (de Waal, 1988; Pika et al., 2005; Pollick & de Waal, 2007). Recent studies have demonstrated considerable cognitive complexity and flexibility in bonobo gestures, with evidence of intentional, goal-directed signalling and spatial reference (Genty & Zuberbuhler, 2014; Genty et al., 2015; Halina et al., 2013; Pika & Zuberbuhler, 2008).

Along with their gestural communication, bonobos are a highly vocal species whose varied vocalisations accompany many of their social behaviours. Unlike other modalities, vocalisations are a particularly advantageous as they are functional in both close and long-range communication. Despite some early descriptions of the bonobo vocal repertoire (de Waal, 1988; Bermejo & Omedes, 1999), the majority of notable research on bonobo vocal communication has been produced in recent years by Z. Clay and colleagues. Their work has provided greater insight into vocal communicative function via quantitative analyses and playback experiments. Studies, primarily focused on their vocal behaviour during feeding and sex, have found that: bonobos can extract meaningful information about food items from sequences of calls (Clay & Zuberbuhler, 2009, 2011a);

some call types can be used in functionally flexible ways (Clay *et al*., 2015); calls emitted during reproductive versus non-reproductive sexual encounters do not significantly differ from one another, a rare occurrence in primate vocal behaviour (Clay & Zuberbuhler, 2011b); and, importantly, that calls emitted during sexual encounters provide cues to individual identity (Clay & Zuberbuhler, 2011b).

a. Vocal repertoire

Two primary studies have attempted to describe the entirety of the bonobo vocal repertoire (de Waal, 1988; Bermejo & Omedes, 1999), and neither provided a quantitative analysis of call types' acoustic features or contextual. Nevertheless, both studies described the bonobo vocal repertoire as being highly graded and generally similar in form and function to the chimpanzee repertoire, but with notably higher pitched vocalisations. Working in captivity, de Waal (1988) described a dozen main call types, however he also described multiple sub-types which could differ slightly in form or use, including: hoots (staccato high-hoot, legato high-hoot, contest hoot, low hoot), peeps (food peep, alarm peep, peep-yelp), barks (wieew bark, whistle bark), grunts, moans, laughs and screams. Bermejo & Omedes, 1999, working in the wild, naturally found many similarities to de Waal's descriptions, but described an additional three call types, often provided different names for call types, and described a greater number of sub-types. In describing and classifying the bonobo vocalisations collected, I considered the call type descriptions of de Waal (1988) and Bermejo & Omedes (1999), and attempted to reconcile them when possible (Table 1). Additional guidance on call type structure and use was also taken from more recent studies, namely Clay & Zuberbuhler (2009). The Material and Methods section of this thesis will provide more specific detail and spectrographic images on each call type collected and used for analyses.

Call type	Call variants (Sub-types)	Contextual use[*]
Alarm peep	-	Alarm, disturbances
Bark	Wieww bark; Composed bark	Agonistic interactions, display, alarm, social excitement, inter-party communication, feeding excitement
Grunts	Grunts; greeting grunts	General feeding, submissive greeting, reconciliation
Hiccup	-	General feeding, grooming; play
High-hoot	Staccato; legato	Agonistic interactions, display, sexual interaction, nesting, social excitement, inter-party communication
Laugh	Pant laugh; croak	Play
Low hoot	-	Agonistic interactions, display, disturbances
Peep	-	General feeding, grooming, play
Peep-yelp	-	General feeding, grooming, play
Pout moan		Appeasement, begging, grooming
Scream	Full scream; scream-whistle; bark-scream; peep scream; rasp scream	Agonistic interactions, sexual interactions
Soft bark	-	General feeding, foraging.
Whistle	Whistle; Whistle-bark; whine-whistle	Agonistic interactions, display, social excitement, food excitement, general feeding, inter-party communication

Table 1. Vocal repertoire. Table 1 describes the bonobo vocal repertoire based on the descriptions of de Waal (1988) and Bermejo & Omedes (1999). As there were differences between these two studies, this table does not precisely replicate either study but instead attempts to comprehensive compile and reconcile them. *Not all call variants or sub-types will be used in all of the contexts listed for a given call type.

This thesis focused on the acoustic structure of individual calls of multiple call types. However, it should be mentioned that some call types are regularly emitted as one discrete unit, while others are produced as stereotyped sequences of call units, and a number of them are flexible, emitted either individually or in sequence. The high-hoot call type is the most restricted to sequences, regularly starting with a build-up of low hooting, barks or legato high-hoots and followed by a series of high-hoots that increase in decibel intensity and decrease in time between calls until they reach a crescendo, after which intensity and emission rate slowly decrease (Figure 4).
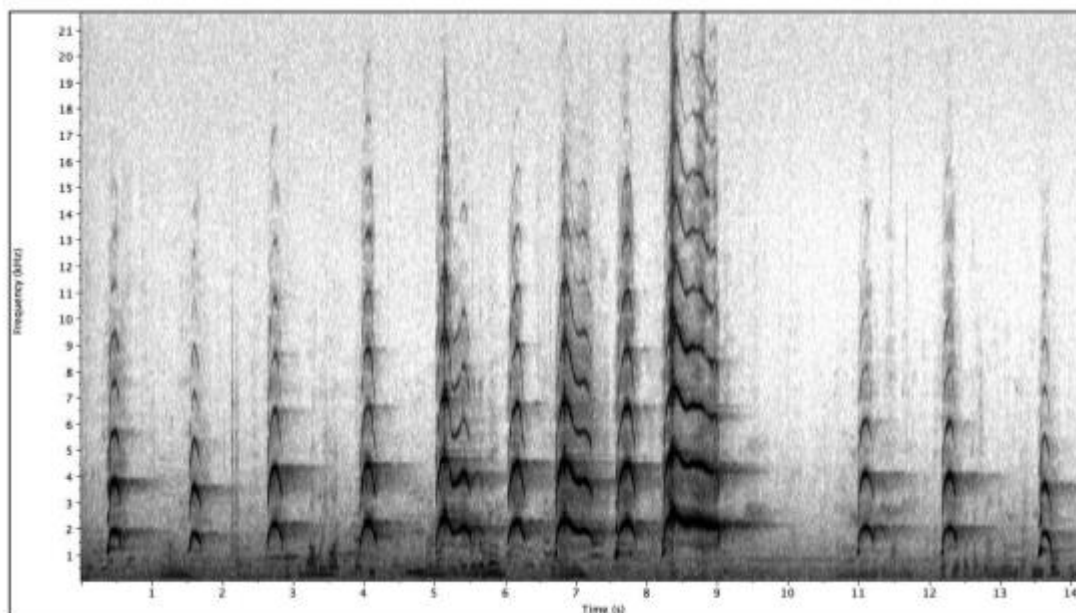
Figure 4. Vocal sequence. This figure shows the build up of a high-hoot sequence emitted by an adult male during feeding.

4. Why study bonobos?

Bonobos are separated from the other African ape species by the Congo River. It is debated precisely how the bonobo speciation occurred, however it is thought that somewhere in the Pleistocene era (between 2.0 and 0.8 million years ago) a group of the proto-*Pan* species became segregated by the Congo River from the rest of the population (Takemoto et al., 2015). This division caused the proto-*Pan* group to split and eventually evolve into two separate species, the chimpanzee and the bonobo. Despite the extreme genetic similarity between the two species, due to the relatively recent speciation, research has uncovered some remarkable distinctions in behaviour and society between bonobos and chimpanzees (Boesch, 2002). However, when compared to the breadth of study on and knowledge of the chimpanzee, bonobos have been largely understudied until very recently. As such, it is vastly important to continue to focus research on the bonobo, fill in the blanks and create a comprehensive understanding of the similarities and differences between the two sister-species. Only then will we clearly be able to see through this rare evolutionary window and learn how and which selective pressures created the two unique *Pan* species we see today. In addition to this, as bonobos and chimpanzees together are human's closest living relatives, understanding them equally is vital to creating a holistic model of human evolution.

F. Research Objectives

In this thesis, I aspired to contribute to the overall understanding of the vocal communication system of one of our closest living relatives, the bonobo. Specifically, I focused on the dynamics of identity information across the bonobo vocal repertoire and whether this encoded information can be used for individual recognition.

1. Individual vocal signatures and the effect of call type – manuscript 1

The first study of this thesis provides the first quantitative investigation of individual vocal signatures across multiple call types of the bonobo vocal repertoire. By studying five common call types, I was able to: 1) investigate the strength of identity information and its stability across a large portion of the repertoire; and 2) as the call types have different functions, explore which selective pressures likely affected the evolution of individual vocal signatures. Multiple statistical procedures were performed to address these questions, which allowed for a robust and comprehensive view of individuality in the bonobo repertoire. Employing multiple and varied techniques also allowed for direct comparison to similar studies on different species, and thus enabled robust hypotheses as to why evolution favoured higher levels individuality in the vocalisations used certain contexts over others.

As much of the previous research on bonobo vocalisations has relied on qualitative descriptions of the vocal repertoire, this study additionally provides a quantitative, statistical investigation of the acoustic differences between these five common call types and their different contextual uses.

2. The influence of familiarity and relatedness on call structure - manuscript 2

Following the first study, I was interested to see if the strong social bonds found in bonobos could influence the acoustic features that code individual information. The primary aim of this second manuscript was to investigate if individuals living in the same group showed acoustic convergence in their calls when compared to pairs living apart. In doing so, I also controlled for relatedness between individuals, resulting in a secondary investigation of the effect of relatedness on vocal similarity. In total, this study compared vocal similarity in four types of pairs of individuals: related pairs who also live in the same group; related pairs who do not live in the same group; unrelated pairs who live in the same group; and unrelated pairs who do not live in the same group.

The complexity of non-human primate communication, particularly as it relates to the highly complex vocal communication systems of other species such as humans and birds, has long been an intense area of study. With this study, I endeavoured to contribute to this debate by linking evidence of vocal similarity to vocal flexibility and vocal production learning.

3. Vocal recognition of familiar individuals – manuscript 3

The final manuscript of this thesis investigated if bonobos were capable of using the identity information encoded in vocalisations to recognise familiar individuals. A series of playback experiments were conducted exploring the differences in bonobos' reaction intensity to a familiar versus an unfamiliar voice.

The playback experiments, which were conducted at three European zoos, took advantage of the fact that bonobos are sometimes moved between zoos for population management. As such, calls played in the experiment only came from individuals who were previously familiar to the bonobo subjects. This experimental

design allowed us to simultaneously test the length of time bonobos can recognise the voice of a past social partner after differing periods of separation.

Over the course of my PhD research project my time was shared between the University of St. Etienne/Lyon, the University of St. Andrews and the three European zoos where the vocal recordings were collected and the playback experiments performed.

Manuscript 1 has been submitted to *Animal Behaviour.*
Manuscript 2 is in preparation for submission.
Manuscript 3 has been published in *Scientific Reports.*

MATERIALS AND METHODS

II. Materials and Methods

A. Subjects & Zoos

This thesis conducted research with captive bonobos living in three separate European zoos. All three zoos are certified by the European Associate of Zoos and Aquaria (EAZA), who set rigorous husbandry standards for all animals under its purview, including for diet, healthcare, facility sizes, and breeding programmes. In all three zoos, bonobos lived in stable multi-male, multi-female breeding groups with a mixture of related and unrelated individuals of all age categories and with similar daily routines and habitat structures. As such, despite being housed in separate locations, the three studied bonobo groups have markedly similar living conditions, making them ideal for comparative study. In this section, I describe the precise conditions at each zoo, including details on their enclosures, diet and daily schedule, followed by a description of the subjects included in the study.

1. Zoos

a. Apenheul

Apenheul Zoo is located in Apeldoorn, the Netherlands and is home to a group of ten bonobos. At the time of study, the social group was comprised of seven females and three males aged 3-35 (see Table 2, page 40, for details). The bonobo habitat at Apenheul consists of an indoor enclosure and outdoor island (Figure 4a). The indoor enclosure contains one large, main hall (630 m^3) and a smaller room (80 m^3) that are visible to the public and three off-exhibit rooms of varying sizes. The outdoor island (4,760 m^2) is covered by grass, shrubbery, multiple large trees (which were surrounded by live-wire to prevent the bonobos from climbing them), a marshy section with tall water plants, a small waterfall and

two large climbing structures (Figure 4b). Between 8 a.m. and 10 a.m. daily the bonobos were shifted into the off-exhibit rooms, while the zoo keepers fed them, checked their general wellbeing, conducted training exercises and cleaned the large hall. From 10 a.m. to 4 p.m. the bonobos only had access to on-exhibit rooms and the outdoor island, if the outside temperature was over 5°C; if the temperature was over 15°C the bonobos were exclusively kept outside during those hours.

Figure 4. The bonobo habitat at Apenheul Zoo. Figures labelled with an 'a)' are the two inside enclosures visible to the public and the figure labelled with a 'b)' is of the outside island.

The group had multiple scheduled feedings per day of a combination of fruits, vegetables and a single feeding of 'monkey chow' (grain based commercial product) in the afternoon. Seeds and 'browse' (edible leaves and branches) were spread each morning around both in the indoor and outdoor enclosures, and the bonobos regularly foraged on many of the plants growing in the outdoor habitat.

b. Planckendael

Planckendael is an animal park located in Mechelen, Belgium and is part of the Royal Zoological Society of Antwerp (KMDA). At the time of study, Planckendael housed nine bonobos, five females and four males aged 2-28 (see Table 2, page 40). The bonobo habitat consists of a single, large hall (600 m^3) visible to the public, five smaller off-exhibit rooms (25 m^3 each) and an outdoor island (a 3000 m^2) (Figure 5a). The outdoor island is covered by large grassy areas, two distinct sections of dense bushes and small trees, a large climbing structure, and a small rocky area (Figure 5b). Each morning the bonobos were shifted into the off-exhibit rooms to allow the keepers to check their wellbeing, perform training exercises and to clean the main hall. By 10 a.m. every day the bonobos were moved into the on-exhibit, main hall exclusively, unless the outside temperature exceeded 15°C, in which case they were given access to the island. Between 4 and 5 p.m. daily the bonobos were shifted back into the off-exhibit rooms for 30 minutes for feeding and to give the keepers temporary access to the main hall.

Figure 5. The bonobo habitat at Planckendael. The figure labelled with an 'a)' is of the large inside enclosure and the figures labelled with a 'b)' are of the outside island. The bottom right image is of vocal data collection in progress.

The group had multiple scheduled feedings per day of a combination of fruit and vegetables. In addition to the scheduled feeds, the group was regularly given branches to browse on and enrichment boxes filled with nuts, seeds and/or carrots, after the keeper had access to the main hall in the morning and afternoon. When outside, the group would regularly forage on grass and plants growing on the island.

c. La Vallée des Singes

La Vallée des Singes (VDS) is a primate park located in Romagne, France and houses the largest captive group of bonobos in Europe. At the time of study, there were 17 individuals in the social group, nine females and eight males aged 1-45 (see Table 2, page 40). The inside bonobo enclosure at VDS is comprised of a single main hall (1400 m^3) that is visible to the public (Figure 6a), which can be divided into two separate rooms of equal size by a large sliding door, as well as seven off-exhibit rooms of varying dimensions. The outdoor enclosure is an island divided into two sections by a high, solid wall (total island size is 1 hectare; Figure 6b). At the time of the study the bonobos were primarily using only the larger of the two sections, which has a large grassy area and a dense forested area, in which the bonobos had full access to the trees. In the mornings, the bonobos were limited to the multiple off-exhibit rooms while the main hall was cleaned. During opening hours, the bonobos were restricted to the main hall and given access to the island if the outside temperature was above 5°C between 10 a.m. to 4 p.m. Daily between 4 and 5 p.m. the keeper would check the bonobos wellbeing and conduct training exercises.

Figure 6. The bonobo habitat at La Vallée des Singes. The figures labelled with an 'a)' are two views of the main inside hall and the figure labelled with a 'b)' is of the outside island.

The bonobo group had multiple scheduled feedings per day of a combination of fruit and vegetables. In the morning, seed and browse was spread out in the main hall, and in the afternoons, the bonobos were fed monkey chow and a protein (chicken or eggs) or a grain (cooked rice or pasta).

2. Subjects

The following table provides an overview all individuals living in each of the three study populations.

Name	Sex	Age	Rank	Zoo	Father	Mother	Years at zoo
Bolombo	M	16	Low	Apen.	Yenge	Kosana	3
Besede*	F	8	Low	Apen.	Clyde	Binti	1
Hortense	F	35	High	Apen.	Unknown	Unknown	5.5
Jill	F	28	High	Apen.	Bosondjo	Laura	17
Kumbuka	F	14	Mid	Apen.	Hani	Molaso	14.5
Makasi*	M	4	n/a	Apen.	Zamba	**Zuani**	4.5
Monyama*	F	3	n/a	Apen.	Unknown	**Jill**	3.5
Yahimba*	F	4	n/a	Apen.	Zamba	**Kumbuka**	4.5
Zamba	M	15	Mid	Apen.	Kidogo II	**Hortense**	5.5
Zuani	F	23	High	Apen.	Unknown	Unknown	16
Busira	F	10	Low	Planck	Birogu	Eja	2
Djanoa	F	18	Mid	Planck	Santi	Yala	11.5
Habari*	M	7	Low	Planck	Vifijo	**Djanoa**	7
Lina	F	28	High	Planck	Vernon	Loretta	5
Lingoye*	F	8	Mid	Planck	Kirembo	**Lina**	5
Louisoko	M	15	Mid	Planck	Masikini	**Lina**	5
Lucuma	M	11	Mid	Planck	Kirembo	**Lina**	5
Nayoki*	F	2	n/a	Planck	Unknown	**Djanoa**	2
Vifijo	M	19	Low	Planck	Kidogo II	Hortense	4.5
Bondo*	M	22	Low	VDS	Mato	**Lisala**	2
Daniela	F	45	High	VDS	Camillo	Margrit	4
David	M	12	Mid	VDS	Kirembo	**Daniela**	4
Diwani	M	17	Mid	VDS	Masikini	**Daniela**	4
Kelele*	M	9	Low	VDS	Ludwig	Salonga	3
Khaya	F	12	Mid	VDS	Keke	Banya	4
Khalessi*	F	1	n/a	VDS	Unknown	**Khaya**	1
Kirembo	M	21	Low	VDS	Masikini	Kombote	4
Lingala	F	10	Mid	VDS	Mwindu	Jill	2.5
Lisala	F	33	Mid	VDS	Masikini	Catherine	2
Loto*	M	4	n/a	VDS	Kuno	**Ulindi**	1
Lucy	F	10	Low	VDS	Bosondjo	Lorel	1.5
Luebo*	M	6	n/a	VDS	Birogu	**Lisala**	2
Moko*	M	1	n/a	VDS	Unknown	**Ukela**	1
Nakala*	F	5	n/a	VDS	Ludwig	**Ukela**	3
Ukela	F	28	High	VDS	Bono	Natalie	3
Ulindi	F	20	Mid	VDS	Bono	Natalie	1

Table 2. Bonobo individuals and group. This table shows the composition of each zoo population. Rank was assigned to each individual based on interactions during agonistic encounters and feeding priority. Each rank assignment was then

discussed and verified with a minimum of two keepers at each zoo. Individuals marked with an asterisk were not included in the thesis and were not given a rank assignment. Individuals in bold in the "Mother" column are currently housed with their offspring. 'Age' and 'Years at zoo' were calculated as of January 1, 2014, and years were calculated in 6-month increments.

For all the studies included in this thesis, stable vocal patterns were required. Therefore, infant, juvenile or adolescent individuals, whose voices are not fully matured and are subject to change due to their ongoing development, were excluded from this work. A recent study showed that male bonobos show a sharp increase in testosterone around eight years old, indicating the onset of sexual maturity (Behringer et. al 2014), while bonobo females are thought to mature between the ages of 6-10, based mainly on natal group dispersion (Furuchi, 1989) but also with corroborating hormonal evidence (Behringer et. al 2014). Therefore, all individuals included in the study were over the age of 10, with the exception of two females who turned 10 during data collection. Additionally, one adult male at La Vallée des Singes, Bondo, had an unusually low calling rate and was excluded from study. This resulted in analysed vocal recordings from a total of 22 individuals; Apenheul (*n*=6), Planckendael (*n*=6) and Vallée des Singes (*n*=10), 14 females and 8 males, ranging in age from 10 to 45, with a mean age of 20 (See Table 2).

B. Vocalisations

The vocal recordings collected from all three zoos formed the backbone of this thesis and were used to investigate the primary hypotheses in all three manuscripts, either by studying the vocalisations extracted from the recordings or by using them in playback experiments.

The overall aim when recording was to collect a sufficient number of vocalisations from independent events for statistical analysis from each individual in each group. Additionally, I aimed to collect a balanced sample of calls from each individual encompassing the variety of observed contexts in which vocalisations were emitted and a good sample of the different call types.

1. Data collection

All recordings were taken *ad libitum* using a Zoom H4 Digital Multi-track Recorder (44.1 kHz sample rate, 16 bits per sample, .wav files) - recording in stereo, with one channel devoted to a Sennheiser MKH70-1 ultra-directional microphone recording all bonobo vocal behaviour and the second channel connected to a micro-tie audio recording device, model AKG MPA III (the bottom right image in Figure 5 shows recording in process). The micro-tie recording device allowed me to record behavioural observations in situ and with temporal synchronisation to the bonobo vocalisations recorded by the directional microphone.

Recordings at Apenheul Zoo were collected from May 14, 2013 – July 6, 2013, as well as on March 12, 2014, between 9 a.m. and 5 p.m. daily, and amounted to roughly 175 hours of total recording time. Due to limited microphone access to the indoor enclosures during the zoo's opening hours, the vast majority of the recordings were taken when the bonobos were on the outside island, which was daily due to the time of year recordings were done. Inside recordings were restricted to before 10 a.m. and after 4:30 p.m.

Recordings at Planckendael Zoo were collected from March 20, 2013 – May 10, 2013 and from February 20 – March 4, 2014, from 10 a.m. to 6 p.m. primarily on weekdays and amounted to roughly 190 hours of total recording time. I had microphone access to the main indoor enclosure and outdoor island during recording hours, as well as access to off-exhibit indoor enclosures after 4 p.m.

Recordings at Vallée des Singes were collected from October 28, 2013 – November 25, 2013 between 9 a.m. and 5 p.m. daily, and amounted to roughly 115 hours of total recording time. I had unlimited microphone access to all indoor and outdoor enclosures during recording hours.

2. Data extraction

After daily recordings were completed, the sound files were uploaded onto a MacBook Pro (15-inch, Mid-2012), which was used for all acoustic and statistical analyses. Individual calls or call sequences, that could unequivocally be assigned to a single caller, were cut and extracted from the larger sound file using Raven Pro 1.3 and Audacity 1.3.14. For each call file, all behavioural observations made via the micro-tie were transcribed, and the file was saved in mono format, discarding the second channel and saving only the bonobo vocalisation. To be included in the acoustic analysis, each call had to meet the following requirements:

1) be unmistakably assigned to a single caller;

2) be clear of any overlap, with other calls or background noise, and of good quality allowing for accurate manual and automatic measurements;

3) emitted by an adult

4) the social or individual context a call was given in had to be unambiguous;

5) be assignable to a call type.

The final two points will be discussed in detail below.

a. Call types

Calls were classified as different call types based primarily on two descriptive studies of the bonobo vocal repertoire, de Waal (1988) and Bermejo & Omedes (1999), with additional guidance on call types given during feeding from Clay & Zuberbuhler (2009). Call classifications were a considerable endeavour, considering that previous studies only provided qualitative description with little to

no quantitative analyses of the acoustic differences between the call types. Calls were categorised based on visual comparison between our data and spectrograms provided by these studies, as well as discrimination by ear. Additionally, for some call types de Waal (1988) and Bermejo & Omedes (1999) provided general acoustic information, such as average duration or peak frequency, which was also used as a guide.
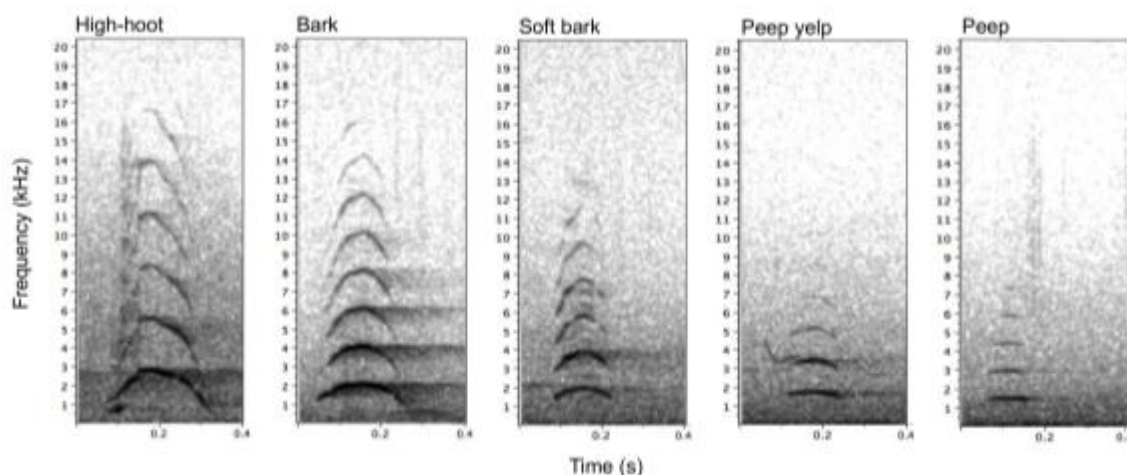
Figure 7. Five common call types studied in this thesis. Call types are labelled above the spectrograms.

There were five call types that comprised the vast majority (75%) of the clear, good quality calls that I collected from the adult bonobos. These five call types were produced with sufficient regularity by the majority of the adult bonobos and were therefore included for analyses. The five call types are (Figure 7; See Results of Manuscript 1 for comprehensive description of the acoustic parameters of each call type):

1) High-hoots (14% of dataset): described as 'staccato high-hoots' by de Waal (1988) and 'high-hoots' by Bermejo & Omedes (1999), but excluded legato high-hoots;

2) Barks (18% of dataset): described as 'barks' and 'wieew barks' by de Waal (1988) and as 'barks' and 'composed barks' by Bermejo & Omedes (1999);

3) Soft barks (17% of dataset): described as 'soft barks' by Bermejo & Omedes (1999), 'food barks' by Clay & Zuberbuhler, (2009) and not described by de Waal (1988);

4) Peep-yelps (18% of dataset): described as 'peep-yelps' by de Waal (1988); Bermejo & Omedes (1999); Clay & Zuberbuhler (2009); and

5) Peeps (11% of dataset): described as 'peeps' by de Waal (1988); Bermejo & Omedes (1999); Clay & Zuberbuhler (2009).

There were other call types that were described by de Waal (1988), Bermejo & Omedes (1999), and Clay & Zuberbuhler (2009) that were identified in our dataset, for example, yelps (8% of adult call dataset), hiccups (2%), legato hooting (2%), whine (1%), whistle (1%), low hooting (0.5%) (Figure 8a-c). These additional call types were not investigated, as they were emitted less regularly by many of the individuals and thus had insufficient numbers for analyses.
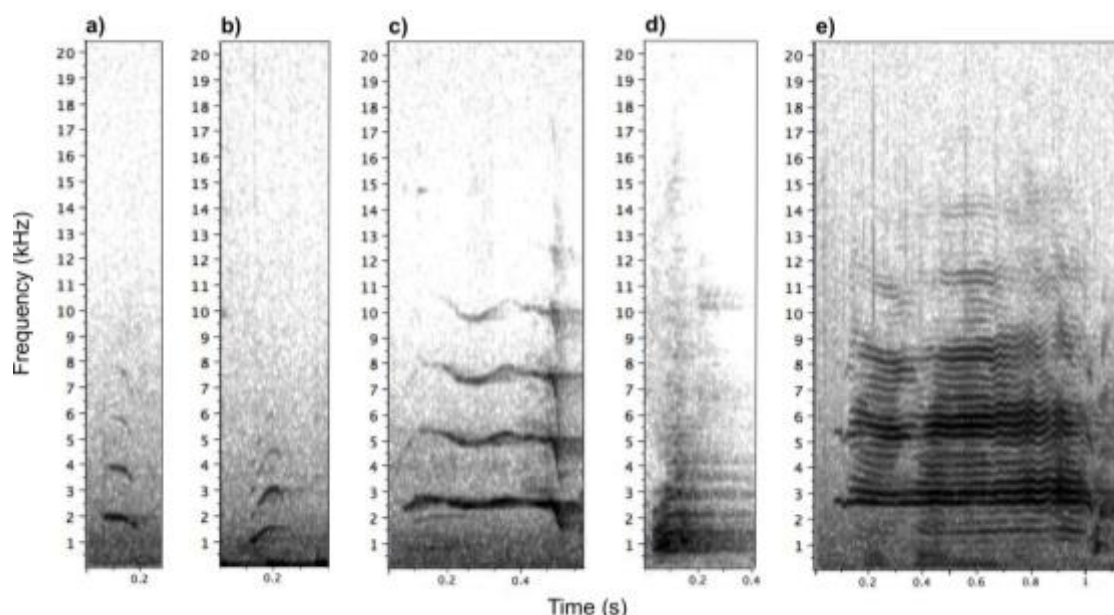
Figure 8. Examples of other call types not included in this thesis. a) Yelp, b) Hiccup, c) Legato high-hoot, d) Grunt and e) Scream.

The bonobo repertoire contains a broad range of vocalisations from atonal, noisy screams (Figure 8e) to tonal barks with a clear harmonic structure (Figure 7). The vast majority of bonobo call types, including the five general types listed above, fit into the latter category, allowing for clear analysis of the fundamental frequency,

as well as the distribution of energy amongst the harmonics. However, two regularly observed call types, screams and grunts, had noisy features characterized by non-linear phenomena and an unclear harmonic structure (Figure 8d & 8e). These two call types would therefore require a different set of acoustic measurements to accurately describe their structure, and as such were excluded from the studies presented in this thesis.

b. Call contexts

If a caller could be identified, social or individual contextual information was recorded. Calls given in the following eight contexts were retained for analyses (Figure 9):

1) 'Contact' – calls given when a subject was resting or moving but nothing else, often elicited vocal responses from other individuals in the group;

2) 'Grooming' (Figure 9a) – calls given during bouts of grooming between individuals;

3) 'Pre-feeding' – calls given directly prior to or at the start of scheduled feedings, when group excitement was extremely high (excitement described as increased sexual activity, displacements, pacing, pilo-erection, vocal activity, and displays);

4) 'Feeding' – calls given during scheduled feedings;

5) 'Foraging' (Figure 9b) – calls given when foraging for or eating food found outside or in the indoor enclosure, specifically outside of scheduled feeding times when food was anticipated;

6) 'Aggression' (Figure 9c) – calls given during agonistic encounters, including mild aggression (no physical contact), aggression (mild physical contact, such as hitting, kicking or grabbing between only two individuals), conflict (a range of physical contact, including biting, and often between more than two individuals) - agonistic interactions were also classified into calls from aggressors, from bystanders or from victims (victim calls were not included in this thesis as they were almost exclusively noisy, atonal screams);

7) Change of environment – calls given when individuals were shifted from one enclosure to another or from inside to outside; and

8) External event – calls given when an individual was visibly startled by or responding to a disturbance external to the group. The most common example was in response to birds, specifically waterfowl, which would often hide amongst the shrubbery on the outside habitats.

Figure 9. Examples of bonobo behavioural contexts. Three examples of contexts in which calls were emitted: a) Grooming, b) Foraging, c) Aggression.

There were some contexts that were identified throughout data collection but could not be included in analyses. Calls given during sex were excluded, as they were largely screams, which were excluded based on their acoustic features, as was the case for the calls of victims in agonistic encounters. This was also the case for calls emitted during the context of play, which were largely non-linear, non-harmonic grunts and panting laughs. Calls given during evening nesting were only recorded at one zoo (due to restrictions on research hours at two of the three zoos), which prevented a systematic comparison across populations and therefore were excluded.

3. Acoustic analysis

After the data were collected and organised, I performed acoustic analyses on the vocal data set.

a. Visualising sound

Raven Pro 1.3 was used to analyse all calls in the dataset. Each individual call was visualised in both a spectrogram and waveform sound representation. All spectrograms were viewed with a window size of 512 points (Figure 10a).
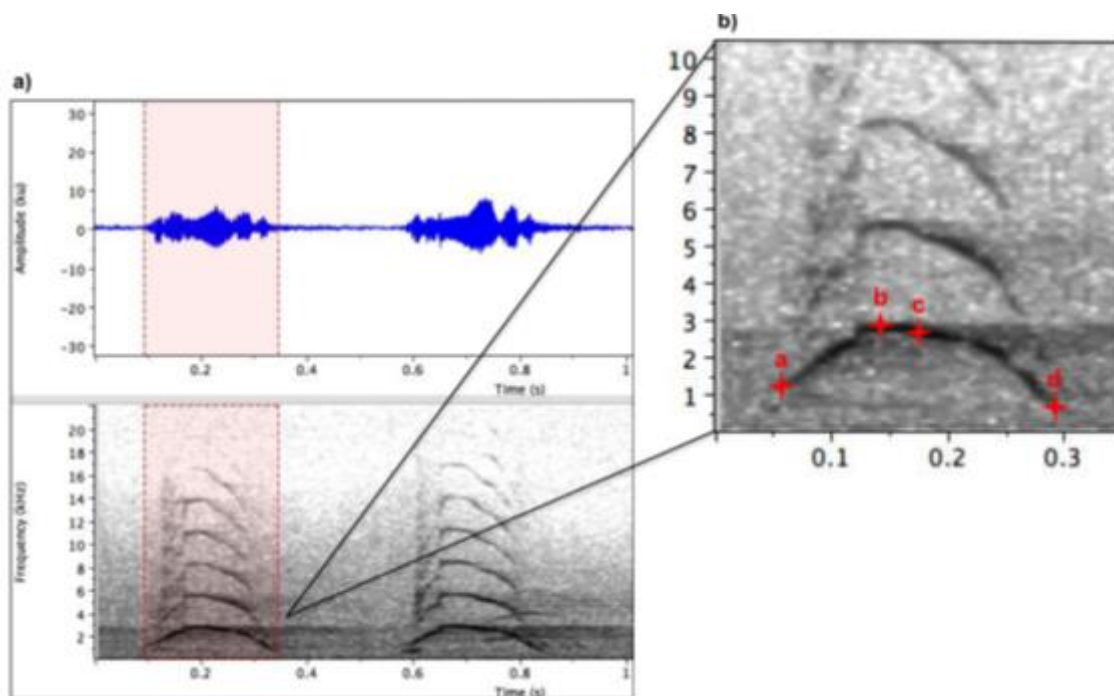
Figure 10. Acoustic measurements. The left figure, a), shows the call selection process, the boundaries of the red shaded area mark the start and end point of the call (for which call duration was calculated) and all automatic measurements are calculated by the Raven software from within those boundaries. The right figure, b), shows the points along the fundamental frequency (F_0) where the manual measurements were taken – point 'a' is the F_0-Start; 'b' is the F_0-Midpoint; 'c' is the F_0-Maximum; and 'd' is the F_0-End. These were then used to calculate other measurements, such as slope of F_0.

b. Acoustic measurements: frequency and temporal domains

For each call in the dataset a total of thirteen manual measurements were taken or calculated, and eight automatically computed measurements by Raven Pro were retained. Each chosen measurement was carefully considered. In recent years, advanced acoustic analysis software, such as Seewave (Sueur et al., 2008), has been developed and can create a holistic picture of vocalisations through many measurements. However, these types of software packages normally require completely clean sounds, such as those recorded from animals inside sound boxes in laboratory conditions. To create such an environment for a full-grown ape would be extremely costly and not at all functional, as bonobo vocalisations are produced in response to complex social interactions. As such, all recordings were collected from bonobos in their regular captive environments, which included ambient environmental noise (trees, birds, rain, ventilation systems, human zoo guests, etc.) and differing distances from the microphone (the bonobos ranged from 0.5 to 20 metres from the microphone). These conditions required us to discard any pure decibel measurements, as they would be affected by the unstable recording distances, and were replaced by measurements that investigated the distribution energy within a call selection as a function of time or frequency. Additionally, Raven Pro's automatically computed measurements were complimented with manual measurements describing the fundamental frequency.

To take a call's measurements, I first established the temporal boundaries of the call using both the spectrogram and waveform views (Figure 10a). Once a call was selected the fundamental frequency was identified and a variety of manual measurements (listed below) were recorded. Thirteen manual parameters describe the fundamental frequency and its temporal modulation (Figure 10):

1) F_0-Start: fundamental frequency at beginning of the call, in Hz;

2) F_0-Midpoint: fundamental frequency at the temporal midpoint of the call, in Hz;

3) F_0-End: fundamental frequency at end of the call, in Hz;

4) F_0-Maximum: highest frequency reached on the fundamental, in Hz;

5) F_0-Maximum Time: time that F_0-Maximum is occurs, in seconds;

6) F_0-Max proportion: point over the duration of the call at which F_0-Maximum is reached, calculated as a proportion (F_0-Maximum Time/Call Duration) (as not all calls are of the same duration, this measurement was more comparable across calls than F_0-Maximum Time);

7) Call Duration (s);

8) Ascending Slope: calculated as (F_0-Maximum – F_0-Start/F_0-Maxproportion –0);

9) Descending Slope: calculated as(F_0-End – F_0-Maximum/1 –F_0Maxproportion);

10) Slope–Start to Midpoint of F_0: (calculated as: F_0 at midpoint of call duration – F_0-Start/Time at midpoint of call duration – 0)

11) Slope–Midpoint to End of F_0: (calculated as F_0-End – F_0 at midpoint of call duration/Call duration – time at midpoint of call duration);

12) Onset transition ΔHz: calculated as (F_0-Midpoint – F_0-Start), in Hz;

13) Offset transition ΔHz: calculated as (F_0-End – F_0-Midpoint), in Hz.

Eight automatically computed parameters, describing the distribution of energy of among the entirety of the frequency spectrum, were also recorded, and are as follows (Figure 10a):

1) Q1-Frequency: the frequency at which the call is divided into two frequency intervals containing 25% and 75% of the call's energy, in Hz;

2) Q1-Time: the time point along the call at which the call is divided into two time intervals containing 25% and 75% of the call's energy, in seconds;

3) Q3-Frequency: the frequency at which the call is divided into two frequency intervals containing 75% and 25% of the call's energy, in Hz;

4) Q3-Time: the time point along the call at which the call is divided into two time intervals containing 75% and 25% of the call's energy, in seconds;

5) Centre Frequency: the frequency at which the call is divided into two frequency intervals of equal energy, in Hz;

6) Centre Time: the time along the call at which the call is divided into two time intervals of equal energy, in seconds;

7) Maximum Frequency: the frequency at which the maximum energy occurs in the call, in Hz;

8) Maximum Time: first time point along the call where maximum amplitude, from waveform, occurs, in seconds.

After each frequency or temporal parameter was measured or calculated, the data set was reduced per the need of each study (see each following section for details).

C. Investigating individual vocal signatures and the effect of call type

Here I will describe the methods used to investigate the identity information found in the five most common bonobo call types (**manuscript 1**).

1. Data set

In total, 1,850 individual calls were retained for analyses and comprised the five main call types: high-hoots, barks, soft barks, peep-yelps and peep. There was an average of 88 total calls per individual regardless of call type (S.D.= 37.9, minimum=45, maximum=227).

Of the 22 adult individuals from whom vocal data were collected, 21 were included in the analyses for manuscript 1. A single female from Planckendael (Busira, See Table 2) was excluded from the analyses due to low call numbers in all but one of the call types.

2. Acoustic parameters

In this thesis discriminant function and principle component analyses are used throughout, and the dependent variables, which are the acoustic parameters, were carefully selected. In these types of analyses, dependent variables are interrelated, or correlated, and can describe different aspects of the same underlying dimension. The analyses use the correlations between the variables to reduce them into a smaller number of factors that describe the maximum amount of common variance (Field, 2009). However, if the variables are perfectly correlated it is not possible to determine each one's unique contribution to the data. Therefore, it is advisable to not retain variables that are very highly correlated with one another.

For this study, any variables that correlated at 0.9 or above, on the correlation matrix, at least one was removed. This resulted in 16 of the acoustic measurements, as were described in the previous section, being retained for the analyses: nine manual parameters (F_0-Start, F_0-End, F_0-Maximum, F_0-Maxproportion, Call Duration, Ascending Slope, Descending Slope, Slope–Start to Midpoint of F_0 and Slope–Midpoint to End) and seven automatically computed parameters (Q1 Frequency, Q1 Time, Q3 Frequency, Q3 Time, Centre Frequency, Maximum Frequency and Maximum Time).

3. Statistics

In this section I will describe the two primary analyses employed in the first manuscript to investigate identity information. It is important to note that there were other statistics used, such as the chi-square test, throughout the manuscript. These additional statistics are well known, commonly used and do not require extensive explanation. The statistical procedures described here, are often used in studying individual vocal signatures, but are perhaps are less common in other areas, and therefore, I will provide some description.

a. Permuted discriminant function analysis

A permuted discriminant function analysis (pDFA) performed on the statistical software R (Version 3.2.0) was used to: 1) provide a quantitative description of the distinctiveness of each of the five call types; 2) investigate if an individuals' vocalisations could be reliably classified to him/herself in each of the five call types based only on the acoustic features outlined.

Discriminant function analysis (DFA) is a common statistical tool used to test if select variables can be used to correctly discriminate or predict between two or more groups or categorical class. However, DFAs are most reliable when the group is comprised of multiple independent data points. When analysing non-independent data, Mundry and Sommer (2007) argued that using a classical DFA is a case of pseudo-replication and can vastly inflate results. To deal with this problem, the authors created a statistical method called a *permuted* discriminant function analysis, which can reliably analyse non-independent data (Mundry & Sommer, 2007). In the dataset used in this thesis each individual bonobo contributed multiple vocalisations, or data points, to each analysis. Whether considering the categorical class to be call type or individual, both contained multiple data points from a single individual and were therefore non-independent. As such, I have used the permuted discriminant function analysis in manuscript 1.

A pDFA begins with a classical DFA procedure with its standard two steps: 1) a discriminant function is constructed from a training data set, 2) the model is tested on an independent data set to obtain a percentage of calls correctly classified regarding their call type. The pDFA then goes further by building a number of randomised data sets, followed by the classic DFA procedure on each of the randomised data set and the original data set. A significance value is generated for the original data set by dividing the number of randomised data sets that classified the data at the same percentage rate or higher as the original data set

by the total number of data sets (Mundry and Sommer, 2007; Mathevon et al., 2010).

b. Potential for individual coding

The potential for individual coding (PIC) is a technique that is often used by researchers investigating identity information in vocal behaviour (e.g. Robisson et al., 1993; Bee, et al., 2001; Charrier, et al., 2001; Bouchet et al., 2011; Salmi et al., 2014). PIC analysis compares levels inter- and intra-individual variability to create an index of individuality and can be calculated for single call type or the whole repertoire, as well as for each acoustic parameter separately.

I first investigated the dataset as a whole, regardless of call type. To do so I followed the following procedure:

1) calculated the mean and standard deviation (S.D.) for each acoustic parameter for each individual's set of calls (each individual calculated separately);

2) used these values to calculate the coefficient of variation for each parameter for each individual using the formula, CV= 100 x (S.D. / Mean);

3) calculated the intra-individual coefficient of variation by taking the mean of the individual CV values calculated in the previous step, CV_{intra}= mean of individual CV values;

4) calculated the inter-individual coefficient of variation for each parameter by calculating the mean and standard deviation for each parameter over the entire data set, regardless of individual identity using the formula CV_{inter}= 100 x (S.D. / Mean);

5) calculated the potential for identity coding, PIC = CV_{inter}/CV_{intra};

6) steps 1-5 were then repeated for each of the five call types separately.

These types of univariate analyses have fallen out of favour in recent years, as researchers have come to a greater appreciation that most variables are influenced by multiple factors (multivariate analysis). However, univariate analysis

can still be useful in exploring and understanding a dataset, and especially when used in combination with multivariate analyses. Additionally, for this study, I endeavoured to understand the individual vocal signatures in the bonobo repertoire from an evolutionary perspective, and as such comparison with other non-human primate studies was essential. As PIC analysis was used in all other studies investigating identity vocal signalling in multiple call types of a species' repertoire, including it in this study was necessary.

D. Investigating the effect of familiarity and relatedness on call structure

In this section, I will describe the methods used to investigate the affect familiarity and relatedness have on the vocal similarity between pairs of individuals in the bark call type (**manuscript 2**).

1. Data set

a. Call type

For this study only the bark call type was investigated. Barks were chosen because: 1) of the five call types investigated in manuscript 1, barks (along with high-hoots) showed high levels of individual coding, and I aimed to test the effect of familiarity and relatedness on a call type that we could reasonably ensure was communicating identity information; 2) importantly, the bark call type was emitted in multiple contexts and regularly by all adult subjects, as such there were a statistically sufficient number from each individual to be able to include all 22 adults in the analysis.

A total 431 calls were retained for analyses with a mean of 20 calls per individual (range of 5–55).

b. Subjects

All 22 adult subjects were included in this study and thus comprised 231 unique dyads. For each dyad a variety of information was included. To investigate the effect that relatedness had on vocal similarity, relatedness was evaluated in two ways: 1) using the raw relatedness coefficient (r), or 2) related categories (not related, $r = 0$; distantly related: $r \leq 0.125$; or closely related: $r > 0.125$). All relatedness coefficients (r) were calculated and provided by the European Association for Zoos and Aquariums (EAZA) bonobo studbook manager.

To investigate the effect of familiarity, I included: 1) whether individuals currently live together or not, 2) if they have ever lived together over the course of their life, and 3) the amount of time they lived together. Unfamiliar dyads were completely unknown to one another and had never lived together. Familiar dyads included individuals that were currently in the same group or had previously lived together. In both cases, the amount of time spent living together was rounded in 6 month increments and ranged between 6 – 252 months (with a mean of 71 months for familiar individuals, and only a single familiar dyad having lived together for less than 12 months) and included 25 past familiar dyads.

Individuals who were familiar with one another, but no longer living in the same group, presented a possible problem for our analysis, as I equated individuals living in the same group to be familiar and individuals living in separate groups to be unfamiliar. Ideally, they would have been completely removed from the analysis; however, almost all 25 past familiar pairs fell into the related category and removing them would have greatly unbalanced the number of data points between the unrelated and related groups. Therefore, for the primary analysis these past familiar pairs were included, and I controlled for their possible confound by including both length of familiarity in months and whether they currently live together or not in the model. However, for the further two models the data were

treated differently, and I was thus able to remove them from these secondary analyses without unbalancing the data (details of the analyses are given below).

c. Acoustic parameters

As the vocal dataset used for this study was considerably smaller than for manuscript 1, the number of acoustic and temporal parameters included was reduced. To do so, I used a stricter standard for reducing correlated variables and removed parameters that correlated above 0.8 (instead of 0.9 as in manuscript 1). This resulted in a total of 12 measurements being retained for analysis – six manual acoustic parameters (F_0-Start, F_0-End, F_0-Maximum, Call Duration, Slope–Start to Midpoint of F_0, Slope–Midpoint to End of F_0) and six automatically computed parameters (Q1-Frequency, Q1-Time, Q3-Frequency, Maximum Frequency and Maximum Time).

2. Analysis

To quantify vocal similarity between individuals, I used two separate measurements. First, I calculated the mean of each acoustic variable describing the bark call type for each individual; I then used these means to calculate the absolute difference for each variable for each dyad. This method provided a straightforward way to assess if and how relatedness and familiarity affected each acoustic variable. To provide a more holistic and comprehensive view of vocal similarity a secondary measure, Euclidean acoustic distance, was created. This was done by using a principle component analysis to reduce the 12 acoustic variables into two factors. For each individual the centroid between the two factors was calculated and then used to calculate the Euclidean acoustic distance between each dyad.

The two measures of vocal similarity were used as dependent variables in four separate linear mixed models. The first model (*primary model*) was used to

provide a comprehensive view of the effect of both relatedness and familiarity on each of the acoustic variables (as measured by the absolute difference), as well as the composite variable, the Euclidean acoustic distance. This first model also included a number of fixed and random factors, and tested for any interaction between the two main independent variables. The following three secondary models were used to test the more subtle nuances in the data set, and only used the Euclidean acoustic distance as the dependent measure.

a. Primary model

The primary model provided the most thorough multivariate investigation as to which variable or variables affected vocal similarity. A linear mixed model was performed with *relatedness* (categorised as unrelated: $r=0$, $n=147$ dyads; distantly related: $r \leq 0.125$, $n=49$; or closely related: $r > 0.125$, $n=35$) and *group* (currently living at the same zoo, $n=156$; or currently living at different zoos, $n=75$) as the main independent variables, and the Euclidean acoustic distance and the absolute differences for all twelve acoustic measurements as the dependant variables. The following variables were also included as fixed factors: familiarity measured as time spent together in months, sex composition of the dyad, and the absolute difference in age between each individual in each dyad. Additionally, the identity of the two individuals in the dyad were considered as random effects.

b. Three models for secondary analyses

By using further analyses, I was able to explore the effect that relatedness and familiarity have on vocal similarity in greater detail and from different perspectives. As there were multiple models run for the secondary analyses, the Euclidean acoustic distance was the only dependant variable used to measure vocal similarity. Additionally, as mentioned, the past familiar pairs were not included in these further analyses.

The first two models of the secondary analysis had two primary goals: 1) to investigate vocal similarity in the familiar and related groups independently from one another and 2) to investigate familiarity and relatedness using continuous variables, specifically the relatedness coefficient and time spent together. This was done by running two linear mixed models limited to the unrelated and unfamiliar dyads separately. For the unrelated dyads (*n*=136), I investigated the relationship between the familiarity of dyads (time spent together in months, continuous variable and limited to pairs currently living in the same group) and the Euclidean acoustic distance. For the unfamiliar dyads (*n*=131), I investigated the relationship between the relationship coefficient (as a continuous variable) and the Euclidean acoustic distance. For these two models, sex composition of the dyad (coded as different sex or same sex) and the absolute difference in age between each individual in each dyad were considered as fixed factors, and the identity of the two individuals in the dyad were considered as two random effects.

The final analysis was performed to investigate vocal similarity by viewing the data as four separate groups: 1) dyads that live together and are related (Familiar-related, *n*=27); 2) dyads that live together and are not related (Familiar-unrelated, *n*=48); 3) dyads that are related but completely unfamiliar (Unfamiliar-related, *n*=43); 4) dyads who are unrelated and unfamiliar (Unfamiliar-unrelated, *n*=88). As in the primary model, the following were included as fixed factors: time spent together in months, sex composition of the dyad and the absolute difference in age between each individual in each dyad. The identity of the two individuals in the dyad was considered as two random effects. This model provided another view as to which factor, relatedness or familiarity, had the strongest effect on vocal similarity. If relatedness had the greatest effect we would expect to see the two groups with related individuals being more acoustically similar than unrelated groups. If familiarity had the strongest effect we would expect to see the two groups with familiar individuals with the highest acoustic similarities. Alternatively, if relatedness and familiarity both had an effect, we should see the familiar-related

group as the most acoustically similar, the unfamiliar-unrelated as the least and the unfamiliar-related and familiar-unrelated groups as roughly equal.

E. Investigating individual recognition – Playback experiments

Playback experiments are a classic tool in which subjects are played a recorded sound, typically a conspecific or predator vocalisation, and researchers record the subject's reactions. This experimental design has been used to test a variety of questions regarding animal vocal communication, but is commonly used to investigate referentiality in a species' alarm or food calls (Evans & Evans, 1999; Manser et al., 2001; Seyfarth & Cheney, 1980; Zuberbuhler, 2000) or to test a species' recognition capacity (Briefer et al., 2012; Levrero et al., 2015; Rendall et al., 1996; Townsend et al., 2010). To test individual recognition in bonobos, I used a series of playback experiments to observe the bonobos' reactions to the voices familiar and unfamiliar individuals (**manuscript 3**).

In all playback experiments, there are two primary risks that can negatively affect the experiment. Firstly, if the subject identifies the played back sound as an experimental manipulation, either by seeing the researcher or by problems with the sound being broadcast. Secondly, if the subject becomes habituated to the sound being played, meaning that they have been overexposed to the sound stimulus and cease to respond to it. In the following section I will describe the playback experimental protocol used to test individual recognition in bonobos and how the experiment was designed to avoid those common problems.

1. Playback experiments

a. Subjects

In captivity bonobos can be moved between zoos for breeding programmes and population management. As detailed records of all captive bonobos are well maintained, I was able to investigate the life history of each adult bonobo within the three zoo populations (Apenheul, Planckendael and La Vallée des Singes) to identity individuals who were previously housed together. This resulted in identifying 15 individuals who had been housed in the past with another bonobo from within the subject group (Table 3). I decided to only test vocal recognition in pairs of individuals who previously, instead of currently, lived together for two reasons. Firstly, to make the experiment realistic we could not play vocalisations from an individual who was visible to other group members, as it would likely cause confusion. Therefore, if we wanted to do so we would have needed to separate the individual from the rest of the group. At the time of the experiments, the individuals in each group were rarely separated from one another and doing so would cause undue stress and likely distort any results. As such, we decided to test vocal recognition with only past familiar individuals to avoid disturbing the groups' normal compositions. Secondly, using the voices of past familiar individuals also allowed us to test long-term vocal recognition.

All individuals included in the experiment, both as observed subjects or broadcast voices, were over the age of 7 when housed with a past familiar social partner and a minimum of 10 years old at the time of the experiment. Observed subjects were between the ages of 10 and 45 (mean age = 21; median age = 19) and were balanced across sex (male n = 7; female n = 8) and zoo (Apenheul n = 5; Planckendael n = 6; La Vallée des Singes n = 4) (Table 3). As mentioned, the experimental design also allowed us to test the dynamics of long-term vocal recognition, as each of our familiar pairs had been separated from one another

for varying numbers of years (separation time: 2–3 years, n = 4; 4.5–5.5 years, n = 8; 8–9 years, n = 3).

Subject	Age	Sex	Zoo	Rank	Familiar Call Conditions			
					Separation Time (months)	Name of familiar	Sex of fami-liar	Subject and familiar relation
Jill	29	F	Apen	High	33	Lingala	F	Mother-daughter
Hortense	36	F	Apen	Mid	100	Vifijo	M	Mother-son
Zuani	22	F	Apen	High	33	Lingala	F	Unrelated
Kumbuka	14	F	Apen	Mid	33	Lingala	F	Unrelated
Zamba	16	M	Apen	Mid	100	Vifijo	M	Full Brothers
Lina	28	F	Planck	High	54	Khaya	F	Unrelated
Louisoko	16	M	Planck	Mid	54	Khaya	F	Unrelated
Lucuma	11	M	Planck	Low	54	Khaya	F	Unrelated
Djanoa	19	F	Planck	Mid	65	Hortense	F	Unrelated
Vifijo	19	M	Planck	Low	100	Hortense	F	Mother-son
Busira	10	F	Planck	Low	32	Lisala	F	Unrelated
Daniela	45	F	VDS	High	54	Lina	F	Unrelated
Kirembo	21	M	VDS	Low	54	Lina	F	Unrelated
Diwani	17	M	VDS	Low	54	Louisoko	M	Paternal half-brother
David	12	M	VDS	Mid	54	Louisoko	M	Unrelated

Table 3. Playback back experiment subjects. Detailed information on the relationship between the subject and the familiar subject, whose calls were used as playback stimuli to test long-term vocal recognition.

b. Experimental realism

Bonobos are an extremely intelligent species, and captive bonobos are exposed to regular human interaction. As such, my colleagues and I went to great lengths to 'convince' the bonobos that the played back vocalisations were emitted by a real bonobo and were not an experimental manipulation. To do so, we pretended as if the past-familiar bonobos were being transferred into each group, and we followed all the normal procedures that occur during actual bonobo transfers (which included: driving a truck/tractor up to the bonobo building carrying an animal crate, the crate was placed to the opening to the separation cage or carried into the building, and the doors were opened and closed to move the imaginary bonobo into the separation room). As all of the subjects had previously experienced a real transfer event in similar conditions, we expected that the mimicked transfer was the best way to provide a context for the bonobos as to why they would hear a foreign or past familiar bonobo vocalisation.

c. Playback Stimuli & Experimental protocol

Calls used for the playback stimuli were taken from the databank of vocal recordings amassed from the three zoos. The mock transfer event provided us with a context in how to choose the call types used for the playback stimuli. Calls were selected on the basis of acoustic similarity to vocalisations recorded during an actual transfer event – where two individuals were transferred together to La Vallée des Singes in 2012 – and could be generally described as peep-yelps. Individual call sequences contained 4–6 calls and had decreasing intercall intervals along the progression of the sequence. Each playback stimulus contained a unique, acoustically distinct, set of calls. The call sequences used for the unfamiliar trials at Apenheul came from a single female unknown to all individuals in the group, this was also the case for La Vallée des Singes. At Planckendael, there was not a single individual within our database that was

unknown to everyone; therefore, at this zoo each broadcast individual was familiar to some and unfamiliar to others.

At each zoo, the experiment consisted of a single mock transfer performance followed by a total of five playback trials. Three to four observers were set-up around the cages, each observing and video recording a single, randomly assigned individual. By testing multiple individuals in each playback trial, we exposed the bonobos to fewer playback stimuli and thus reduced the risk of habituation. While this technique created a possible non-independence of the reactions of tested individuals, together with the fact that multiple individuals' reactions were recorded from the same playback trial, I later controlled for it by entering trial number as a random factor in statistical models.

After the mock transfer, as well as between each playback trail, we waited until the group had returned to baseline behaviours (e.g. resting, foraging, grooming) before proceeding to the next trial. This ensured that the bonobo's responses were in direct reaction to the most recent stimulus and independent from previous ones. In total, each subject was tested once in each of the two experimental conditions. The order in which each bonobo heard the stimulus for each condition was counterbalanced (8 individuals heard the familiar first, while 7 heard the unfamiliar first).

2. Analysis

a. Measuring behavioural reactions

To inform which behavioural measures would be included, we relied on reported behaviours from when bonobos encounter neighbouring groups in the wild (Hohmann & Fruth, 2002), my own observations during a transfer of a female into the Apenheul group, and on previous studies investigating vocal recognition in a variety of species (Bergman, 2010; Insley, 2000; Matthews & Snowdon, 2011;

Ramos-Fernandez, 2005; Rendall et al., 1996). We assessed a variety of measures on body and head movements, in relation to the speaker and in general. Social interactions and vocal responses were also coded; however, our playbacks elicited no interactions between individuals (neither aggressive or affiliative) and only one vocalisation (a single call by an unidentified individual). Therefore, the following eight behavioural variables, measured in the 60 seconds following the playback, were included:

1) Latency to the first behaviour displayed after the start of the playback (start of a behaviour or cessation of a behaviour, for instance if they were eating and stopped);

2) Latency to the first locomotion after the start of the playback broadcast;

3) Total duration of locomotion in any direction;

4) Total duration of locomotion toward the speaker;

5) Number of separate locomotion occurrences;

6) Duration of time spent looking toward the speaker;

7) Number of times an individual looked toward the speaker;

8) Total number of head movements (each change of head direction was counted as one movement event).

After the experiments were completed each video was given a randomly assigned number, and I then coded all videos blind to the experimental condition (familiar or unfamiliar) a minimum of four weeks later. In addition, F. Levrero blind coded 67% of the videos to ensure inter-observer reliability, and coding results were compared for each variable separately. The intraclass correlation coefficient (ICC) measures agreement between two raters or coders, and is considered strong between 0.700 – 0.800 and in almost perfect agreement above 0.800. The ICC was above 0.860 for all variables.

b. Statistical analysis

I first performed a principal component analysis on the eight dependent behavioural measures and retained a single composite score (PC1). To test for an effect of familiarity versus non-familiarity on the bonobos' behavioural response, a linear mixed effect model with PC1 as the dependent measure (using statistical software R, Version 3.2.0) was used. The following variables were also included as fixed effects: subject rank, subject sex, and subject age; and random effects: individual identity, playback trial number, and zoo location.

Testing for the effect of separation time on individual recognition was challenging, given the small sample size in each of the separation time categories (separation time: 2–3 years, $n = 4$; 4.5–5.5 years, $n = 8$; 8–9 years, $n = 3$). Therefore, to do so, I used two different linear models, one more conservative than the other. The first model investigated the differences in reaction (using the PC1 score) only to the voices of past familiar partners between the three separation groups, and was followed by post-hoc multiple comparison tests. The second, more conservative model, considered individual differences in reactivity to the playbacks by calculating the absolute difference in each individual's PC1 score between the unfamiliar and familiar conditions. This absolute difference score was then used in the model as the dependent variable to compare the differences between the three separation groups.

RESULTS

Individual vocal signatures across the bonobos' graded repertoire

Sumir Keenan[1,2], Nicolas Mathevon[1], Florence Nicolè[3], Jeroen MG Stevens[4], Klaus Zuberbühler[2,5] and Florence Levréro[1]

[1]Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Saint-Etienne, France

[2]University of St. Andrews, Department of Psychology & Neuroscience, St. Andrews, Scotland, UK

[3]Université de Lyon/Saint-Etienne, Laboratoire de Biotechnologies Végétales appliquées aux Plantes Aromatiques et Médicinales, FRE-CNRS 3727, Saint-Etienne, France

[4]Royal Zoological Society of Antwerp, Centre for Research and Conservation, Antwerp, Belgium

[5]Université de Neuchâtel, Department of Comparative Cognition, Neuchâtel, Switzerland

MANUSCRIPT 1: Individual vocal signatures across the bonobo' graded repertoire.

Question

Do bonobo vocalisations contain identity information?

Methods

~ **500** hours recording bonobo vocalisations at **3** European zoos

Subjects

♂ *n* = 8

♀ *n* = 13

Five call types collected

High-hoot
n = 333

Bark
n = 431

Soft bark
n = 413

Peep-yelp
n = 420

Peep
n = 253
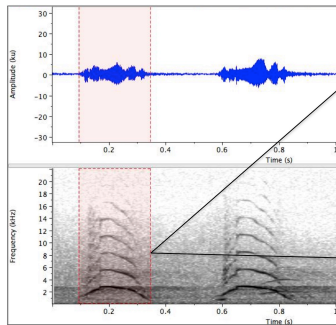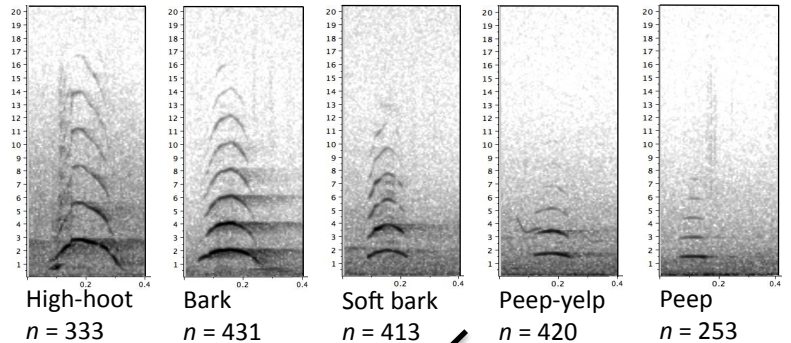
Acoustic analysis – 16 frequency and temporal measurements taken on each.
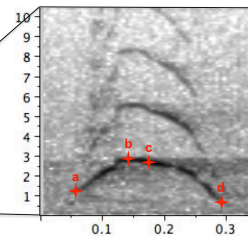
'**Permuted discriminant function analysis (pDFA)**'
&
'**potential for individual coding**' (PIC)
were used to measure individuality.

Results

All investigated bonobo call types contained identity information, but individual vocal signatures were stronger in some call types than in others.

Confusion matrices showing the accuracy of individual classification by pDFA for each call type. All call types were significantly (*p*<0.05) greater than chance, which was 10%.

High-hoot
Classification: **55%**

Bark
Classification: **41%**

Soft bark
Classification: **28%**

Peep-yelp
Classification: **25%**

Peep
Classification: **22%**

Actual individual

Predicted individual
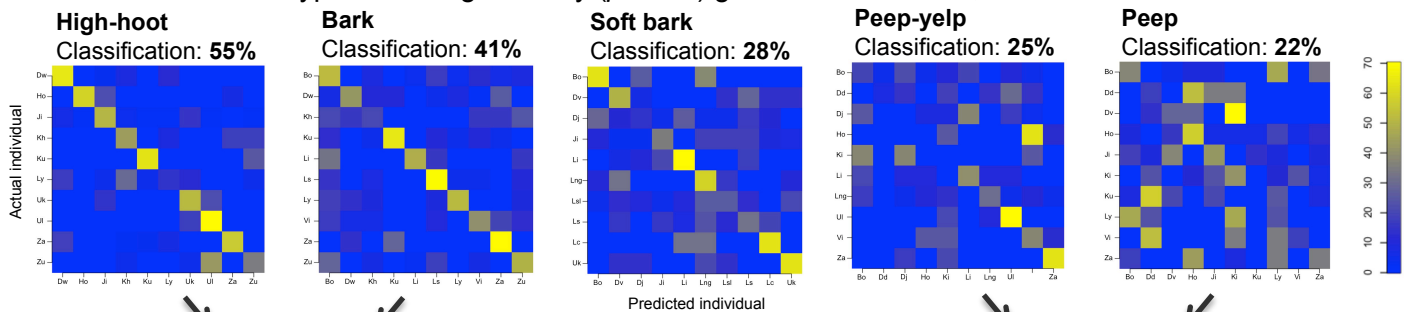
"Loud" call types showed stronger individuality. Such as used during aggressive interactions.

"Close" call types showed weaker individuality. Such as used during grooming.
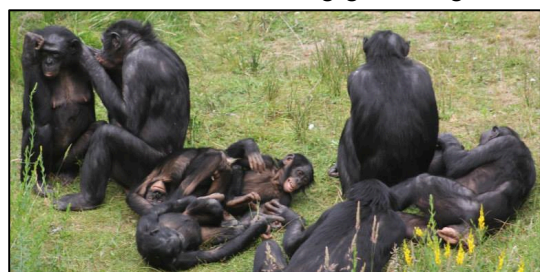
Individual vocal signatures across the bonobos' graded repertoire

Sumir Keenan, Nicolas Mathevon, Florence Nicolè, Jeroen M.G. Stevens, Klaus Zuberbühler and Florence Levréro

ABSTRACT

The vocalizations of a wide range of species allow for individual recognition, a vital process in mediating social interactions. Many hypotheses have proposed which selective pressures acted on the evolution of individual vocal signatures and whether this identity information is stable across a species' repertoire. Two main hypotheses have emerged, one suggesting that calls used to communicate over long distances should have the clearest individual vocal signatures, while the other suggests call types used during important social interactions should demonstrate the strongest signatures. However, only a limited number of studies have investigated these hypotheses in more than a single call type of a species' repertoire. Here we investigate individual vocal signatures in the five most common call types of the bonobo, *Pan paniscus,* graded vocal repertoire. For each call type, we evaluated the reliability of vocal signatures by investigating intra- and inter-individual acoustic variability, as well as using a permuted discriminant function analysis to separate individuals based on their calls' acoustic features. Although all call types demonstrated the potential for identity coding, the strength of these vocal signatures were not consistent across the five call types. Calls used during distance communication and during situations of high excitement or stress had the most distinct individual signatures; while call types used primarily during foraging and grooming had the weakest individual signatures. This suggests that calls used for distance communication evolved stronger levels of individual identity, enabling individual recognition even in situations where the likelihood for signal degradation is high.

However, and importantly, we also found that the five call types were used flexibly and across multiple contexts. We, therefore, argue that multiple selective pressures likely acted concurrently in the evolution identity information in the bonobo repertoire.

INTRODUCTION

Vocalisations are an integral aspect of the social lives of many species and can transmit a wealth of information to listeners about the environment and the individual calling (Seyfarth & Cheney, 2003). Individual differences in the acoustic and temporal structure of species-specific call types, individual vocal signatures, inform on a signaller's identity and have been found in a wide range of mammal and bird species, including in: bats (Voight-Heucke, Taborsky & Dechmann, 2010), hyenas (Mathevon, Koralek, Weldele, Glickman & Theunissen, 2010), seals (Charrier, Mathevon & Jouventin, 2003; Insley, 2000), deer (Vannoni & McElligott, 2007), elephants (Soltis, Leong & Savage, 2005), marmots (Matrosova, Blumstein, Volodin & Volodina, 2011), and a number of birds (e.g. Aubin & Jouventin, 2002; Catchpole & Slater, 2008; Charrier, Jouventin, Mathevon & Aubin, 2001; Dentressangle, Aubin & Mathevon, 2012; Kondo, Izawa & Watanabe, 2010; Vignal, Mathevon & Mottin, 2004) and non-human primates (e.g. Chapman & Weary, 1990; Fischer, Hammerschmidt, Cheney & Seyfarth, 2001; Levréro & Mathevon, 2013; Price, Arnold, Zuberbuhler & Semple, 2009; Wich, Koski, de Vries & van Schaik, 2003). In addition to establishing the presence of individual vocal signatures within a given species, many of these studies have focused on the selective pressures that favoured the evolution of inter-individual differences in vocalisations.

It has long been hypothesised that the evolution of a species' vocal repertoire, including the variability and stability of identity information within vocal signals, is heavily influenced by environmental factors (Marler, 1967, Marler, 1975). The 'distance communication hypothesis' suggests that in low visibility environments, such as forests, vocal signals used to communicate over long distances should have the clearest individual vocal signatures in a repertoire (Mitani, Gros-Louis & Macedonia, 1996). In more recent years, further non-mutually exclusive hypotheses have proposed other factors that have potentially influenced the selection of identity information and acoustic variability in vocal signals, such as

the social function of a call ('social function hypothesis'; Price et al., 2009; Snowdon & Hausberger, 1997), the physiological and behavioural effects a call structure can have on a listener (Owren & Rendall, 2001), individual spacing (Mathevon, Aubin, Vielliard, da Silva, Sebe & Boscolo, 2008), and the size of the social group (Pollard & Blumstein, 2011).

Despite the widespread and long-term interest as to why selection favoured individuality in the vocal systems of so many species, the vast majority of studies only systematically investigated a limited portion of a species' vocal repertoire. To date only six studies have compared more than two call types in any species and all focused on non-human primates (red-capped mangabeys (Bouchet, Blois-Heulin, Pellier, Zuberbuhler & Lemasson, 2012), mouse lemurs (Leliveld, Scheumann & Zimmermann, 2011), Campbell's monkeys (Lemasson & Hausberger, 2011), rhesus macaques (Rendell, Owren & Rodman, 1998) and gorillas (Salmi, Hammerschmidt & Doran-Sheehy, 2014)), except one, which studied a bird species, the South Polar skua (Charrier et al., 2001). These studies have primarily found support for either the distance communication hypothesis or the social function hypothesis, with the exception of Salmi et al. (2014), who found that individuality is equally strongly encoded across all call types in female gorillas.

In the present study, we focus on the bonobo (*Pan paniscus*) vocal repertoire, investigating the reliability of individual vocal signatures between different call types. Two descriptive studies have qualitatively described the bonobo repertoire, both noting that the highly graded call types are used flexibly and not restricted to use in a single context (Bermejo & Omedes, 1999; de Waal, 1998). Despite growing interest in the acoustic communication of this species, our current knowledge of the information content of their calls is limited (Clay & Zuberbuhler, 2009, 2011; Clay et al., 2015; White, Waller, Boose, Merrill & Wood, 2015).

Bonobo society is characterised by female-dominance and complex fission-fusion dynamics, which require sophisticated social knowledge and communication (Clay et al., 2015; Furuichi, 2011), and it has been suggested that individual vocal recognition is an essential tool in enabling successful social navigation (White et al., 2015). The natural habitat of bonobos is comprised of dense African forest, making vocal communication the most effective modality, with evidence that some loud call types can be heard from a distance of 500 metres in the forest (Hohmann & Fruth, 1994). These vocalisations are regularly used by community members to communicate with one another when a group fuses into foraging parties throughout the day and enable the sub-parties to convene around a resource, such as a fruit-bearing tree or nesting tree, despite potential distances between groups (Bermejo & Omedes, 1999; Hohmann & Fruth 1994, 1995; White, 1996; White et al., 2015). Loud calls are also exchanged when different communities encounter one another, and appear to influence inter-community interaction (Furuchi, 2011 & Hohmann & Fruth, 2002). These results suggest that in both intra- and inter-community interactions, individual vocal distinctness would be highly adaptive.

Considering the importance of long distance communication in bonobos, we hypothesised that individual vocal signatures would be more evident in loud vocalisations than in softer ones used for close range communication. Alternatively, descriptive studies suggest that the different call types in the bonobo vocal repertoire are used very flexibly, each one regularly used in multiple contexts (Bermejo & Omedes, 1999; Clay et al., 2015; de Waal, 1988), in contrast to what has been reported for most monkey species (e.g. Gautier & Gautier, 1977). Therefore, it is equally possible that well-defined individuality is a general feature of all bonobo calls, as has been shown in gorillas (Salmi et al., 2014). Here, we tested these two hypotheses by characterising and comparing individual vocal signatures in five different call types, representative of the bonobo's long- and short-range vocal repertoire. Additionally, as previous studies investigating the bonobo vocal repertoire were qualitative, we first

provide a quantitative description of the most common call types found in captive bonobos.

METHODS

Subjects

The bonobos observed for this study are members of three separate captive groups housed in three European zoos: Apenheul (Apeldoorn, the Netherlands), Planckendael Zoo (Mechelen, Belgium) and la Vallée des Singes (Romagne, France; see Supplemental Table 1 for group composition at each zoo. (*Supplemental Tables are following the Reference section of this manuscript*.) At each zoo groups are housed in large indoor enclosures with varying access to off-exhibit rooms and outdoor islands. Vocal recordings and observations were taken from all areas at all three zoos, with the exception of the off-exhibit enclosures at Apenheul.

To avoid the potential confound of comparing adult voices to immature ones, only bonobos over the age of 10 were included in this study. Additionally, one adult male and one female, who had overall low calling rates, were excluded. This led to a total of 21 individuals ranging in age from 10 – 45, with a mean age of 20.5 years old, and comprised of 13 females and 8 males.

Data Collection

Vocal recordings

Vocal recordings were collected between March 2013 and May 2014, beginning no earlier than 8 a.m. and finishing no later than 6 p.m. With a minimum of 115 total recording hours per zoo (Apenheul ~ 175 recording hours; Planckendael ~ 190 recording hours; Vallée des Singes ~ 115 recording hours).

Audio recordings were taken using a Zoom H4 Digital Multi-track Recorder (44.1 kHz sample rate, 16 bits per sample, .wav files) - recording in stereo, with one

channel devoted to a Sennheiser MKH70-1 ultra-directional microphone recording any bonobo vocal behaviour and the second channel connected to a micro-tie recording device, model AKG MPA III, for comments by the researcher. This allowed for temporal synchronising of each vocalisation to information on vocaliser identity and call context as recorded by the researcher.

Acoustic Analysis

Only vocalisations that could unequivocally be assigned to a single caller were retained for analysis. Only calls of good, measureable quality were included, while calls that overlapped with background noise or other calls were removed from the dataset. In total, 1,850 individual calls were retained for analyses (with an average of 88 total calls per individual (N=21), S.D.= 37.87, minimum=45, maximum=227). Raven Pro 1.3 was used to measure automatic and manual parameters on each call. A correlation matrix was produced and very highly correlated variables were removed (0.9 and above), resulting in a total of 16 measurements being retained for analysis – nine manual parameters describing the fundamental frequency and its temporal modulation (measurements were made using Raven's default spectrogram view, with the exception of window size, which was set to 512): F_0-Start (fundamental frequency at beginning of the call, in Hz); F_0-End (fundamental frequency at end of the call, in Hz); F_0-Maximum (highest frequency reached on the fundamental, in Hz); F_0-Maxproportion (point over the duration of the call at which F_0-Maximum is reached, calculated as a proportion: time of F_0-Maximum (s)/Call Duration (s)); Call Duration (s); Ascending Slope (calculated as: F_0-Maximum – F_0-Start/F_0-Maxproportion – 0); Descending Slope (calculated as: F_0-End – F_0-Maximum/1 – F_0Maxproportion); Slope–Start to Midpoint of F_0 (calculated as: F_0 at midpoint of call duration – F_0-Start/Time at midpoint of call duration – 0); and Slope–Midpoint to End (calculated as F_0-End – F_0 at midpoint of call duration/Call duration – time at midpoint of call duration), and seven automatically computed parameters describing the distribution of energy across the frequency spectrum of the entire call: Q1 Frequency (the first quartile, i.e. the frequency at which the

call is divided into two frequency intervals containing 25% and 75% of the call's energy, in Hz); Q1 Time (the first quartile of the call's energy, in seconds); Q3 Frequency (the third quartile of the call's energy, in Hz); Q3 Time (the third quartile of the call's energy, in seconds); Centre Frequency (the median, i.e. the frequency at which the call is divided into two frequency intervals of equal energy, in Hz); Maximum Frequency (the frequency at which the maximum energy occurs in the call, in Hz); and Maximum Time (first time point along the call where maximum amplitude, from waveform, occurs, in seconds) (Figure 1).
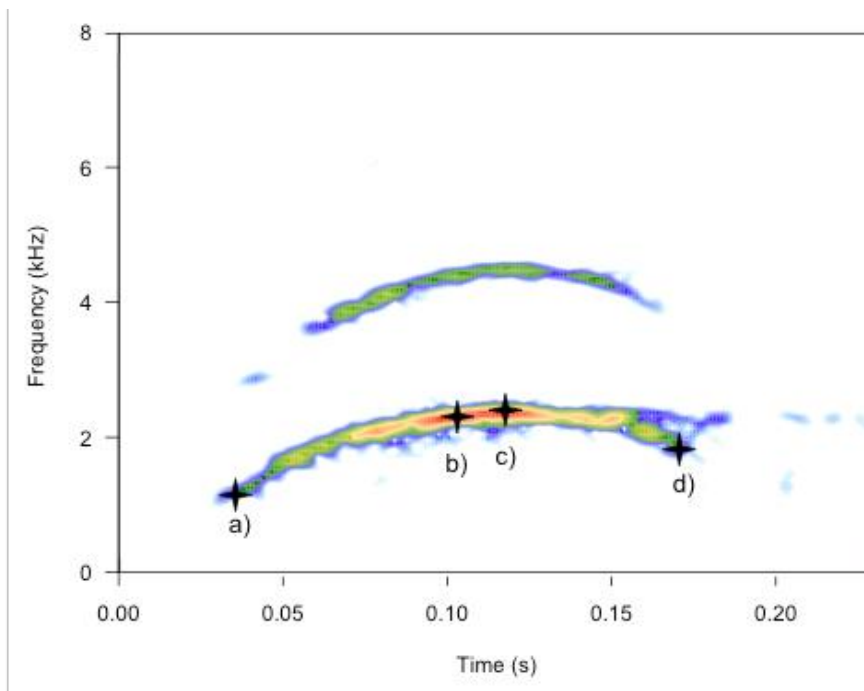
Figure 1. Acoustic and temporal measurements. Example of manual measurements taken on a soft bark call: a= F_0-Start, b=F_0-Midpoint (not included in analysis but used to calculate other measurements), c=F_0-Maximum, d= F_0-End. Call duration=time at b) – time at a).

Call Types

Each call was assigned to one of several types based on classifications described in previous studies (Bermejo & Omedes, 1999; Clay & Zuberbuhler, 2009; de Waal, 1988). For each call type, we calculated its percentage of occurrence in the entire dataset (i.e., relative to the 2,373 measureable calls, of all call types, collected from the 21 adults included in the study). Any call types

that represented less than 10% of the total number of measurable calls collected were excluded from the analysis (the majority of excluded call types were between 0.2 and 2.5% of the dataset, only a single call type neared the 10% threshold at 8%), as these uncommon call types were not produced in sufficient numbers from most individuals. In addition, two call types were excluded from the current study, screams and grunts. Their noisy features, characterized by non-linear phenomena and unclear harmonic structure, distinguish them from the rest of the adult vocal repertoire and would, therefore, require a different set of measurements to describe their acoustic features. As such, five call types, which represented 78% of the total calls collected over the study period (excluding screams and grunts), were retained for analysis: 1) high-hoots (14%; described as 'staccato-high hoots' by de Waal (1988) and 'high hoots' by Bermejo & Omedes (1999)); 2) barks (18%; described as 'barks' and 'wieew barks' by de Waal (1988) and as 'barks' and 'composed barks' by Bermejo & Omedes (1999)); 3) soft barks (17%; described as 'soft barks' by Bermejo & Omedes (1999), 'food barks' by Clay & Zuberbuhler (2009) and not described by de Waal (1988)); 4) peep-yelps (18%; described as 'peep-yelps' by de Waal (1988), Bermejo & Omedes (1999), and Clay & Zuberbuhler (2009)); and 5) peeps (11%; described as 'peeps' by de Waal (1988), Bermejo & Omedes (1999), and Clay & Zuberbuhler (2009)) (see Figure 2 in Results for spectrogram of each call type; see Supplemental Table 2 for full acoustic description of each call type).

Call Contexts

If a caller could be identified, social and individual contextual information was also recorded. Each call was assigned to a single context of the following eight contexts: 1) 'pre-feeding' – calls given directly prior to or at the start of scheduled feedings when group/social excitement was extremely high (excitement described as increased sexual activity, displacements, pacing, pilo-erection, vocal activity, and/or displays); 2) 'feeding' – calls given during scheduled feedings; 3) 'foraging' – calls given when foraging for or eating food

found outside or in the inside enclosure outside of scheduled feeding times; 4) 'aggression' - calls given during agonistic encounters, including mild aggression (no physical contact), aggression (mild physical contact, such as hitting, kicking or grabbing between only two individuals), conflict (a range of physical contact, including biting, and often between more than two individuals) - agonistic interactions were also classified into calls from victims, from aggressors or from bystanders; 5) grooming - calls given during bouts of grooming; 6) contact – calls given when a subject was resting or moving but nothing else; 7) change of environment – calls given when individuals were shifted from one enclosure to another or from inside to outside; and 8) external event – calls given when an individual was visibly startled by or responding to a disturbance external to the group.

There were other contexts that were identified throughout the study that were not included in the following analyses. For example, calls given during tantrums or play were almost exclusively emitted by infants and juveniles, while calls given during evening nesting were only recorded at one zoo (due to restrictions on research hours at two of the three zoos), which prevented a systematic comparison across subjects and therefore were excluded. Additionally, calls given in the contexts of sex and by victims in agonistic encounters were excluded as they largely consisted of screams.

Statistical Analysis

Call type distinctiveness

To confirm whether our dataset could be reliably classified into the five main call types proposed by previous studies, we performed a multivariate analysis to determine the statistical distinction between the 1,850 individual calls (high-hoots: $n=333$, barks: $n=431$, soft barks: $n=413$, peep-yelp: $n=420$, peeps: $n=253$; See Supplemental Table 2 for acoustic description of each call type). The raw values of the 16 acoustical parameters of interest were centred and

normalised by transforming them into z-scores, as the parameter set consisted of different units.

Mundry and Sommer (2007) have convincingly argued that using traditional discriminant function analyses (DFA) when analysing non-independent data (e.g. if the same individual contributed multiple calls) is a case of pseudo-replication and can inflate results. To address this issue, we employed a permuted discriminant function analysis (pDFA, Mathevon et al., 2010; Mundry & Sommer 2007) using the 16 acoustic variables (analysis performed in R, Version 3.2.0). The pDFA takes the classic DFA a step further by comparing the distribution of percent correct classifications obtained after 100 iterations (mean effect size) to the distribution of percent correct values obtained by initially randomly assigning the call type to each individual call (for the current analysis this distribution was obtained via 1000 created data sets where the call type of each call was randomly permutated).

Individual vocal signatures
We assessed the reliability of individual signatures for each call type independently using two different approaches: the first approach used a pDFA (same method as described above, except that calls within each call type were classified according to the identity of the vocaliser), and secondly by calculating the acoustic variation and potential for individual coding (PIC) for each call type overall, as well as for each acoustic parameter describing the calls (Robisson, Aubin & Bremond, 1993).

For the pDFA analysis, we retained individuals who had a minimum of 14 calls for each call type (with the exception of one individual who had 13 calls in the peep call type) (See Supplemental Table 1 for the number of calls contributed by each individual for each call type). Not every individual had a sufficient number of calls for each call type to be included in all five call type analyses. The peep call type had the fewest number of contributing individuals (n=10);

therefore, to allow for direct comparison across call types, we randomly chose ten individuals for each of the other four call types (however, balancing the data for individual sex and group), with the majority of individuals contributing to 2 or 3 call type analyses (mean = 2.5, max = 4, min = 1; Supplemental Table 1).

For the acoustic variation and PIC analyses, we included all individuals with 12-20 calls for each call type. For individuals who had more than 20 calls in a given call type, we randomly chose calls to be excluded (however ensuring where possible that a single calling event or calls given in a single day were not over-represented). This ensured that no single individual was overrepresented, as no individual contributed more than 20 calls for each call type.

The amount of variability across the five call types was determined by calculating the inter- and intra- individual coefficient of variation (CV). The CV values were then used to calculate the potential of individual coding (PIC) (according to Robisson et al., 1993). Additionally, recent studies investigating individual distinctiveness in non-human primates have calculated CV and PIC values, including them here enables direct comparison (Bouchet et al., 2012; Lemasson & Hausberger, 2011; Salmi et al., 2014). As the coefficient of variation (CV) can only be calculated with variables on a ratio scale, two variables (that were on an interval scale) were not included (Slope from Start to midpoint of F0 and Slope from Midpoint to end of F0). Therefore, for each of the remaining 14 parameters in each call type separately we first calculated the CV(Inter) = (100 X S.D. calculated across all individuals) / (Mean (calculated across all individuals)) and the CV(Intra) = the mean of individual CV values, where the CV (= 100 X S.D./Mean) was calculated for each individual separately. PIC values could then be obtained for each parameter (PIC = CV(Inter)/ CV (Intra)). PIC values over 1 indicate a potential to code for individual information. To assess the variation of the five call types as a whole, we took the mean of the CV(Inter) and the CV(Intra) across the 14 parameters for each call type separately. These means were then used to calculate the PIC values for each

call type. To test for call type differences in the levels of variation and potential for individual coding, Friedman tests were done with the CV and PIC values for each parameter for each call type respectively.

We recognise that assessing individuality by using coefficient of variations and potential for individuality is arguably an outdated method, and therefore we have based the backbone of our conclusions on the pDFA results. However, the very limited numbers of studies that have investigated individual vocal signatures in multiple call types have all used CV and PIC analysis, therefore to allow for direct comparison with these studies we have also included these results.

RESULTS

Call types – Classification and context
The results of the pDFA showed that the acoustic structure of calls supports the initial, human-driven, classification into call types, with an accuracy of 57% (chance level of 20%), which was significant at p=0.001 (Figure 2).
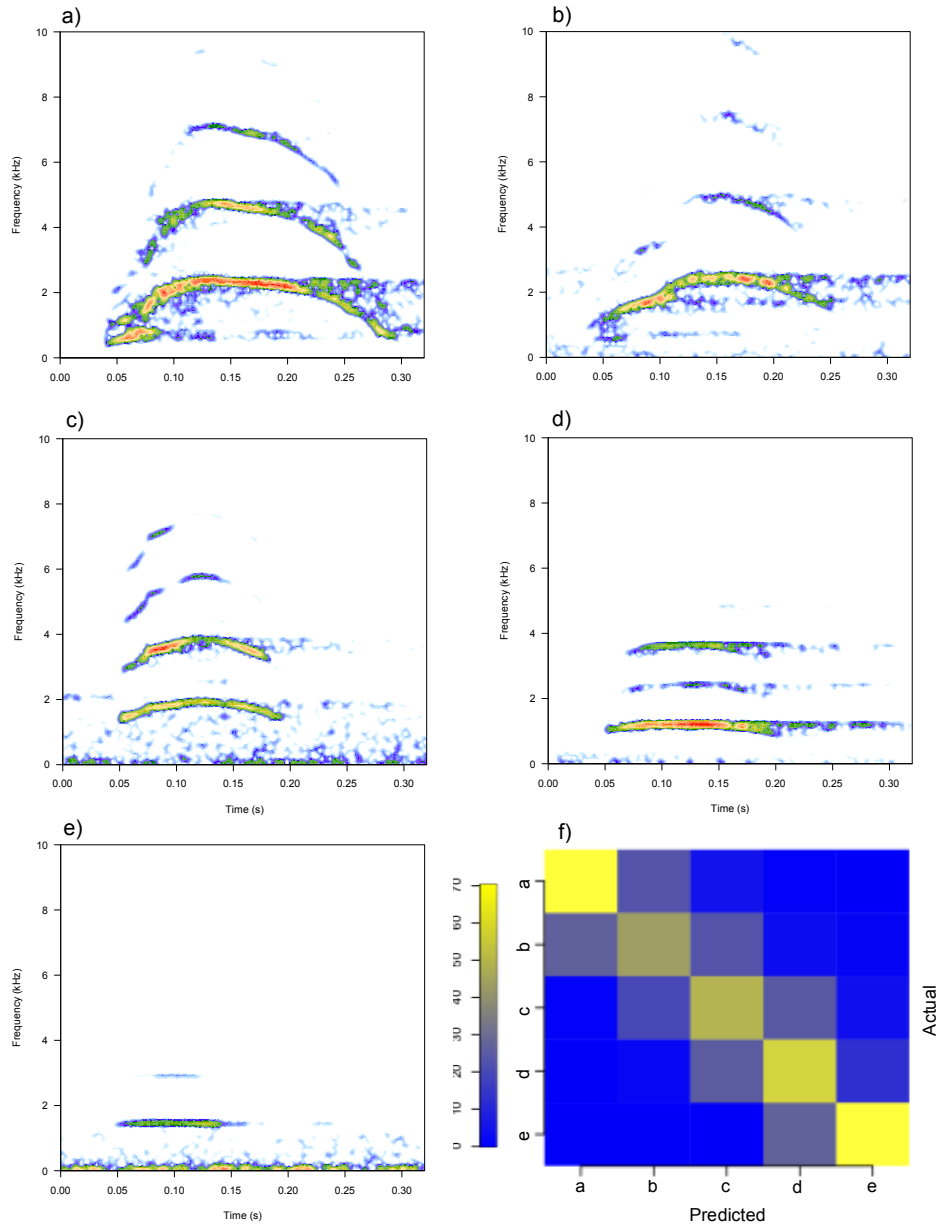
Figure 2. Five call types. Spectrographic illustrations of the five most common call types of the bonobo repertoire. Spectrogram images depict: a) High-hoot; b) Bark; c) Soft bark; d) Peep-yelp; e) Peep. Panel f) shows the confusion matrix obtained from the permuted DFA classifying the five call types, which are labelled as just listed. On the confusion matrix, the diagonal shows the rate at which a call type was correctly assigned – specifically, what percentage the actual call type and the predicted call type match. The brightness of each diagonal yellow square corresponds to the strength of classification. The off-diagonal cells show percentage of misclassification for each call type and which other call type they were misclassified as.

A scatterplot using discriminant function 1 and 2 to visualise the grouping of the call types, fails to show any distinct boundaries, or a graded structure, between the five call types (Figure 3).

	Function 1	Function 2
F_0-Start	-0.121	**0.436**
F_0-End	-0.015	0.146
F_0-Maximum	**0.616**	0.024
F_0-Maximum proportion	0.076	**0.293**
Call Duration	**0.445**	-0.003
Ascending Slope	**0.59**	-0.452
Descending Slope	**-0.612**	0.151
Slope-Start to Midpoint	**0.543**	-0.183
Slope-Midpoint to End	**-0.467**	-0.166
Q1 Frequency	0.244	0.111
Q1 Time	0.195	0.068
Q3 Frequency	0.299	-0.012
Q3 Time	0.34	-0.044
Center Frequency	0.303	0.006
Maximum Frequency	0.214	0.057
Maximum Time	0.238	0.019

Figure 3. Scatterplot of call types. Scatterplot demonstrating the graded nature of the acoustic and temporal properties of the five call types. Discriminant function 1 and 2 generated from the permuted DFA on the five call types, legend shows factor loadings on each function of 16 acoustic and temporal variables. White asterisks mark the centroid of each call type.

Despite the gradedness of the five call types, overall the call types were used significantly differently in each of the observed contexts ($\chi 2$ (28) = 930.281, n=1850, $p < 0.001$; Table 1), and post-hoc tests revealed that each call type was used significantly differently from each other call type (see Supplemental Table 3 for details). The high-hoots were primarily used during times of social excitement, when the potential for conflict was high or when a conflict was already occurring and included during: pre-feeding, changes in environment, external events (or alarm), and by aggressors in agonistic encounters. Barks were given in similar contexts to high-hoots, however, with a marked increase of calls being given during feeding events. Soft barks were given mostly in

association with feeding and foraging, but also during pre-feeding events. Peep-yelps and peeps were used similarly with the main difference being an increased usage of peeps during grooming and contact and a decreased usage during feeding or foraging (Table 1).

	Percentage of calls given in each context				
	High-hoot	Bark	Soft Bark	Peep-yelp	Peep
Aggression	23.12	19.03	2.66	0.00	0.00
Change in environment	18.02	13.46	9.93	2.86	1.58
External Event (Alarm)	13.81	6.96	2.18	2.14	0.79
Pre-feeding	**38.74**	**34.34**	21.55	17.38	21.74
Feeding	1.80	18.10	**38.26**	**36.19**	**28.46**
Forage	4.50	3.25	18.40	15.95	9.09
Groom	0.00	0.93	1.45	6.43	9.88
Contact-general	0.00	3.94	5.57	19.05	**28.46**

Table 1. Call type usage. The percentage of each call type used in each context. The two contexts in which each call type is most commonly used are highlighted, with the bold number being the most common.

Individual vocal signature

Five separate pDFA analyses were run on each call type to evaluate the level of individual distinctiveness within each call type. Individuals were classified correctly 55% of the time for high-hoots (chance level = 10%; $p = 0.001$), 41% for barks (chance level = 10%; $p = 0.001$), 28% for soft barks (chance level = 10%; $p = 0.001$), 25% for peep-yelps (chance level = 10%; $p = 0.004$) and 22% for peeps (chance level = 10%; $p = 0.006$) (all percentages were from cross-validated data sets; Figure 4). While all call types showed individual

classification at a rate significantly greater than chance, there is a clear pattern of decreased classification as calls increasingly become more 'soft' or 'quiet'.
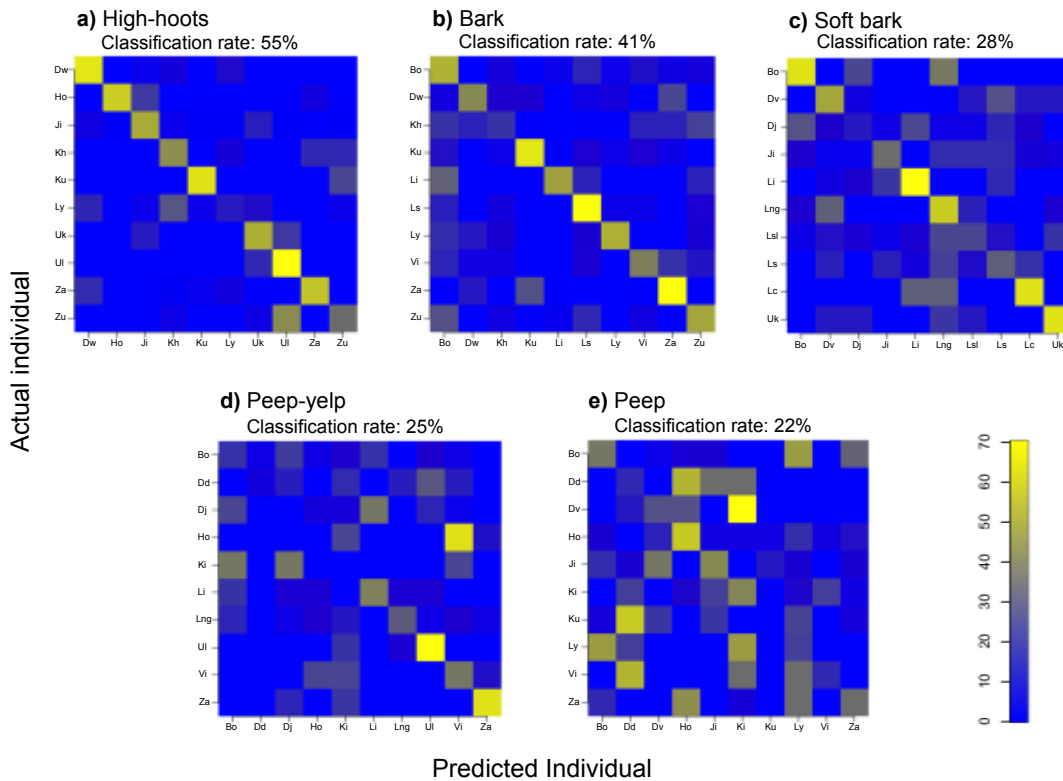
Figure 4. Individual vocal signatures in five different call types. Each confusion matrix shows the results of five separate permuted DFAs investigating the strength individuality in each call type. The confusion matrix shows the probability that an individual's calls were correctly classified – the legend shows the percent accuracy for each individual. (Random classification at 10%.) (Details on each individual, age, sex, rank, zoo and number call contributed to each analysis can be found in Supplementary Table 1.)

Variability and vocal signature

The level of acoustic variability significantly differed between the call types. The level of intra-individual variation (measured by CV(Intra)(χ^2 (13)=55.926, p<0.001)) was highest for peeps (58%) and decreased along the repertoire largely in line with the pDFA classification results: peep-yelps (41%), soft barks (36%), high-hoots (31%), and barks (28%). The level of inter-individual variation (measured by CV(Inter)(χ (χ^2 (13)=54.623, p<0.001) followed the same pattern:

peeps (60%), peep-yelps (45%), soft barks (39%), high-hoots (38%), and barks (33%). All of the call types had PIC values over 1, indicating that all five have some capacity to code for individuality, however the PIC was significantly stronger in some call types (χ^2 (13)=49.114, p<0.001). PIC was highest in high-hoots (1.22), followed by barks (1.18), peep-yelps (1.10), soft barks (1.08) and peeps (1.03) (Table 2). Therefore, we can conclude that both intra- and inter-individual variation was highest and the potential for individual coding was lowest in calls used as contact calls, and both measures were intermediary in calls emitted during feeding and foraging. Variation was lowest, and potential for individual coding highest, in loud call types (high-hoots and barks) predominantly used during aggressive encounters, group tension (high potential for aggression) and when social excitement (as previously described) was high.

We then used this information to investigate which of the studied acoustic parameters had the highest potential for individual coding. When looking across all call types, all of the parameters had PIC values over 1, with the onset frequency of the fundamental frequency (F_0-Start) (1.20) and the ascending slope (1.22) having the highest PIC values (Table 2). However, when investigating each call type separately, these two parameters did not always have the highest potential for individual coding. The onset frequency of the fundamental frequency (F_0-Start) had the highest PIC in both the peep and barks, while the maximum frequency reached on the fundamental frequency (F_0-Maximum) had the highest PIC in peep-yelps and high-hoots and the call duration and the frequency of the fundamental at the end of the call (F_0-End) equally had the highest PIC for soft barks. Additionally, for each call type separately, not all acoustic parameters reached the minimum threshold (PIC≤1) of identity coding (Table 2).

	PIC values by *Call Types*					
	Overall PIC value for each parameter	Peep	Peep-yelp	Soft bark	Bark	High-hoot
Individual Parameters						
F_0-Start	1.20	1.23	1.24	1.12	1.29	1.38
F_0-End	1.11	1.20	1.18	1.20	1.19	1.43
F_0-Maximum	1.20	1.21	1.25	1.12	1.25	1.46
F_0-Maximum proportion	1.08	0.91	1.03	1.04	1.06	1.18
Call Duration	1.16	1.05	1.13	1.20	1.19	1.28
Ascending Slope	1.22	0.97	1.15	1.06	1.19	1.14
Descending Slope	1.03	1.06	1.17	1.08	1.02	1.21
Q1 Frequency	1.17	1.17	1.11	1.12	1.17	1.38
Q1 Time	1.07	0.99	1.01	1.03	1.15	1.06
Q3 Frequency	1.14	1.17	1.08	1.18	1.23	1.34
Q3 Time	1.09	1.00	1.05	1.06	1.12	1.14
Center Frequency	1.12	1.15	1.11	1.10	1.22	1.36
Maximum Frequency	1.08	1.12	1.08	1.08	1.11	1.28
Maximum Time	1.10	1.02	1.03	1.03	1.10	1.08
Overall PIC for call type		1.03	1.10	1.08	1.18	1.22

Table 2. Potential for individual coding (PIC). Investigating the potential for individuality in each call type as well as each parameter across all call types and within each call type separately. A PIC greater than or equal to 1 suggests potential for individuality. For each call type the parameter with the highest PIC is highlighted.

DISCUSSION

Our study demonstrated that all five common call types within the bonobo repertoire showed significant individual vocal signatures. However, the level of individual distinctiveness differed across call types. Loud calls, such as high-hoots and barks, used in the wild for distance communication (Bermejo & Omedes, 1998; Hohmann & Fruth, 1994, 1995; Mori, 1983) and during situations of high excitement or stress (both in the wild and in captivity), showed the strongest vocal signatures. These two call types both had the highest call type accuracy, the highest individual accuracy and the highest potential for

individual coding, which should allow identity information to propagate long distances through forests and reach remote conspecifics. On the other hand, the quieter close calls, peeps and peep-yelps, used mostly during foraging, contact or grooming, had lower potential for individuality and the lowest individual classification accuracy. The weaker identity information in the peep and peep-yelp call types would make them unsuitable for distance communication, as they would likely suffer from propagation-induced degradation, as has been shown in birds (Mouterde, Theunissen, Elie & Mathevon, 2014). Nonetheless, in a recent experimental study, we demonstrated that bonobos could vocally recognise familiar individuals from peep-yelps (Keenan, Mathevon, Stevens, Guery, Zuberbuhler & Levrero, 2016), suggesting that all call types, even the ones with statistically low individual coding, contain enough identity information for recognition, at least at short-range distances.

At first sight, it appears that the vocal signatures present in the bonobo repertoire have evolved in line with predictions made by the 'distance communication hypothesis' (Marler, 1967), as the loud calls had clearer and more pronounced individual distinctiveness. This suggests that calls used for distance communication evolved stronger levels of individual identity, potentially enabling individual recognition even in situations where the likelihood for signal degradation is high. Additionally, this suggests that being able to identify individuals when communicating at a distance, both within a community and with other communities, is important for these social interactions. This is in line with other species that live in fission-fusion societies in dense rain forest environments, including the bonobos' congener, the chimpanzee (Mitani et al., 1996).

Confirming previous research on individual call types (e.g. Clay et al., 2015), all five call types were used flexibly between contexts and often in direct social interactions. As such, and according to the social function hypothesis, we

predicted to see little difference in the strength of vocal signatures across call types. Our data did not support this hypothesis, suggesting that social function does not affect individual distinctiveness, at least in bonobos. However, it is difficult to distinguish within the bonobo repertoire which call types or contexts should be considered less social than others. The strongest support for the social function hypothesis has come from a few monkey species (e.g. Bouchet et al., 2012; Lemasson & Hausberger, 2011), where functionally referential alarm calls show reduced levels of individual distinctiveness, arguably because the most important encoded information is predator type and urgency but not necessarily the vocaliser's identity. Research within great apes has thus far not produced any evidence of predator-specific alarm calls; therefore, to accurately test the social function theory in bonobos, we would first need to define what is considered less social vocal behaviour. Alternatively, we can consider that these two hypotheses concurrently impacted the evolution of individual distinctiveness, as the call types with higher vocal signatures are most often used in distance communication and in contexts of high social importance. For example, the high-hoots and barks likely have multiple social functions, during close aggressive interactions these calls could function to warn or intimidate a victim who is in very close, visual range and at the same time be broadcasting to recruit individuals outside visual range, such has been found in chimpanzees (Slocombe & Zuberbuhler, 2007).

A recent study investigating individual distinctiveness in both close and long range call types of female western gorillas (Salmi et al., 2014) found, as in our data and the two monkey studies (e.g. Bouchet et al., 2012; Lemasson & Hausberger, 2011), that all eight call types analysed had potential for individual coding (PIC > 1). However, while the PIC values were in a similar range for both the gorilla call types (PIC between 1.07 – 1.23; Salmi et al., 2014) and the bonobo call types (PIC between 1.03 – 1.22; this study), the pDFA results were dramatically different, with vocal individuality correctly classified at similar rates across call types in gorillas (between 74-85%, chance level 30%; Salmi et al.,

2014) and at very different rates across bonobo call types (22-55% chance level 10%, this study). Overall, Salmi et al. (2014), as opposed to the results of this study, found no significant difference between the potential for individual coding or the pDFA classification rates between call types, and thus concluded that neither the social function or the distance hypotheses accurately reflected the evolutionary pressures acting on female gorilla calls. Their results are in contrast to results found in the two monkey species, Campbell's monkeys (Lemasson & Hausberger, 2011) and Red-capped mangabeys (Bouchet et al., 2012). When comparing three monkey species, Bouchet et al. (2013) argued further support for this hypothesis, as across the three studied species (Campbell's monkeys, De Brazza's monkeys and red-capped managbeys) all demonstrated the highest levels of individuality in their contact calls when compared to alarm calls or threat calls. However, other differences in identity coding between the three species call types, suggested that the complexity of the species' social structure greatly influenced the complexity of the species' repertoire, including the stability of identity information across call types.

Salmi et al. (2014) suggested that the difference between gorillas and monkeys could be due in part to the gradedness of the female gorilla vocal repertoire. However, the bonobo call types examined here also show a graded call system, with intra- and inter-individual variation similar to what was found in gorillas, specifically when compared to the two monkey species (Intra: bonobo = 28-58% (this study), gorilla = 35-56% (Salmi et al., 2014), Red-capped mangabey = 11-31% (Bouchet et al., 2012), Campbell's monkey = 20-50% (Lemasson & Hausberger, 2011); Inter: bonobo = 33-60% (this study), gorilla = 40-60% (Salmi et al., 2014), Red-capped mangabey = 20-50% (Bouchet et al., 2012), Campbell's monkey = 45-227% (Lemasson & Hausberger, 2011) suggesting that a graded call repertoire likely does not dictate the stability of individual distinctiveness across call types.

Despite some clear differences, the individual vocal signatures of bonobos are in many ways the most similar to the results found in gorillas. This is not surprising as both apes show a graded vocal repertoire, live in forest environments and have repertoires lacking the functionally referential alarm calls that characterise the vocal systems of many monkey species. Alternatively, the distinct differences in the social structures of the two apes (largely stable polygynous groups in gorillas (Robbins, Bermejo, Cipolletta, Magliocca, Parnell, & Stokes, 2004) and flexible fission-fusion societies in bonobos (Kano, 1992) may account for the differences found in their individual vocal distinctiveness.

In conclusion, a surface inspection of our data suggests that the vocal signatures present in the bonobo repertoire have evolved in line with the distant communication hypothesis, as the loud calls have more clear and pronounced individual distinctiveness. However, as all of our call types demonstrated some level of individuality the distance communication hypothesis alone does not fully explain individuality in the bonobo repertoire. When comparing our results with other studies it appears to confirm that no one theory alone can explain the individual distinctiveness found in a repertoire, but what is likely occurring is a combination of several selective influences acting concurrently. We could argue that variable individual distinctiveness in a species repertoire is influenced largely by social structure, and in extension to that: whether individuals need to communicate with one another over large distances, if groups are territorial and, importantly, if the species is regularly preyed upon. As such the ubiquity of individuality across the bonobo vocal repertoire may result from their fission-fusion social dynamics, graded vocal system, the specific usage of some call types and the dense forest environment, which puts additional pressure on call types that function to send accurate individual information over long distances.

Ethics Statement:

All work was performed in accordance with the relevant guidelines and regulations, and all experimental protocols were approved by the Institutional Animal Ethical Committee of the University of Lyon/Saint-Etienne, under the authorization no. 42-218-0901-38 SV 09 (Lab ENES).

REFERENCES

Aubin, T. & Jouventin, P. (2002). How to identify vocally a kin in a crowd? The penguin model. *Advances in the Study of Behavior, 31*, 243-277. doi: 10.1016/S0065-3454(02)80010-9

Bergman, T. & Sheehan, M. (2013). Social knowledge and signals in primates. *American Journal of Primatology, 75*, 683-694. doi:10.1002/ajp.22103

Bermejo, M. & Omedes, A. (1999). Preliminary Vocal Repertoire and Vocal Communication of Wild Bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica, 70*, 328-357. doi:10.1159/000021717

Bouchet, H., Blois-Heulin, C., Pellier, A., Zuberbuhler, K. & Lemasson, A. (2012). Acoustic variability and individual distinctiveness in the vocal repertoire of red-capped mangabeys (*Cercocebus torquatus*). *Journal of Comparative Psychology*, *126*, 45-56. doi:10.1037/a0025018

Bouchet, H., Blois-Heulin, C. & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 1-15. doi:10.3389/fpsyg.2013.00390

Catchpole, C. K. & Slater, P.J.B. (2008). Bird song: Biological themes and variations (2nd ed.). Cambridge, U.K.: Cambridge University Press.

Charrier, I., Jouventin, P., Mathevon, N. & Aubin, T. (2001). Individual identity coding depends on call type in the South Polar skua, *Catharacta maccormicki*. *Polar Biology*, *24*, 378-382. doi:10.1007/s003000100231

Charrier, I., Mathevon, N. & Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Animal Behaviour*, *65*, 543-550. doi: 10.1006/anbe/2003.2073

Chapman, C.A. & Weary, D.M. (1990). Variability in spider monkeys' vocalizations may provide basis for individual recognition. *American Journal of Primatology, 22*, 279-284. doi:10.1002/ajp.1350220407

Clay, Z. & Zuberbuhler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour, 77*, 1387-1396. doi:10.1016/j.anbehav.2009.02.016

Clay, Z. & Zuberbuhler, K. (2011). Bonobos extract meaning from call sequences. *PloSOne, 6*, e18786. doi:10.1371/journal.pone.0018786

Clay, Z., Archbold, J. & Zuberbuhler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ, 3*, e1124. doi: 10.7717/peerj.1124

Dentressangle, F., Aubin, T, & Mathevon, N (2012). Males use time whereas females prefer harmony: individual call recognition in the dimorphic blue-footed booby. *Animal Behaviour*, *84*, 413-420. doi: 10.1016/j.anbehav/2012/05/012

de Waal, F.B.M. (1988). The communicative repertoire of captive bonobos (*Pan paniscus*) compared to that of chimpanzees. *Behaviour, 106*, 183-251. doi: 10.1163/156853988X00269

Fischer, J., Hammerschmidt, K., Cheney, D. L. & Seyfarth, R.M. (2001). Acoustic features of female chacma baboon barks. *Ethology, 107*, 33-54. doi:10.1046/j.1439-0310.2001.00630.x

Furuchi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthrophology* 20, 131-142. doi: 10.1002/evan.20308

Gautier, J.P. & Gautier, A. (1977). Communication in old world monkeys. In: T.A. Sebeok (Ed.), *How animals communicate* (pp. 890-964). Bloomington, Indiana: Indiana University Press.

Hohmann, G. & Fruth, B. (1994). Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology 15*, 767–782. doi:10.1007/BF02737430

Hohmann, G., & Fruth, B. (1995). Loud calls in great apes: sex differences and social correlates. In: E. Zimmermann (Ed.), *Current Topics in Primate Vocal Communication* (pp.161–184). New York, New York: Plenum Press.

Hohmann, G. & Fruth, B. (2002). Dynamics is social organization of bonobos (*Pan paniscus*). In: Boesch, C., Hohmann, G. & Marchant, L.F. (Eds.) *Behavioural Diversity in Chimpanzees and Bonobos* (pp.138-150). Cambridge, U.K.: Cambridge University Press.

Insley, S. J. (2000). Long-term vocal recognition in the northern fur seal. *Nature*, *406*, 404–5. doi:10.1038/35019064

Kanō, T. (1992). *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford, California: Stanford University Press.

Keenan, S., Mathevon, N., Stevens, J.M.G., Guery, J.P., Zuberbuhler, K. & Levrero, F. (2016). Enduring voice recognition in bonobos. *Scientific Reports, 6*, 22046. doi: 10.1038/srep22046.

Kondo, N., Izawa, E., & Watanabe, S. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. *Behaviour*, *147*, 1051–1072. doi:10.1163/000579510X505427

Leliveld, L.M.C., Scheumann, M. & Zimmermann, E. (2011). Acoustic correlates of individuality in the vocal repertoire of a nocturnal primate (*Microcebus murinus*). *The Journal of the Acoustical Society of America, 129*, 2278-2288. doi:10.1121/1.3559680

Lemasson, A. & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *The Journal of the Acoustical Society of America, 129*, 3341-3352. doi:10.1121/1.3569704s

Levréro, F. & Mathevon, N. (2013). Vocal signature in wild infant chimpanzees. *American Journal of Primatology, 75*, 324-332. doi:10.1002/ajp.22108

Marler, P. (1967). Animal Communication Signals. *Science, 157*, 769-774. doi:10.1126/science.157.3790.769

Marler, P. (1975). On the origin of speech from animal sounds. In: J.F. Kavanagh & J. Cutting (Eds.), *The Role of Speech in Language* (pp. 11-37). Cambridge, Massachusetts: MIT Press.

Mathevon, N., Aubin, T., Vielliard, J., da Silva, M.L., Sèbe, F. & Boscolo. D. (2008). Singing in the rain forest: How a tropical bird song transfers information. *PlosONE, 3*, e1580. doi: 10.1371/journal.pone.0001580

Mathevon, N., Koralek, A., Weldele, M., Glickman, S.E. & Theunissen, F.E. (2010). What the hyena's laugh tells us: sex, age dominance and individual

signature in the giggling call of *Crocuta crocuta*. *BMC Ecology, 10*. doi:10.1186/1472-6785-10-9

Matrosova, V.A., Blumstein, D.T., Volodin, I.A. & Volodina, E.V. (2011). The potential to encode sex, age, and individual identity in the alarm calls of three species of Marmotinae. *Naturwissenschaften, 98*, 181–192. doi:10.1007/s00114-010-0757-9

Mori, A. (1983). Comparison of the communicative vocalizations and behaviors of group ranging in Eastern gorillas, chimpanzees and pygmy chimpanzees. *Primates, 24*, 486-500. doi :10.1007/BF02381682

Mitani, J. C., Gros-Louis, J., & Macedonia, J. M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology*, *17*, 569–583. doi:10.1007/BF02735192

Mouterde, S., Theunissen, F., Elie, J.E. & Mathevon, N. (2014). Learning to cope with degraded sounds: Female zebra finches can improve their expertise at discriminating between male voices at long distance. *Journal of Experimental Biology*, 217, 3169-3177. doi: 10.1242/jeb.104463

Mundry R. & Sommer C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour, 74*, 965-976. doi:10.1016/j.anbehav.2006.12.028

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, *10*, 58–71. doi:10.1002/evan.1014

Pollard, K.A. & Blumstein, D.T. (2011). Social group size predicts the evolution of individuality. *Current Biology, 21*, 413-417. doi:10.1016/j.cub.2011.01.051

Price, T., Arnold, K., Zuberbuhler, K. & Semple, S. (2009). Pyow but not hack calls of the male putty-nosed monkey (*Cercopithcus nictitans*) convey information about caller identity. *Behaviour, 146*, 871-888. doi:10.1163/156853908X396610

Rendall, D., Owren, M. J., & Rodman, P. S. (1998). The role of vocal tract filtering in identity cueing in rhesus monkey (Macaca mulatta) vocalizations. *Journal of Acoustic Society of American*, *103*(1), 602–614.

Robbins, M. M., Bermejo, M., Cipolletta, C., Magliocca, F., Parnell, R.J. & Stokes, E. (2004). Social Structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*), *American Journal of Primatology, 64*, 145-159. doi:10.1002/ajp.20069

Robisson, P., Aubin, T. & Bremond, J. C. (1993). Individuality in the voice of the emperor penguin, *Aptenodytes forsteri*: Adaptation to a noisy environment. *Ethology*, *94*, 279-290. doi:10.1111/j.1439-0310.1993.tb00445.x

Salmi, R., Hammerschmidt, K. & Doran-Sheehy, D.M. (2014). Individual distinctiveness in call types of wild western female gorillas. *PLoSONE, 9*, e101940. doi:10.1371/journal.pone.0101940

Seyfarth, R.M. & Cheney, D.L. (2003). Signalers and receivers in animal communication *Annual Review of Psychology, 54*, 145-173. doi:10.1146/annurev.psych.54.101601.145121

Slocombe, K. E. & Zuberbuhler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *PNAS, 104*, 17228-17233. doi: 10.1073/pnas.0706741104

Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication II: Rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour*, *70*, 589–599. doi:10.1016/j.anbehav.2004.11.016

Snowdon, C.T & M. Hausberger (1997). *Social influences on vocal development*. Cambridge, U.K.: Cambridge University Press.

Vannoni, E. & McElligott, A. G. (2007). Individual acoustic variation in fallow deer (Dama dama) common and harsh groans: A source-filter theory perspective. *Ethology*, *113*, 223–234. doi:10.1111/j.1439-0310.2006.01323.x

Vignal, C., Mathevon, N. & Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature*, *430*, 448-451. doi: 10.1038/nature02645

Voigt-Heucke, S. L., Taborsky, M. & Dechmann, D.K.N. (2010). A dual function of echolocation: bats use echolocation calls to identify familiar and unfamiliar individuals. *Animal Behaviour*, *80*, 59–67. doi:10.1016/j.anbehav.2010.03.025

White, F. J. (1996). Comparative socio-ecology of *Pan paniscus*. In W.C. McGrew, L.F. Marchant & T. Nishida (Eds.), *Great Ape Societies* (pp. 29–41). Cambridge, U.K.: Cambridge University Press.

White, F. J., Waller, M., Boose, K., Merrill, M.Y. & Wood, K.D. (2015). Function of loud calls in wild bonobos. *Journal of Anthropological Sciences, 93*, 1-13. doi:10.4436/JASS.93003

Wich, S.A., Koski, S., de Vries, H. & van Schaik, C.P. (2003). Individual and contextual variation in Thomas langur male loud calls. *Ethology 109*, 1-13. doi:10.1046/j.1439-0310.2003.00837.x

SUPPLEMENTAL TABLES

Supplemental Table 1. Group composition and numbers of calls contributed per individual. ✓Indicates individuals included in pDFA analysis**.** *Indicates individual included in PIC and variability analysis. ᵒIndicates individuals who were not included in the study, which was all individuals under the age of 10 and an adult male and female both with low calling rates.

Name	Name Abbr.	Sex	Age	Rank	Zoo	Bark	High -hoot	Peep	Peep -yelp	Soft bark
Hortense	Ho	F	35	High	Apen	*15	✓*18	✓*32	✓*15	11
Jill	Ji	F	28	High	Apen	*55	✓*52	✓*25	*51	✓*44
Zuani	Zu	F	23	High	Apen	✓*26	✓*45	0	*15	*26
Bolombo	Bo	M	16	Low	Apen	✓*37	0	✓*25	✓*29	✓*15
Zamba	Za	M	15	Mid	Apen	✓*14	✓*48	✓*18	✓*23	*18
Kumbuka	Ku	F	14	Mid	Apen	✓*32	✓*19	✓*27	*12	*20
Besedeᵒ	n/a	F	8	Low	Apen	n/a	n/a	n/a	n/a	n/a
Yahimbaᵒ	n/a	F	4	n/a	Apen	n/a	n/a	n/a	n/a	n/a
Makasiᵒ	n/a	M	4	n/a	Apen	n/a	n/a	n/a	n/a	n/a
Monyamaᵒ	n/a	F	3	n/a	Apen	n/a	n/a	n/a	n/a	n/a
Lina	Li	F	28	High	Planck	✓*17	2	6	✓*21	✓*28
Vifijo	Vi	M	19	Low	Planck	✓*27	0	✓*14	✓*19	6
Djanoa	Dj	F	18	Mid	Planck	6	11	5	✓*23	✓*32
Louisoko	Ls	M	15	Mid	Planck	✓*32	6	3	10	✓*26
Lucuma	Lc	M	11	Mid	Planck	*13	0	3	*15	✓*14
Busiraᵒ	n/a	F	10	Low	Planck	n/a	n/a	n/a	n/a	n/a
Habariᵒ	n/a	M	7	Low	Planck	n/a	n/a	n/a	n/a	n/a
Lingoyeᵒ	n/a	F	8	Mid	Planck	n/a	n/a	n/a	n/a	n/a
Nayokiᵒ	n/a	F	2	n/a	Planck	n/a	n/a	n/a	n/a	n/a
Daniela	Dd	F	45	High	VDS	1	0	✓*14	✓*38	*35
Lisala	Lsl	F	33	Mid	VDS	2	4	3	7	✓*35
Ukela	Uk	F	28	High	VDS	*14	✓*18	1	10	✓*19
Bondoᵒ	n/a	M	22	Low	VDS	n/a	n/a	n/a	n/a	n/a
Kirembo	Kir	M	21	Low	VDS	2	2	✓*22	✓*15	8
Ulindi	Ul	F	20	Mid	VDS	*24	✓*17	8	✓*16	*14
David	Dv	M	12	Mid	VDS	5	9	✓*15	*29	✓*19
Diwani	Dw	M	17	Mid	VDS	✓*50	✓*23	5	9	8
Khaya	Kh	F	12	Mid	VDS	✓*24	✓*17	8	*17	6
Lucy	Ly	F	10	Low	VDS	✓*22	✓*30	✓*13	*17	7
Lingala	Lng	F	10	Mid	VDS	*13	*12	6	✓*29	✓*22
Keleleᵒ	n/a	M	9	Low	VDS	n/a	n/a	n/a	n/a	n/a
Lueboᵒ	n/a	M	6	n/a	VDS	n/a	n/a	n/a	n/a	n/a
Nakalaᵒ	n/a	F	5	n/a	VDS	n/a	n/a	n/a	n/a	n/a

Loto⁰	n/a	M	4	n/a	VDS	n/a	n/a	n/a	n/a	n/a
Khalessi⁰	n/a	F	1	n/a	VDS	n/a	n/a	n/a	n/a	n/a
Moko⁰	n/a	M	1	n/a	VDS	n/a	n/a	n/a	n/a	n/a

Supplemental Table 2. Acoustic description of each call type.

Acoustic parameters		Call Type				
		High-hoot	Bark	Soft Bark	Peep-Yelp	Peep
		N=333	*N*=431	*N*=413	*N*=420	*N*=253
F_0-Start	MEAN	1031.73	1352.51	1543.31	1364.85	1342.38
	S.D.	*490.48*	*586.69*	*340.01*	*357.09*	*376.35*
F_0-End	MEAN	1285.32	1420.46	1405.03	1363.51	1354.77
	S.D.	*404.58*	*423.75*	*393.69*	*375.72*	*384.72*
F_0-Maximum	MEAN	2392.79	2219.37	1928.1	1607.05	1401.18
	S.D.	*332.98*	*365.34*	*286.36*	*331.53*	*382.08*
F_0-Max proportion	MEAN	0.468	0.443	0.45	0.512	0.308
	S.D.	*0.131*	*0.123*	*0.149*	*0.202*	*0.36*
Call Duration	MEAN	0.239	0.193	0.172	0.158	0.106
	S.D.	*0.067*	*0.057*	*0.048*	*0.044*	*0.038*
Ascending Slope	MEAN	3236.07	1978.76	858.14	479.93	101.31
	S.D.	*1968.15*	*1200.45*	*564.47*	*391.53*	*147.78*
Descending Slope	MEAN	-2083.48	-1464.78	-953.52	-510.31	-78.25
	S.D.	*847.28*	*782.78*	*630.38*	*494.19*	*114*
Slope-Start to Midpoint of F_0	MEAN	11124.01	8787.26	4146.4	2821.28	654.03
	S.D.	*5093.9*	*5819.3*	*3021.21*	*2310.53*	*1583.44*
Slope-Midpoint to End of F_0	MEAN	-8891.67	-8096	-5849.33	-2757.86	-459.76
	S.D.	*4194.48*	*4878.49*	*4131.16*	*2635.8*	*1306.48*
Q1-Frequency	MEAN	1953.65	1832.98	1796.7	1509.9	1346.47
	S.D.	*679*	*501.88*	*414.84*	*414.53*	*404.16*
Q1-Time	MEAN	0.084	0.072	0.065	0.057	0.037
	S.D.	*0.048*	*0.048*	*0.042*	*0.038*	*0.031*
Q3-Frequency	MEAN	2952.44	2639.05	2354.58	2009.77	1723.85
	S.D.	*953.46*	*871.97*	*777.02*	*757.62*	*696.86*
Q3-Time	MEAN	0.171	0.136	0.118	0.106	0.072
	S.D.	*0.063*	*0.058*	*0.049*	*0.044*	*0.042*
Centre Frequency	MEAN	2436.3	2155.83	2028.72	1708.41	1488.27
	S.D.	*759.6*	*592.43*	*573.09*	*571.71*	*516.4*
Maximum Frequency	MEAN	2324.69	2075.19	2026.32	1743.69	1477.03
	S.D.	*952.46*	*742.84*	*658.31*	*674.38*	*547.96*
Maximum Time	MEAN	0.124	0.102	0.091	0.08	0.054
	S.D.	*0.075*	*0.059*	*0.046*	*0.042*	*0.029*

Supplemental Table 3. Results from post-hoc comparisons between each call type on their different contextual usage. * There were no peep-yelps or peeps given in the aggression context, therefore the df for that comparison is 6 as compared to 7 for the other comparisons.

	df	χ^2	Likelihood Ratio	N	p
Highhoot – Bark	7	76.296	94.65	764	< 0.001
Highhoot – Soft Bark	7	290.838	346.057	746	< 0.001
Highhoot – Peepyelp	7	419.85552	526.550615	753	< 0.001
Highhoot – Peep	7	346.154476	437.113823	586	< 0.001
Bark – Soft Bark	7	153.934834	166.575643	844	< 0.001
Bark – Peepyelp	7	265.360653	309.841552	851	< 0.001
Bark – Peep	7	216.074299	252.697659	684	< 0.001
Soft Bark – Peepyelp	7	73.984461	82.075174	833	< 0.001
Soft Bark – Peep	7	119.831321	126.653002	666	< 0.001
Peepyelp – Peep	6*	21.447089	21.840479	673	0.002

Vocal convergence in bonobos - genetic and social influences

Sumir Keenan[1,2], Jeroen MG Stevens[4], Nicolas Mathevon[1], Nicky Staes[4], Klaus Zuberbühler[2,5] and Florence Levréro[1]

[1]Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Saint-Etienne, France

[2]University of St. Andrews, Department of Psychology & Neuroscience, St. Andrews, Scotland, UK

[3]Université de Lyon/Saint-Etienne, Laboratoire de Biotechnologies Végétales appliquées aux Plantes Aromatiques et Médicinales, FRE-CNRS 3727, Saint-Etienne, France

[4]Royal Zoological Society of Antwerp, Centre for Research and Conservation, Antwerp, Belgium

[5]Université de Neuchâtel, Department of Comparative Cognition, Neuchâtel, Switzerland

Original manuscript, in preparation for submission

MANUCSRIPT 2: Vocal convergence in bonobos – genetic and social influences.

Question

Does relatedness and/or familiarity affect vocal similarity between individuals?

Methods

Subjects – 22 individual bonobos living at 3 zoos ➤ **231 unique dyads**

Related, n = 85 | Familiar, n = 156
Unrelated, n = 147 | Unfamiliar, n = 75

Acoustic analysis ➤ **12** acoustic measurements
431 individual bark calls collected
Mean of **20** calls per individual

Measures of vocal similarity ➤ 1. **Absolute acoustic distance** – Mean of each 12 acoustic variables for each 22 individuals was taken.

Example:
Mean of measurement | **Absolute acoustic distance**
'Start Frequency of F0' | Individual A ⬌ Individual B
| 1595.37 Hz 229.66 1365.71 Hz
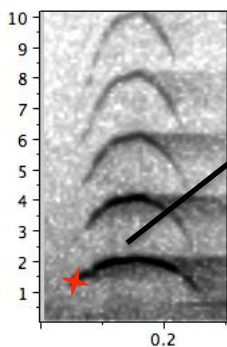
2. **Euclidean acoustic distance (ED) –**
To create a *composite measure* of acoustic features - Two factors retained from a PCA with 12 acoustic measurements, the centroid between these two variables calculated for each individual. Centroids used to calculate ED for dyads.

	Factor loading score	
Acoustic variables	**PC1**	**PC2**
F_0-Start (Hz)	0.696	- 0.374
F_0-End (Hz)	0.490	- 0.450
F_0-Maximum (Hz)	0.603	- 0.301
Call Duration (s)	0.455	0.733
Slope–Start to Midpoint of F_0 (Hz/s)	- 0.522	- 0.043
Slope–Midpoint to End of F_0 (Hz/s)	0.196	0.150
Q1 Frequency (Hz)	0.764	- 0.469
Q1 Time (s)	0.528	0.566
Q3 Frequency (Hz)	0.460	- 0.265
Q3 Time (s)	0.483	0.724
Maximum Frequency (Hz)	0.622	- 0.364
Maximum Time (s)	0.545	0.615

The measures of vocal similarity were used as dependent variables in linear mixed models.

Results

Vocal convergence occurs between both related and familiar individuals independently.

1. Absolute acoustic distance –

For each acoustic variable a ✓ indicates significant vocal convergence.

	Related	Same social group
F_0-Start (Hz)	✓	✓
F_0-End (Hz)		✓
F_0-Maximum(Hz)	✓	✓
Call Duration (s)		
Slope–Start to Midpoint of F_0	✓	
Slope–Midpoint to End of F_0	✓	
Q1-Frequency (Hz)	✓	✓
Q1-Time (s)	✓	
Q3-Frequency (Hz)		
Q3-Time (s)		
Maximum Frequency (Hz)		✓
Maximum Time (Hz)	✓	

All significance values are p < 0.05

2. For Euclidean acoustic distance –

Euclidean acoustic distance – Mean, 95% C.I.

Relatedness
Unrelated   Distantly related   Closely related
$r \leq 0.125$ $r > 0.125$

Social group
Different social group   Same social group

Increasing vocal similarity
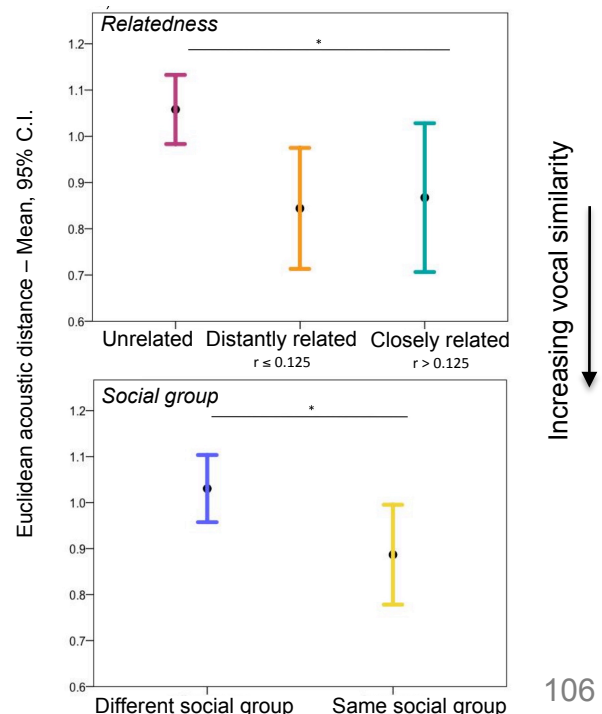
Vocal convergence in bonobos - genetic and social influences

Sumir Keenan, Jeroen M.G. Stevens, Nicolas Mathevon, Nicky Staes, Klaus Zuberbühler and Florence Levréro

ABSTRACT

The vocal communication systems of non-human primates have historically been described as composed of innate vocal structures that are relatively fixed and not subject to acoustic modification through social learning. Given the importance of vocal learning in human language, and the abundant evidence of flexible vocal behaviour in birds, cetaceans, bats and pinnipeds, this has been considered to be an evolutionary paradox. However, in very recent years, evidence of vocal dialects and vocal convergence in some primate species has suggested that their vocal structures may not be as ridged as originally thought. While this research has provided promising leads that vocal learning does occur in non-human primates, only a few studies have been able to conclusively exclude the effects of genetics or environment on the observed vocal patterns. Here we investigate vocal convergence in the bark vocalisation of bonobos, *Pan paniscus*, living at three European zoos. Vocal similarity between individuals was measured by calculating the differences between twelve acoustic variables describing the bark call type. To provide a holistic and comprehensive view, we additionally used a principle component analysis to reduce the acoustic variables into two factors, which were then used to calculate a Euclidean acoustic distance for each pair. These measures were used in a linear mixed model to compare vocal similarity between individuals based on their familiarity, while equally considering and controlling for the effect of relatedness. We found significant vocal convergence between related individuals, with vocal similarity increasing with increasing relatedness. Importantly, we also found conclusive evidence of increased vocal similarity between individuals living in the same group as compared to in different groups independent of relatedness. These results further promote the idea of

vocal learning non-human primates, and are the first to demonstrate vocal convergence based on both relatedness and familiarity in bonobos.

INTRODUCTION

Vocal production learning, the modification of a signal as a direct result of social experience, is an essential pillar in the communication systems of a wide and diverse range of taxa (Janik & Slater, 1997; 2000). Birdsong has long been used as the ideal model of vocal learning, with an overwhelming number of species showing extreme plasticity in their vocal capacity, allowing for direct and observable vocal imitation and social learning (Marler & Slabbeboorn, 2004; Catchpole & Slater, 2008; Beecher & Brenowitz, 2005; Bolhuis & Everaert, 2013; Bradbury & Balsby, 2016). Additionally, many bird species have demonstrated vocal dialects - local and regional differences between different groups of a single species that cannot be explained by genetic or geographic differences (Wright & Wilkinson, 2001; Podos & Warren, 2007; Henry, et al. 2015). Evidence of vocal learning has also been found outside birds in a range of distantly related species, but perhaps most convincingly in cetaceans (Janik, 2014), pinnipeds (Schusterman, 2008; Reichmuth & Casey, 2014), and bats (Boughman, 1998; Knörnschild, 2014).

In humans, the vocal communication system is extremely complex and flexible, and the capacity for vocal production learning persists throughout adulthood (Chambers, 2008; Fitch, 2010). As such, in the search for the evolutionary roots of human language comparative studies have sought to demonstrate this capacity in our closest living relatives (Fitch, 2005). Early vocal learning experiments attempted to teach non-human primates, namely chimpanzees, to speak and despite years of training were marked by failure (Hayes, 1951; Premack & Premack, 1972). These studies, along with others on monkeys (Egnor & Hauser, 2004), demonstrated a lack of vocal flexibility and lead to a general consensus that non-human primate vocal production is largely innate, with strong genetic influences (Hammerschmidt & Fischer, 2008) and a fixed acoustic structure (Jürgens, 1995; Tomasello, 2008).

In more recent years, research on the vocal behaviour of non-human primates has uncovered the presence of vocal dialects in a range of species, suggesting that the significant vocal differences found between populations within a species is the result of vocal learning. Unfortunately, many of these studies were not able to exclude the possibility that the dialects were due to genetic or habitat differences between groups (Hammerschmidt & Fischer, 2008; Snowdon, 2009), although with notable exceptions in chimpanzees (Crockford et al., 2004) and Japanese macaques (Tanaka et al., 2006). The evidence that genetic and ecological differences can affect vocal behaviour in non-human primates, has led some to argue that without controlling for these factors the use of vocal dialects as evidence for vocal production learning is largely inferential (Snowdon, 2009). These inconsistencies have led many researchers to conclude that in non-human primates, the capacity for communicative social learning may instead be mostly present in their gestural communication (Call & Tomasello, 2007; Arbib et al., 2008) or in their ability for contextual vocal learning – in which an established signal is learned to be associated to a new context as the result of social experience (Janik & Slater, 2000).

An alternate approach in the investigation of vocal production learning in non-human primates has been to study vocal flexibility by way of vocal convergence between individuals or whole groups. Chimpanzees (Mitani & Gros-Louis, 1998), Diana monkeys (Candiotti, et al., 2012) and Japanese macaques (Suguira, 1998), can flexibly modify their vocalizations to converge (vocal matching) in certain contexts, such as with individuals with whom they are chorusing. While Campbell's monkeys and mandrills show stable vocal convergence with closely affiliated individuals (Lemasson et al., 2011) or group members (Levrero et al., 2015), and chimpanzees (Watson et al., 2015) and pygmy marmosets (Elowson & Snowdon, 1994) have demonstrated stable group-wide, vocal convergence over time and across multiple individuals after the merging of captive populations. These evidence for vocal flexibility in non-human primates, although

more subtle than what has been found in other taxa, strongly suggest the potential for vocal production learning.

In this study, we had the opportunity to investigate the vocal dynamics between individual bonobos with the advantage of knowing the full life history of each individual, including their entire pedigree and the length of familiarity between all dyads. We thus investigated the presence of vocal convergence between pairs of individuals from the same population as compared to different populations, while equally considering and controlling for the effect genetic relatedness may have on vocal similarity in three captive groups of bonobos (*Pan paniscus*).

In the wild, bonobos are native to the dense forest regions of the Democratic Republic of Congo, and live in a male-philopatric society characterised by complex fission-fusion dynamics and close social bonds between related and unrelated individuals (Furuchi, 2011). As a forest dwelling species, that are regularly in low-visibility environments, bonobos are hugely reliant on vocalisations for inter- and intra- group communication (Hohmann & Fruth 1994; White 1996; Hohmann & Fruth, 2002; Furuchi, 2011). Here we investigate vocal convergence in the bark call type, a loud call type that is used in both distance and close communicative exchanges. In the wild, barks are most commonly given to the discovery of a valued food source, during feeding, and during agonistic encounters, but also are used during inter-community encounters (Bermejo & Omedes, 1999).

The source-filter hypothesis argues that the acoustic features of an individual's calls are created and dictated by the morphological features of both the source of sound production (the vocal cords) and by the filter through which the sounds pass (the vocal tract) (Fitch & Hauser, 1998; Reby & McComb, 2003; Taylor & Reby, 2010). As genetically related individuals are more likely to have inherited similar vocal morphology than unrelated individuals, we hypothesise that relatedness will have an effect on vocal similarity. Additionally, based on

evidence from previous studies on non-human primates demonstrating vocal convergence between individuals and/or whole group dialects, we predict that we will likely see a level of convergence between individual bonobos living in the same group. To date, this study is the first to concurrently investigate the effects of relatedness and social familiarity on the vocal similarity between individuals in any ape species.

METHODS

Study groups

The bonobos observed for this study are members of three separate captive groups located at three European zoos - Apenheul (Apeldoorn, the Netherlands), Planckendael Zoo (Mechelen, Belgium) and la Vallée des Singes (Romagne, France).

A recent study showed that male bonobos show a sharp increase in testosterone around eight years old, indicating the onset of sexual maturity (Behringer et. al 2014), while bonobo females are thought to mature between the ages of 6-10, based mainly on natal group dispersion (Furuchi, 1989) but also with corroborating hormonal evidence (Behringer et. al 2014). Therefore, to avoid the potential confound of comparing adult voices to immature ones, only individuals over the age of 10 years were included in the study (with the exception of one female who turned 10 during recording). This resulted in vocal recordings being collected from 22 individuals, Apenheul (n=6), Planckendael (n=6) and la Vallée des Singes (n=10), 14 females and 8 males, ranging in age from 10 to 45, with a mean age of 20.

In total, the 22 subjects comprised 231 unique dyads. For each dyad, a variety of information was included in the study. The life history of captive bonobos can be complex, and individuals are moved between zoos for population management and to mimic bonobos natural female dispersal from her native group. As such, in

addition to investigating vocal similarity between dyads currently living together, we also included information on the time each pair had spent together. This allowed us to account for the time individuals had spent together, as well as assessing the vocal similarity of individuals who were familiar to one another but no longer housed together. Familiar dyads included individuals that were currently in the same group or had previously lived together. In both cases, the amount of time spent living together was rounded in 6 month increments and ranged between 6 – 252 months (with a mean of 71 months for familiar individuals, and only a single familiar dyad having lived together for less than 12 months) and included 25 past familiar dyads. All dyads that were in the same group at the time of the study (calculated as of January 1, 2014) had been living together for a minimum of 12 months.

Relatedness for all dyads was evaluated using the relatedness coefficient (r), which was calculated and provided by the European Association for Zoos and Aquariums (EAZA) bonobo studbook manager.

Data Collection

All vocal recordings were collected between March 2013 and May 2014, between 8 a.m. and 6 p.m., with a minimum of 115 total recording hours per zoo (Apenheul ~ 175 recording hours; Planckendael ~ 190 recording hours; Vallée des Singes ~ 115 recording hours). Audio recordings were taken using a Zoom H4 Digital Multi-Track Recorder - recording in stereo, with one channel devoted to a Sennheiser MKH70-1 ultra-directional microphone recording any bonobo vocal behaviour and the second channel connected to a micro-tie recording device, model AKA MPA III, for comments by the researcher.

Call Type

For the current study, we investigated the effects of genetic relatedness, group membership and familiarity on the bark call type. Calls were classified as barks based on classifications described in previous studies (de Waal, 1988; Bermejo

& Omedes, 1999; Clay & Zuberbuhler, 2009) and included: calls described as barks and composed barks by Bermejo & Omedes (1999); barks, specifically wieew barks, by de Waal (1988); but excluded what Clay & Zuberbuhler (2009) described as food barks. Additionally, our call type classifications were verified using a pDFA analysis (Keenan et al., *Submitted*; See Manuscript 1 of this thesis - 'Individual vocal signatures across the bonobo vocal repertoire' - for details).

The bark call type was a common call type used by all of the study bonobos and comprised almost 20% of the total number of calls we collected. In captivity, barks are used in a variety of contexts but mostly given during times of group excitement (characterised by increased sexual activity, displacements, pacing, pilo-erection, vocal activity and/or displays) and/or when potential for conflict was high or already occurring and specifically included: directly before (34%) or during scheduled feedings (18%); being shifted from indoor to outdoor habitats (14%); and by aggressors in agonistic encounters (19%). The remaining 15% of barks were found in the following contexts: as alarm calls (7%); foraging on grass, trees or bushes (3%); as contact calls (4%); and during grooming with other individuals (1%) (Keenan et al., *Submitted*; See Manuscript 1 of this thesis for details). In the wild, barks are used in similar contexts, during times of excitement, such as when discovering a valued food source, and aggression (Bermejo & Omedes, 1999). Additionally, barks are also used in the wild during distance communication between community sub-parties who separate to search for food and during inter-community encounters (Bermejo & Omedes, 1999).

Analyses

Acoustic Measurements & Analysis

Only vocalisations that could unambiguously be assigned to one caller, and that were of good measurable quality were retained for analysis. In total, 431 individual calls were retained for analyses with a mean of 20 calls per individual (range of 5–55). Raven Pro 1.3 was used to measure automatic and manual parameters on each call. A correlation matrix was produced and highly correlated

variables were removed (0.8 and above), resulting in a total of 12 measurements being retained for analysis – six manual acoustic parameters describing the fundamental frequency and its temporal modulation (measurements were made using Raven's default spectrogram view, with the exception of window size, which was set to 512): F_0-Start (fundamental frequency at beginning of the call, in Hz), F_0-End (fundamental frequency at end of the call, in Hz), F_0-Maximum (highest frequency reached on the fundamental, in Hz), Call Duration (s), Slope–Start to Midpoint of F_0 (calculated as: F_0 at midpoint of call duration – F_0-Start/Time at midpoint of call duration – 0) and Slope–Midpoint to End of F_0 (calculated as F_0-End – F_0 at midpoint of call duration/Call duration – time at midpoint of call duration), and six automatically computed parameters describing the distribution of energy among the frequency spectrum: Q1-Frequency (the frequency at which the call is divided into two frequency intervals containing 25% and 75% of the call's energy, in Hz), Q1-Time (the time point along the call at which the call is divided into two time intervals containing 25% and 75% of the call's energy, in seconds), Q3-Frequency (the frequency at which the call is divided into two frequency intervals containing 75% and 25% of the call's energy, in Hz), Q3-Time (the time point along the call at which the call is divided into two time intervals containing 75% and 25% of the call's energy, in seconds), Maximum Frequency (the frequency at which the maximum energy occurs in the call, in Hz) and Maximum Time (first time point along the call where maximum amplitude, from waveform, occurs, in seconds).

Statistical analysis

To obtain a measure of vocal similarity between dyads, we calculated the mean of each variable for each individual, followed by the absolute difference between these means for each variable for all dyads. For an additional measure of vocal similarity, we standardised each of the 12 variables (into z-scores, as they were different units) and a PCA was calculated. The first two factors were retained for analysis and represented 52% of the total variance (Supplemental Table 1) (*Supplemental Tables are following the Reference section of this manuscript*.)

For each individual, we calculated the centroid of both principal components, which was then used to calculate the Euclidean acoustic distance (E.D.) between each dyad as a comprehensive measure of acoustic similarity.

These measures of vocal similarity were then used as dependent variables in four separate linear mixed models. The first model was used to provide a comprehensive view of the effect of both relatedness and familiarity on each of the acoustic variables (as measured by the absolute difference) as well as the composite variable, the Euclidean acoustic distance. This first model also included a number of fixed and random factors, and tested for any interaction between the two main independent variables. The following three secondary models were used to test the more subtle nuances in the data set and only used the Euclidean acoustic distance as the dependent measure.

SPSS (Version 22) was used to perform all statistical tests. To achieve normality (based on the Kolmogorov-Smirnov statistic and visually investigating P-P plots) and homogeneity of variance (based on Levene's statistic), the dependent variables (the E.D. and the absolute differences for each variable) were square-root transformed. In the first, and primary, linear mixed model relatedness (categorised as unrelated: r=0, n=147 dyads; distantly related: r $\leq$ 0.125, n=49; or closely related: r > 0.125, n=35) and group (currently living at the same zoo, n=156 dyads; or currently living at different zoos, n=75) were the main variables. We further considered the following as possible confounding covariates and included them as fixed factors: familiarity measured as time spent together in months (continuous variable calculated in 6 month increments, and included individuals who had previously lived together but who are not currently in the same group), sex composition of the dyad (coded as 'different sex' or 'same sex'), and the absolute difference in age between each individual in each dyad. Interaction between the two main variables, relatedness and familiarity, was also included in the model. Lastly, we considered the identity of the two individuals in

the dyad as two random effects (LMM for all models set at Maximum likelihood; maximum iterations=100; maximum step-halvings=10).

Individuals who were familiar with one another but no longer living in the same group presented a possible confound for our analysis. Ideally, we would have removed them completely; however, almost all 25 past familiar pairs fell into the highly related category and removing them would have greatly unbalanced the number of data points between the unrelated and related groups. For the primary analysis, these past familiar pairs were included, and we controlled for their possible confound by including both length of familiarity in months and whether they currently live together or not in the model. For the further models the data were treated differently and we were thus able to remove them from the secondary analyses without unbalancing the data.

For the first of the secondary models, we separated the dyads into four groups: 1) dyads that live together and are related (Familiar-related, n=27); 2) dyads that live together and are not related (Familiar-unrelated, n=48); 3) dyads that are related but completely unfamiliar (Unfamiliar-related, n=43); 4) dyads who are unrelated and unfamiliar (Unfamiliar-unrelated, n=88) and investigated the vocal similarity between these four groups with a LMM using the E.D. (square-root transformed) as the dependent variable. All individuals who were familiar but no longer live in the same group (past familiar dyads, n=25) were excluded in this analysis. As before, we considered the following as possible confounding covariates and included them as fixed factors: time spent together in months (continuous variable calculated in 6 month increments), sex composition of the dyad (coded as 'different sex' or 'same sex'), and the absolute difference in age between each individual in each dyad. The identity of the two individuals in the dyad was considered as two random effects.

Lastly, we investigated the effect investigated familiarity and relatedness using continuous variables, specifically the relatedness coefficient and time spent

together. This was done was done with two linear mixed models limited to the unrelated and unfamiliar dyads separately. For the unrelated dyads (n=136) we investigated the relationship between the familiarity of dyads (time spent together in months, continuous variable and limited to pairs currently living in the same group) and the Euclidean acoustic distance (E.D., square-root transformed). Past familiar individuals were excluded from this analysis. For the unfamiliar dyads (n=131), we investigated the relationship between the relationship coefficient (as a continuous variable) and the Euclidean acoustic distance. For these two models sex composition of the dyad (coded as different sex or same sex) and the absolute difference in age between each individual in each dyad were considered as fixed factors and the identity of the two individuals in the dyad were considered as two random effects.

RESULTS

The analyses showed a clear overall effect of both relatedness and group on the vocal similarity between individuals, even when controlling for other effects, such as age or sex differences. The results of our first linear model demonstrated that the acoustic distance (E.D.) is shorter between related individuals as compared to unrelated ones (LMM, n=231, F= 8.710, p=0.003; Figure 1b, Table 1 & Supplemental Table 2), and between individuals at the same zoo compared to dyads who do not live together (LMM, n=231, F=5.455, p=0.020) (Figure 1a; Table 1 & Supplemental Table 2).
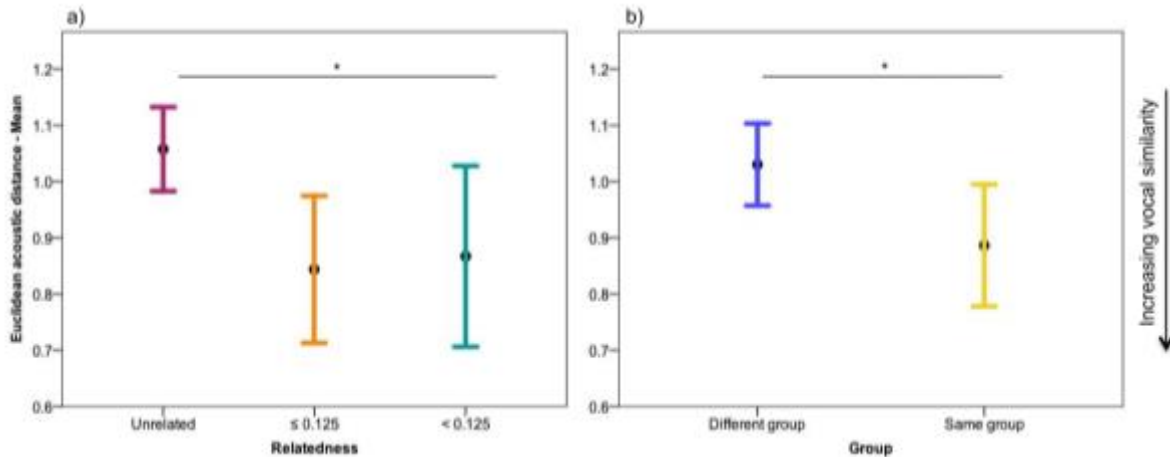
Figure 1. Primary model. a) Euclidean acoustic distances between dyads in the three related categories. b) Euclidean acoustic distances between dyads from the same group compared to different groups. Both graphs show means (shaded circle) with error bars displaying a 95% confidence interval. *$p < 0.05$.

Additionally, when investigating the effect of relatedness in only unfamiliar individuals, we found that the acoustic distance (E.D.) between individuals decreased as the relatedness coefficient increased (LMM, n=131, F=5.533, p=0.020) (Figure 2a). Conversely, when looking at the effect of familiarity on the acoustic distance (E.D.) between pairs within unrelated group members, we found that the time individuals had spent together (with a minimum of 12 consecutive months) did not significantly affect the vocal similarity between individuals (LMM, n=136, F=1.146, p=0.286) (Figure 2b). Considering all these results, the comparison between the four different types of dyads showed an overall significant effect (LMM, n=206, F=4.550, p=0.034) with familiar-related pairs having the smallest acoustic distance (Euclidean acoustic distance mean=0.784, S.D.=0.482), followed by unfamiliar-related (E.D. mean=0.943, S.D.=0.449), familiar-unrelated (E.D. mean=0.922, S.D.=0.427), and lastly unfamiliar-unrelated (E.D. mean=1.079, S.D.=0.433) (Figure 3).

	(F, P)		Current group x Relatedness	Sex	Age difference	Familiarity *(time together)*
Response variables	**Relatedness**	**Group**	**Current group x Relatedness**	**Sex**	**Age difference**	**Familiarity** *(time together)*
Euclidian Distance	*8.710, 0.003*	**5.455, 0.020**	0.978, 0.324	1.015, 0.315	0.679, 0.411	0.063, 0.802
Absolute Difference F0-Start	**5.080, 0.025**	*15.627, 0.000*	0.130, 0.719	0.993, 0.320	0.448, 0.504	**4.751, 0.030**
Absolute Difference F0-End	0.424, 0.516	**6.931, 0.009**	0.059, 0.809	1.951, 0.164	1.637, 0.202	0.593, 0.442
Absolute Difference F0-Maximum	**4.945, 0.027**	**5.422, 0.021**	0.296, 0.587	0.973, 0.325	3.303, 0.070	*14.469, 0.000*
Absolute Difference Call Duration	0.278, 0.599	**6.628, 0.011**∗	0.000, 0.997	3.169, 0.076	0.064, 0.801	*12.130, 0.001*
Absolute Difference Slope - Start to Midpoint of F0	**7.248, 0.008**	0.240, 0.625	0.083, 0.774	0.370, 0.544	2.842, 0.093	0.048, 0.827
Absolute Difference Slope - Midpoint to End of F0	*9.615, 0.002*	0.377, 0.540	0.273, 0.602	0.655, 0.419	**4.222, 0.041**	0.147, 0.702
Absolute Difference Q1-Frequency	*11.198, 0.001*	**6.338, 0.012**	0.115, 0.735	0.188, 0.665	1.310, 0.254	1.698, 0.194
Absolute Difference Q1-Time	**6.476, 0.012**	1.867, 0.173	0.675, 0.412	0.215, 0.643	0.428, 0.513	0.049, 0.825
Absolute Difference Q3-Frequency	2.758, 0.098	0.354, 0.552	0.207, 0.649	**5.735, 0.017**	0.990, 0.321	0.047, 0.828
Absolute Difference Q3-Time	0.004, 0.951	**6.289, 0.013**∗	0.033, 0.856	1.586, 0.209	0.012, 0.912	*17.572, 0.000*
Absolute Difference Maximum Frequency	2.479, 0.117	*9.040, 0.003*	0.651, 0.421	0.183, 0.669	0.077, 0.782	1.233, 0.268
Absolute Difference Maximum Time	**4.951, 0.027**	0.000, 0.986	0.006, 0.940	0.058, 0.810	0.433, 0.511	0.498, 0.481

Table 1. Results of primary linear mixed model testing the effect of relatedness and group on vocal similarity. Numbers in bold are under significance at $p < 0.05$, without Bonferonni corrections. If they are bold and *italicized* then they remain significant under Bonferonni corrections (which in this case would be $p < 0.004$). ∗These two measurements showed significant divergence, and not convergence, between dyads of the same group.
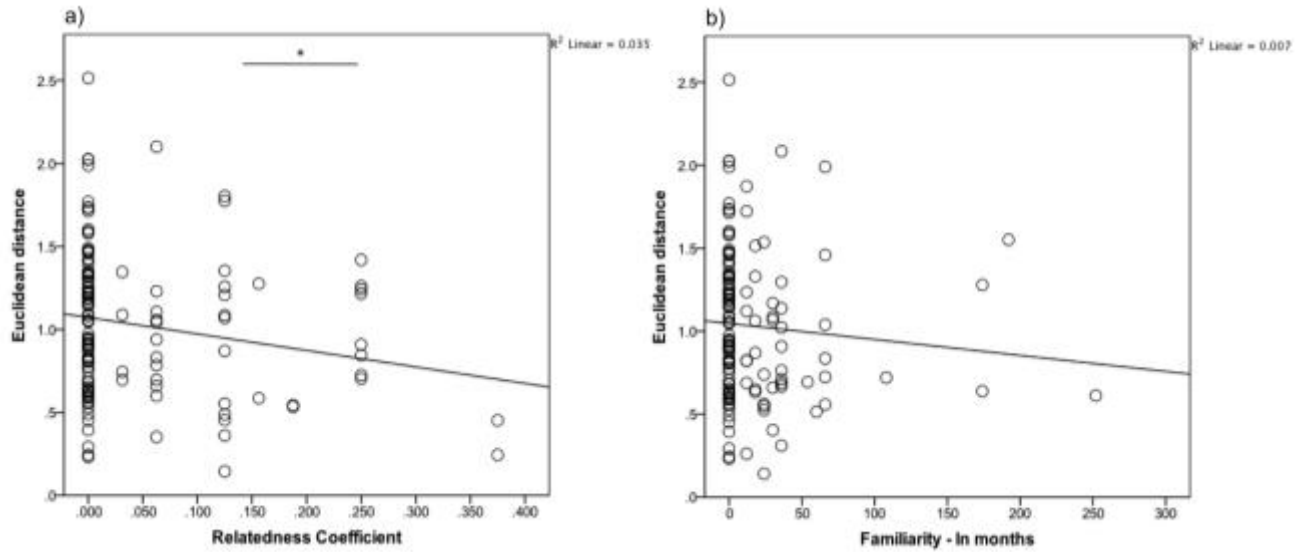
Figure 2. Results from LMM on continuous measures of relatedness and familiarity. a) shows results on relatedness and b) on familiarity (which demonstrated the same linear trend, but was not significant). *$p < 0.05$.
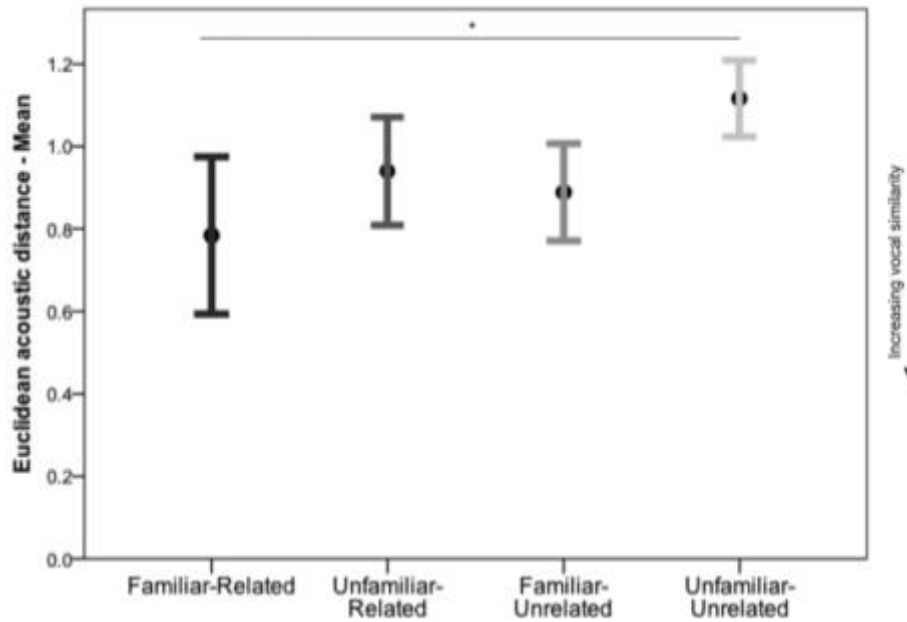

Figure 3. Secondary analysis investigating vocal similarity between four categorical groups that describe relatedness and familiarity. *$p < 0.05$.

The analysis of the absolute difference of each acoustic variable demonstrated that the respective weight of relatedness and familiarity on vocal similarity

between individuals depends on the acoustic variable (Table 1 & Supplemental Table 2; Figure 4). Four acoustic features were more similar between related dyads: Slope– F_0-Start to Midpoint of F_0, Slope–Midpoint to F_0-End (two variables illustrating frequency modulation), Q1 Time, and Maximum Time (two variables linked to the call temporal dynamics) (Figure 4a). Two other acoustic features were more similar between individuals of the same group: F_0-End and Maximum Frequency (two variables describing the fundamental frequency) (Figure 4b). Additionally, both relatedness and group significantly decreased acoustic differences between dyads in three acoustic variables: F_0-Start, F_0-Maximum (two variables describing the fundamental frequency) and Q1 Frequency (a variable illustrating the distribution of energy among the spectrum) (Figure 4c). Lastly, two temporal acoustic variables showed significant *increases* in acoustic differences between individuals of the same group: Call Duration and Q3 Time (Figure 4b).

Unsurprisingly, four variables that were significant for group were also significant for length of familiarity, which was included in the model as fixed factor (Table 1). Additionally, there was no interaction between the two main factors, relatedness and group, for all 12 parameters and the Euclidean acoustic distance. We also considered two additional fixed factors, age difference between individuals and if they were of the same sex or not. Neither of these factors were significant in the Euclidean acoustic distance, however each were significant for a single variable. Individuals of the same sex sounded more similar to one another than to members of the opposite sex in the Q3-Frequency variable and individuals closer in age were more dissimilar in the Slope–Midpoint to End variable (Table 1).
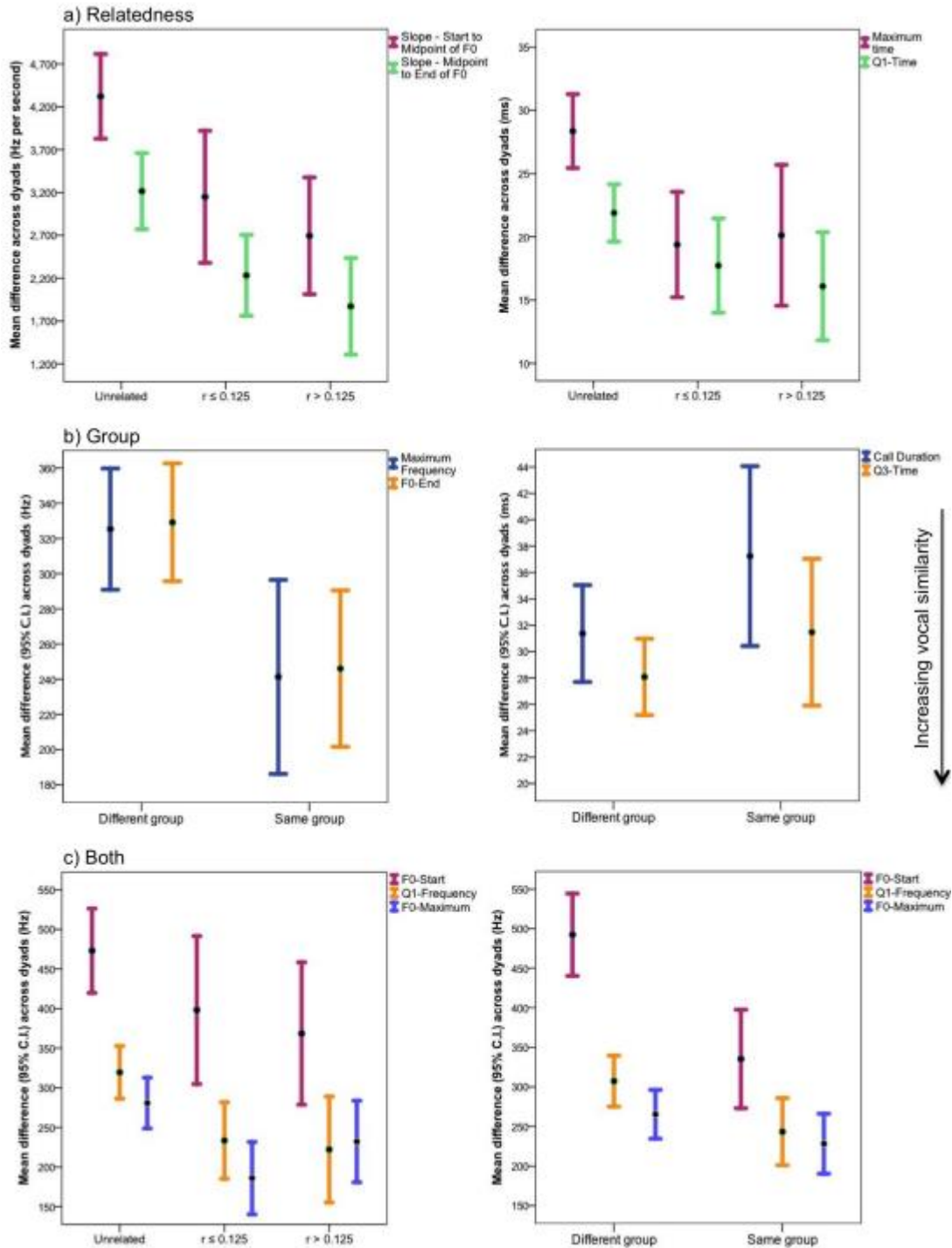
Figure 4. Results from primary model investigating effect of relatedness and group on each acoustic parameter separately. Panel a) displays the acoustic variables that were significantly affected by relatedness. Panel b) displays the acoustic variables that were significantly affected by group. And the final panel, c), shows acoustic variables that were affected by both group and relatedness. (All comparisons shown in this figure are significant, but due to multiple variables on each panel significant asterisks were not included – see Table 1 for all *p* values.)

DISCUSSION

This study demonstrates the first evidence of vocal convergence in bonobos based on both genetic similarity and social familiarity. By using acoustic distances to compare the vocal characteristics between individuals in the bark call type, we found that individuals who are related sound more similar than individuals who are not related and that vocal similarity increased with increasing relatedness. Of the twelve acoustic parameters measured, seven were significantly more similar in related pairs than unrelated pairs. The results were similar when considering the acoustic distances for individuals living in the same group, with individuals living together being more vocally similar than individuals in different groups. However, the strength of similarity did not increase with longer periods of familiarity, suggesting that a period of a year is more than enough time for individuals' acoustic features to converge. When testing the effect of familiarity on each parameter separately, we found that five of the twelve parameters were significantly more similar between dyads of the same social group compared to dyads from different groups. However, two temporal parameters ('Call duration' and 'Q3-Time') showed minimal but significant divergence between familiar pairs. The cause of the divergence in these two features is unclear, however as individual recognition is vital for social functioning (Tibbetts & Dale, 2007; Keenan, et al., 2016), preventing complete vocal convergence and maintaining individual identity information would be highly adaptive.

Additionally, we compared vocal similarity between four different types of dyads: familiar-related, familiar-unrelated, unfamiliar-related and unfamiliar-unrelated. If familiarity and relatedness were causing vocal convergence equally we hypothesised that the greatest vocal similarity would be found in familiar-related dyads, the least in unfamiliar-unrelated and there would be minimal differences between familiar-unrelated and unfamiliar-related. This is precisely what we

found, which further strengthens the evidence that both relatedness and familiarity are affecting vocal similarity.

It has been suggested that vocal similarity between related individuals provides a basis for kin recognition (Gouzoules & Gouzoules, 1990; Kessler et al., 2012, 2014). Kin selection, in which individuals act nepotistically in social interactions, plays an important role in shaping social relationships in a range of animal societies, and depends entirely on the ability to recognise kin from non-kin (Silk, 2002; Widdig, 2007). Support for this hypothesis has come from studies with macaques (Rendall et al., 1996; Pfefferle et al., 2014, 2015), grey mouse lemurs (Kessler et al., 2012) and mandrills (Levrero et al., 2015), which have experimentally demonstrated that individuals are capable of using vocalisations to discriminate between related and unrelated conspecifics. Notably, however, evidence for kin selection and vocal recognition have largely focused on species whose social structures are centred around matrilineal networks of related individuals (Gouzoules & Gouzoules, 1987; Melnick & Pearl, 1987) and who display nepotism during grooming, coalition support and reconciliation, making recognising kin important in social functioning (Silk, 2002).

For species like bonobos and chimpanzees, whose social structure is dependent upon cooperation, coordination and coalition between kin and non-kin, there is considerably less evidence of kin selection in social behaviour, excluding between mother and offspring (Goldberg & Wrangham, 1997; Hohmann, et al., 1999; Langergraber et al., 2007; Tokuyama & Furuichi, 2016). In such social environments, the importance of kin recognition is unclear. However, a study with male-philopatric western gorillas found that after migration from their natal groups, females were significantly more likely to be living in a new group with at least one female relative (Bradley et al., 2007). In female bonobos, the process of choosing a new community during natal dispersal is not well understood (Gerloff et al., 1999; Hashimoto et al., 1998), however, it has been observed that female bonobo migration is characterised by a lack of aggression from resident

females toward immigrants. Therefore, it is possible that during migration females may choose new communities inhabited with female relatives that have migrated before them (Idani, 1991; Sakamaki et al., 2015). Despite the lack of exclusive matrilineal association in these male-philopatric species, these observations in bonobos and results in gorillas suggest that kin recognition and selection may facilitate the process of female natal dispersal. Further study is required to test if kin recognition, facilitated by kin vocal signatures, plays a role in how females choose new communities after natal migration.

The convergence of multiple acoustic features between familiar individuals provides a strong argument for vocal flexibility and thus vocal production learning in bonobos independent of genetic factors. Other factors that could account for vocal convergence between individuals, such as increased similarity between individuals of the same sex or similar age were controlled for in this study and were found to have no effect on the variables in which acoustic convergence was occurring. Additionally, all individuals were living in similar captive zoo environments, and some individuals had lived at more than one of the zoos, making characteristics of the habitat an unlikely explanation for the observed vocal similarities between group members. The historical lack of evidence for vocal learning in non-human primates has been considered a paradox in the evolutionary models of human language (Egnor & Hauser, 2004). The results presented in this study contribute to the growing body of evidence suggesting that vocal production learning is indeed occurring in non-human primates.

For non-human primates vocal convergence has been hypothesised to function in creating vocal dialects, which have been suggested to promote group cohesion, facilitate recognition and advertise group membership (Crockford et al., 2004; de la Torre & Snowden, 2009; Fischer et al., 1998; Henry et al., 2015). However, additional arguments have been made that vocal convergence also has a similar function to that in humans, as described by speech accommodation theory, to reinforce bonding between established group members and to facilitate

social integration of new group members (Giles, 1991; Hammerschmidt & Fischer, 2008; Locke & Snow, 1997). This latter hypothesis is supported by evidence of active vocal convergence during chorusing or with closely affiliated group members in chimpanzees (Mitani & Brandt 1994; Mitani & Gros-Louis, 1998) Diana monkeys (Candiott et al., 2012), pygmy marmosets (Snowden & Elowson, 1999), Campbell's monkeys (Lemasson et al., 2011) and Japanese macaques (Suguira, 1998). Importantly, these two hypotheses are not mutually exclusive and it is possible that vocal convergence in bonobos serves to communicate group membership to both group members and individuals in neighbouring communities, as well as to strengthen social bonds between individuals within a group.

These results provide intriguing insights into the complexity of vocal communication and social interactions in bonobos and open up many avenues for exploration. Investigating vocal similarity within group members but differentiating between pairs with close, affiliative relationships as compared to individuals with distant or agonistic relationships could improve our understanding on how and if vocal convergence functions in social bonding. On the other hand, while we found evidence for vocal convergence, we did not experimentally test if bonobos can use this vocal similarity for kin or group recognition. Conducting such experiments could provide further clues to how vocal similarities could be utilised for successful social navigation both within a group and potentially between neighbouring communities.

Non-human primate communication has largely been studied with the objective of researching the evolution of human language. When compared to the huge human capacity for vocal flexibility, the more subtle flexibility present in the stable repertoires of non-human primates has perhaps been overlooked. Additionally, the cognitive complexity of ape species' social lives, and their tendency for multi-modal communication, has possibly muddled the evidence for vocal production learning present in their vocalizations. Whether the function of the observed

vocal convergence is for recognition or social bonding, we argue that the convergence is indeed evidence of vocal production learning, as we have controlled for other likely explanations.

REFERENCES

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate Vocalization, Gesture, and the Evolution of Human Language. *Current Anthropology*, *49*, 1053–1076.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology and Evolution*, *20*, 143–149.

Behringer, V., Deschner, T., Deimel, C., Stevens, J. M. G., & Hohmann, G. (2014). Age-related changes in urinary testosterone levels suggest differences in puberty onset and divergent life history strategies in bonobos and chimpanzees. *Hormones and Behavior*, *66*, 525–533.

Bermejo, M., & Omedes, A. (1999). Preliminary vocal repertoire and vocal communication of wild bonobos (Pan paniscus) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, *70*, 328–357.

Bolhuis, J.J. & Everaert, M. (Eds.) (2013). *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain*. Cambridge, Mass.: MIT Press.

Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society B: Biological Sciences*, *265*, 227–233.

Bradley, B. J., Doran-Sheehy, D. M. & Vigilant, L. (2007). Potential for female kin associations in wild western gorillas despite female dispersal. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 2179–2185.

Bradbury, J. W., & Balsby, T. J. S. (2016). The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology*, *70*, 293–312.

Call, J. & and Tomasello, M. (Eds.) (2007). *The gestural communication of apes and monkeys*. New Jersey & London: Lawrence Erlbaum.

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, *8*, 382–385.

Catchpole, C. K. & Slater, P.J.B. (2008). Bird song: Biological themes and variations (2nd ed.). Cambridge, U.K.: Cambridge University Press.

Chambers, J. K. (2008). Dialect Acquisition. *Linguistic Society of America*, *68*, 673–705.

Clay, Z., & Zuberbuhler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, *77*, 1387–1396.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild Chimpanzees Produce Group-Specific Calls : a Case for Vocal Learning ? *Ethology*, *243*, 221–243.

de Waal, F. B. M. (1988). The communicative repertoire of captive bonobos *(Pan paniscus*), compared to that of chimpanzees. *Behaviour*, 27, 183–251.

Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*, 649–54.

Elowson, A. M., & Snowdon, C. T. (1994). Pygmy marmosets, Cebuella pygmaea, modify vocal structure in response to changed social environment. *Animal Behaviour*, 47, 1267–1277.

Fitch, W. T., & Hauser, M. D. (1998). Unpacking " Honesty ": Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, A.N. Popper

& R.R. Fay (Eds.), *Acoustic Communication* (pp. 15–64). New York, New York: Springer–Verlag.

Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, *20*,193–230.

Fitch, W. T. (2010). *The evolution of language.* Cambridge, U.K.: Cambridge University Press.

Furuichi, T. (1989). Social Interactions and the Life History of Female *Pan paniscus* in Wamba, Zaire. *International Journal of Primatology*, *10*, 173–197.

Furuchi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthrophology,* 20, 131-142.

Gerloff, U., Hartung, B., Fruth, B., Hohmann, G., & Royal, T. H. E. (2013). Intracommunity relationships, dispersal pattern paternity living analysis, *266*, 1189–1195.

Giles, H., Coupland, N., & Coupland, J. (1991). Accomodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accomodation: Developments in Applied Sociolinguistics* (pp. 1-68). Cambridge, U.K.; Cambridge University Press.

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In Oller, D.K. & U. Griebel (Eds.) *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication* (pp. 367-395). Cambridge, Mass.: MIT Press.

Hashimoto, C., Takenaka, O., & Furuichi, T. (1996). Matrilineal kin relationship and social behavior of wild bonobos (*Pan paniscus*): Sequencing the D-loop region of mitochondrial DNA. *Primates*, *37*, 305–318.

Hayes, C. (1951). The ape in our house. NY: Harper and Brothers.

Henry, L., Barbu, S., Lemasson, A., & Hausberger, M. (2015). Dialects in Animals : Evidence , Development and Potential Functions. *Animal Behavior and Cognition*, *2*, 132–155.

Hohmann, G., & Fruth, B. (1994). Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology*, *15*, 767–782.

Hohmann, G. & Fruth, B. (2002). Dynamics is social organization of bonobos (*Pan paniscus*). In: Boesch, C., Hohmann, G. & Marchant, L.F. (Eds.) *Behavioural Diversity in Chimpanzees and Bonobos* (pp.138-150). Cambridge, U.K.: Cambridge University Press.

Idani, G. (1991). Social relationships between immigrant and resident bonobo (*Pan paniscus*) females at Wamba. *Folia Primatologica*, *57*, 83–95.

Janik, V. M., & Slater, P. J. B. (1997). Vocal Learning in Mammals. *Advances in the Study of Behavior*, 26, 55–99.

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*, 1–11.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60–65.

Jürgens, U, (1995). Neuronal control of vocal production in nonhuman and human primates. Current topics in primate vocal communication. Edited by: Zimmerman E, Newman JD, Jürgens U. 1995, New York: Plenum, 199-202.

Kessler, S. E., Radespiel, U., Hasiniaina, A. I. F., Leliveld, L. M. C., Nash, L. T., & Zimmermann, E. (2014). Modeling the origins of mammalian sociality: moderate evidence for matrilineal signatures in mouse lemur vocalizations. *Frontiers in Zoology*, *11*, 14.

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat, *Saccopteryx bilineata*. *Animal Behaviour*, *84*, 761–769.

Lemasson, A., Ouattara, K., Petit, E. J., & Zuberbühler, K. (2011). Social learning of vocal structure in a nonhuman primate?. *BMC Evolutionary Biology*, *11*, 362.

Levréro, F., Carrete-Vega, G., Herbert, A., Lawabi, I., Courtiol, A., Willaume, E., Kappeler, P.M. & Charpentier, M. J. E. (2015). Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communication*, 6, 7609.

Marler, P.R. & Slabbekoorn, H. (2004) *Nature's Music: The Science of Birdsong.* Elsevier Academic Press.

Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and Call Convergence in Chimpanzees: Tests of Three Hypotheses. *Behaviour*, *135*, 1041–1064.

Podos, J., & Warren, P. S. (2013). The Evolution of Geographic Varation in Birdsong. *Advances in the Study of Behavior*, *37*, 403–444.

Pfefferle, D., Ruiz-lambides, A. V, & Widdig, A. (2014). Female rhesus macaques discriminate unfamiliar paternal sisters in playback experiments: support for

acoustic phenotype matching. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20131628.

Premack, A. J., & Premack, D. (1972). Teaching language to an ape. Scientific American, 227, 92-99.

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, *65*, 519–530.

Tanaka, T., Sugiura, H., & Masataka, N. (2006). Cross-Sectional and Longitudinal Studies of the Development of Group Differences in Acoustic Features of Coo Calls in Two Groups of Japanese Macaques. *Ethology*, *112*, 7–21.

Tibbetts, E. A., & Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, *22*, 529–37.

Tomasello, M. (2008) *Origins of Human Communication*. Cambridge, Mass.; MIT Press.

Sakamaki, T., Behncke, I., Laporte, M., Mulavwa, M., Ryu, H., Takemoto, H., Tokuyama, N., Yamamoto, S. & Furuichi, T. (2015). Intergroup transfer of females and social relationships between immigrants and residents in bonobo (*Pan paniscus*) societies. In T. Furuichi, J. Yamagiwa, & F. Aureli (Eds). *Dispersing primate females* (pp. 127-164). Tokyo: Springer.

Silk, J. B. (2002). Kin Seletion in Primate Groups. *International Journal of Primatology*, *23*, 849–875.

Snowdon, C. T. (2009). *Plasticity of Communication in Nonhuman Primates*. *Advances in the Study of Behavior, 40,* 239–276.

Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, *55*, 673–687.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Cheng, L., West, V. & Slocombe, K. E., (2015). Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. *Current Biology*, 25, 495–499.

White, F.J. (1996). Comparative socio–ecology of *Pan paniscus*. In W. C. McGrew, L.F. Marchant & T. Nishida (Eds.), *Great Ape Societies* (pp. 29–41). Cambridge, UK: Cambridge University Press.

Wright, T. F., & Wilkinson, G. S. (2001). Population genetic structure and vocal dialects in an amazon parrot. *Proceedings of the Royal Society B : Biological Sciences*, *268*, 609–616.

SUPPLEMENTAL TABLES

Supplemental Table 1. This table diplays the loading score of each acoustic variable onto two factors. These two factors were then used to create the Euclidean acoustic distances between dyads.

	Factor loading score	
Acoustic variables	**PC1**	**PC2**
F_0-Start (Hz)	0.696	- 0.374
F_0-End (Hz)	0.490	- 0.450
F_0-Maximum (Hz)	0.603	- 0.301
Call Duration (s)	0.455	0.733
Slope–Start to Midpoint of F_0 (Hz/s)	- 0.522	- 0.043
Slope–Midpoint to End of F_0 (Hz/s)	0.196	0.150
Q1 Frequency (Hz)	0.764	- 0.469
Q1 Time (s)	0.528	0.566
Q3 Frequency (Hz)	0.460	- 0.265
Q3 Time (s)	0.483	0.724
Maximum Frequency (Hz)	0.622	- 0.364
Maximum Time (s)	0.545	0.615

Supplemental Table 2. This table displays the means for the absolute distance between dyads when averaged over all dyads in each category for each variable.

Mean absolute distance between dyads for each parameter	Unrelated (r=0)	Distantly related (r < 0.125)	Closely related (r ≤ 0.125)	Currently in different groups	Currently in same group
Euclidean distance (E.D.)	**1.05793**	**0.84276**	**0.86737**	**1.03046**	**0.88556**
F_0-Start (Hz)	**473.04233**	**398.17414**	**368.61489**	**492.29824**	**335.34336**
F_0-End (Hz)	303.657	292.63321	309.46151	**329.16831**	**246.10004**
F_0-Maximum (Hz)	**280.85345**	**186.0643**	**232.3027**	**265.47738**	**228.24976**
Call Duration (s)	0.03601	0.02911	0.0276	*0.03137*	*0.03724*
Slope–Start to Midpoint of F_0 (Hz/s)	**4322.31403**	**3149.82759**	**2694.49942**	3830.42545	3819.77099
Slope–Midpoint to End of F_0 (Hz/s)	**3215.94916**	**2233.51852**	**1872.38792**	2739.90107	2937.27926
Q1 Frequency (Hz)	**319.58623**	**233.55611**	**222.27665**	**307.32812**	**243.46563**
Q1 Time (s)	**0.02189**	**0.01773**	**0.0161**	0.0208	0.01874
Q3 Frequency (Hz)	433.9624	565.16697	493.13549	475.61001	460.6698
Q3 Time (s)	0.03078	0.02511	0.02816	*0.02808*	*0.03147*
Maximum Frequency (Hz)	319.23759	264.24365	256.61859	**325.35464**	**241.36255**
Maximum Time (s)	**0.02836**	**0.01939**	**0.02012**	0.02533	0.02495

Enduring voice recognition in bonobos

Sumir Keenan[1,2], Nicolas Mathevon[1,3], Jeroen MG Stevens[4], Jean Pascal Guéry[5], Klaus Zuberbühler[2,6] and Florence Levréro[1]

[1]Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Saint-Etienne, 42100, France

[2]University of St. Andrews, Department of Psychology & Neuroscience, St. Andrews, Scotland, UK

[3]Department of Psychology, Hunter College, CUNY, New York, USA

[4]Royal Zoological Society of Antwerp, Centre for Research and Conservation, Antwerp, Belgium

[5]Vallee des Singes Zoological Park, Romagne, France.

[6]Université de Neuchâtel, Department of Comparative Cognition, Neuchâtel, 2000, Switzerland

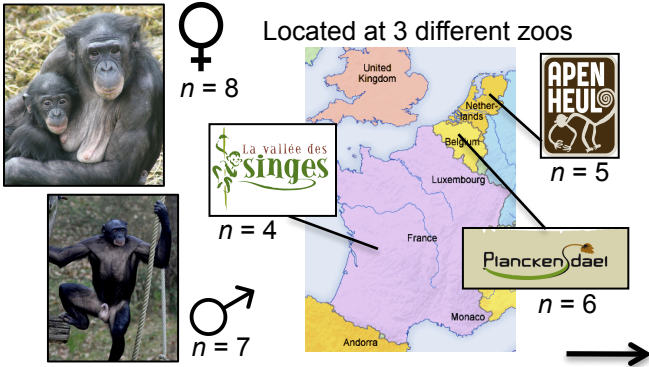MANUCSRIPT 3: Enduring voice recognition in bonobos.

Question

Can bonobos recognise past group members only from their voice?

Methods

15 adult individuals were played the vocalisations of a past group member
and an unfamiliar individual and their reactions were recorded

Subjects

♀
n = 8

Located at 3 different zoos

APEN HEUL
n = 5

La vallée des singes
n = 4

Planckendael
n = 6

♂
n = 7

Familiar pairs had been separated for varying numbers of years allowing testing of *long-term vocal recognition* :
2 – 3 years, n = 4 dyads
4.5 – 5.5 years, n = 8 dyads
8 – 9 years, n = 3 dyads

Each playback vocal stimuli consisted of a unique sequence of 4-6 individual calls
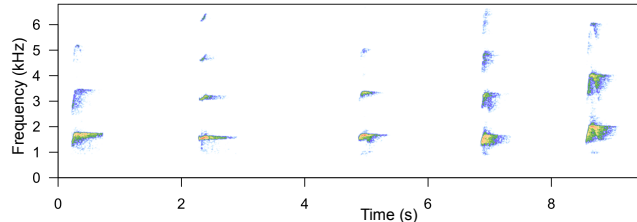
8 measured behaviours (in 60 seconds after a playback) were used to create a single principle component score for each individual for each playback.

	Factor loading score – PC1
Latency to first behaviour (after playback)	– 0.176
Latency to first locomotion	– 0.778
Duration of locomotion (in any direction)	0.873
Duration of locomotion toward speaker	0.771
Number of locomotion events	0.791
Duration looking toward speaker	0.488
Number of head movements (to speaker)	0.573
Number of head movements (any direction)	0.437

Results

Bonobos have significantly stronger reactions to familiar voices than unfamiliar.

Results of primary linear mixed model (LMM) investigating reactions to familiar versus unfamiliar voices

$P=0.008$
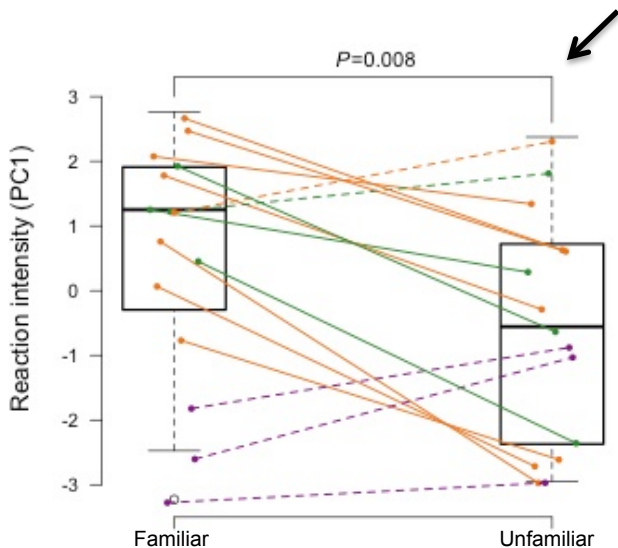
Results of two additional LMM **confirm** long-term vocal recognition, *but upper limit is not clear*
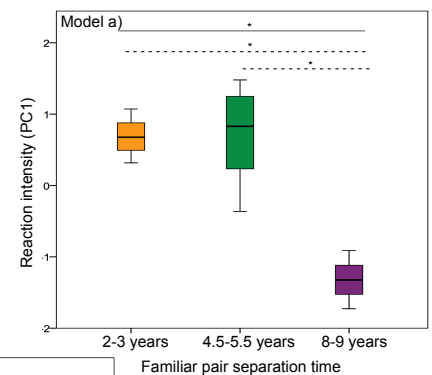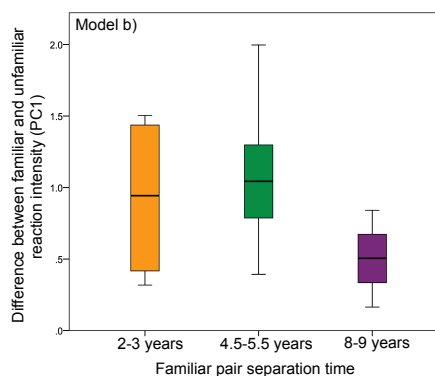
Model a)

Model b)

Model a) shows significant decrease in recognition after 8 years of separation, while model b) does not

139

SCIENTIFIC REP**O**RTS

Enduring voice recognition in bonobos

Sumir Keenan[1,2]**, Nicolas Mathevon**[1,3]**, Jeroen MG Stevens**[4]**, Jean Pascal Guéry**[5]**, Klaus Zuberbühler**[2,6] **& Florence Levréro**[1]

Long-term social recognition is vital for species with complex social networks, where familiar individuals can encounter one another after long periods of separation. For non-human primates who live in dense forest environments, visual access to one another is often limited, and recognition of social partners over distances largely depends on vocal communication. Vocal recognition after years of separation has never been reported in any great ape species, despite their complex societies and advanced social intelligence. Here we show that bonobos, *Pan paniscus*, demonstrate reliable vocal recognition of social partners, even if they have been separated for five years. We experimentally tested bonobos' responses to the calls of previous group members that had been transferred between captive groups. Despite long separations, subjects responded more intensely to familiar voices than to calls from unknown individuals - the first experimental evidence that bonobos can identify individuals utilising vocalisations even years after their last encounter. Our study also suggests that bonobos may cease to discriminate between familiar and unfamiliar individuals after a period of eight years, indicating that voice representations or interest could be limited in time in this species.

The social life of many primate species is characterised by lasting associations between individuals, making individualised social knowledge of primary importance. Individual vocal recognition has been found in many primate species[1–6] and is particularly important for those living in dense forest habitats where vocalisations are often the most efficient communication channel[7]. For primate species with fission-fusion dynamics, community members regularly separate into smaller, fluid parties for hours, days or weeks while often maintaining vocal contact[8]. In many primate species, individuals disperse from their natal community during puberty[9], but often continue to interact with former group members during subsequent inter-community encounters. Thus, successful social navigation within a community and between communities may depend on the ability to recognise both current and previous social partners. As a whole, long-term vocal recognition has only been demonstrated in a limited number of birds and mammals[10–14], including only two monkey species with stable, non-fission-fusion group structures[15–16] and has never been investigated in any ape species.

The present study focuses on the bonobo, an ape living in dense equatorial rainforest with large, overlapping home ranges and complex fission-fusion social networks between related and unrelated individuals[17]. It has been suggested that bonobos use vocalisations to communicate with distant conspecifics[7–8], and some ape species have recently demonstrated a capacity for long-term memory for finding tools in the distant past, suggesting they could also retain social information for similar time periods[18]. We therefore hypothesised that long-term vocal recognition would be present in this species as a valuable adaptation for mediating their social environment. We used a series of playback experiments to test long-term vocal recognition in bonobos by comparing their behavioural responses to the vocalisations of familiar and unfamiliar conspecifics. Bonobos are known to display less aggression towards new individuals than their closest relatives, chimpanzees[19–20]. However, despite their relatively tolerant nature they still can react with mild aggression (with motor and vocal displays), caution or complete avoidance during intercommunity encounters[19–20]. This suggested that bonobos would likely react more cautiously (e.g. fewer approaches to the loudspeaker, less overall movement) to an unfamiliar voice than a familiar one, but that it would fail to induce extreme reactions such as aggression or panic. All familiar individuals had been separated from our tested subjects for varying numbers of years; in captivity individuals are sometimes

[1]Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Saint-Etienne, France. [2]University of St. Andrews, School of Psychology & Neuroscience, St. Andrews, Scotland, UK. [3]Department of Psychology, Hunter College, CUNY, New York, USA. [4]Royal Zoological Society of Antwerp, Centre for Research and Conservation, Antwerp, Belgium. [5]Vallée des Singes Zoological Park, Romagne, France. [6]Université de Neuchâtel, Department of Comparative Cognition, Neuchâtel, Switzerland. Correspondence and requests for materials should be addressed to F.L. (email: florence.levrero@univ-st-etienne.fr)
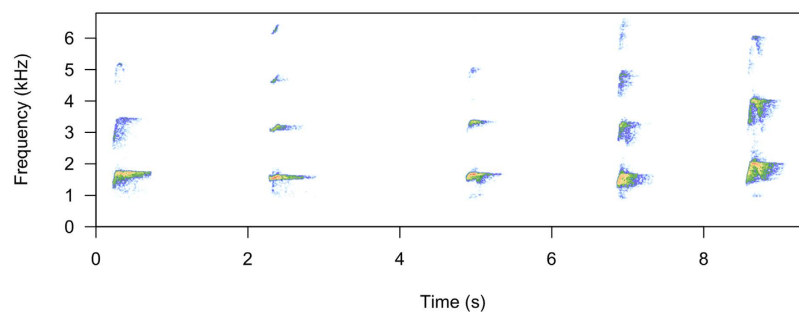
Figure 1. An example of a playback stimulus. This sequence was produced by an adult bonobo female living at Apenheul Zoo and was used to test a familiar condition at Planckendael Zoo.

Behavioural variables	Factor loading score
	PC1
	−0.176
Latency to first locomotion	**−0.778**
Duration of locomotion (in any direction)	**0.873**
Duration of locomotion toward speaker	**0.771**
Number of locomotion events	**0.791**
Duration looking toward speaker	0.488
Number head movements oriented toward speaker	0.573
Total number of all head movements	0.437

Table 1. Factor Loadings of measured behavioural variables on the first Principal Component (PC1). Factors that loaded highly onto PC1 are in bold.

transferred between zoos for population management and breeding programs. This, along with detailed life histories of individuals housed at three European zoos (Apenheul, Netherlands; Planckendael, Belgium, and La Vallée des Singes, France), allowed us to identify 15 individuals who had previously been housed with another individual living at one of the other zoos (Methods).

At each zoo, prior to the start of the playback experiment, we carefully mimicked a transfer of new bonobos and hid a loudspeaker in the enclosure where new individuals are normally held upon arrival. All of the subjects had previously experienced a real transfer event, by being transferred themselves (14 of 15 individuals), being in a group when a new individual was brought in (14 of 15) or both (13 of 15); therefore, this method increased the chances that they believed the broadcast calls were emitted by real individuals. At each zoo the experiment consisted of one mock transfer event followed by five playback trials, which occurred over a single day. After a playback trial, we waited until the whole group returned to 'baseline' behaviours, such as feeding, foraging, grooming or resting, before beginning the next trial (with a minimum of 10 minutes between trials). In keeping with the illusion that a real transfer was occurring, calls used for the playback stimuli were selected on the basis of acoustic similarities to vocalisations recorded during an actual transfer event (Fig. 1, Methods, Supplemental Table S1), and for each trial, the playback stimulus was composed of a unique call sequence (Fig. 1, Methods, Supplemental Audio 1). By using multiple observers we were able to test multiple subjects with each playback trial, this allowed us to minimise the number of playback trials in order to reduce the risk of habituation. Each observer recorded the behaviour of a single subject with a video camera. Each subject was recorded once in the familiar condition and once in the unfamiliar condition (See Methods).

Results

Vocal recognition of previous group members. The bonobos' reactions to the playbacks were assessed using 8 behavioural variables encompassing locomotion, looking direction and latency of behaviours after the playback (Methods). These measurements were then collapsed into a single composite behavioural score using a principle component analysis (Methods-Table 1). We found that bonobos responded more intensely when hearing a familiar voice compared to an unfamiliar voice ($n = 15$, linear mixed model, $t = -0.396$, $P = 0.014$, Methods-Table 2a). When hearing a familiar voice they responded more rapidly, increased their locomotion and approached the speaker more (Fig. 2, Supplemental Video 1). We also tested for any effects of the subject's (the receiver) sex, rank and trial number. Additionally, as some individuals were subadults when housed with their previous social partner, any effect of age was tested (current age was used). Importantly, none of these factors were found to significantly influence the bonobos' responses, highlighting the robustness of our findings.

The effect of separation time on recognition. The playback experiments also allowed us to investigate the dynamics of long-term vocal recognition as familiar pairs (familiar condition) had been separated for varying

	Estimate	Standard Error	t	P
(a) MODEL 1				
(Intercept)	2.539	2.197	1.156	
Trial Condition (Familiar V. Unfamiliar)	−1.151	2.364	−0.396	0.014
Subject Rank	0.801	0.755	1.060	0.321
Subject Sex	−0.696	1.006	−0.692	0.510
Subject Age	0.012	0.057	0.302	0.852
(b) MODEL 2				
(Intercept)	2.491	1.713	1.454	
Separation Time	−0.066	0.013	−5.230	< 0.0001
Subject Rank	0.526	0.513	1.024	0.342
Subject Sex	0.625	0.713	0.876	0.388
Subject Age	0.030	0.041	0.734	0.516
(c) MODEL 3				
(Intercept)	1.737	0.625	2.781	
Separation Time	−0.088	0.071	−1.240	0.226
Subject Rank	−0.398	0.474	−0.840	0.394
Subject Sex	0.444	0.337	1.317	0.200
Subject Age	−0.02	0.022	−0.944	0.352

Table 2. Results of LME models. (**a**) Model 1 tested for the effect of vocal familiarity on bonobos' response to playbacks. (**b,c**) Models 2 and 3 tested the effect of separation time on bonobos' response to previous partner's voice. Model 2 examines responses to the familiar individual playback alone while model 3 uses the relative difference for each individual in response to familiar vs. unfamiliar individuals. n = individuals.
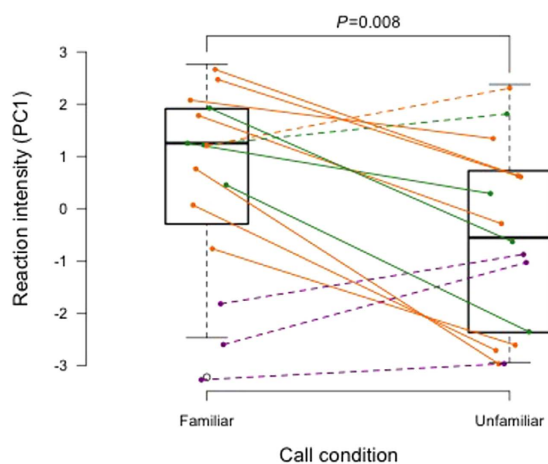
Figure 2. Bonobo reactions to the calls of familiar and unfamiliar individuals. Each individual was tested in both conditions, and each line on the figure links the responses in each condition for the same individual. The colour of the lines corresponds to the separation time between the subject and the former group mate used in the familiar condition. Green lines = bonobos that have been separated for 2–3 years (n = 4); orange lines = separated for 4.5–5.5 years (n = 8); purple lines = separated for 8–9 years (n = 3). The principle component score (PC1) represents an integrated measure of the behavioural response, with higher scores indicating a stronger behavioural reaction to the broadcast calls. Solid lines = bonobos that reacted more to the familiar voice; dashed lines = bonobos that reacted equally to both signals or more to the unfamiliar voice.

numbers of years (separation time: 2–3 years, n = 4; 4.5–5.5 years, n = 8; 8–9 years, n = 3). Two statistical models were run to test the effect of separation time on the bonobos behavioural responses to familiar individuals. The first model investigated the magnitude of variation between the behavioural responses to each of the three separation time categories and found a significantly decreased response towards a past partner's voice in dyads that had been separated for more than 8 years (n = 15; LMM: t = −5.230, P < 0.0001, Methods-Table 2b) (Fig. 3a). Post-hoc tests showed no significant differences in the bonobos' reactions for pairs separated between 2–3 years and 4.5–5.5 years (multiple comparison test: z = −0.651, P = 0.784). Conversely, both were significantly different from the reactions of dyads separated for 8–9 years (multiple comparison tests between 2–3 years and 8–9 years: z = −4.802, P < 0.0001, between 4.5–5.5 years and 8–9 years: z = −6.707, P < 0.0001). As this first model only considered variation in the responses to familiar individuals, we analysed the data with a second model by
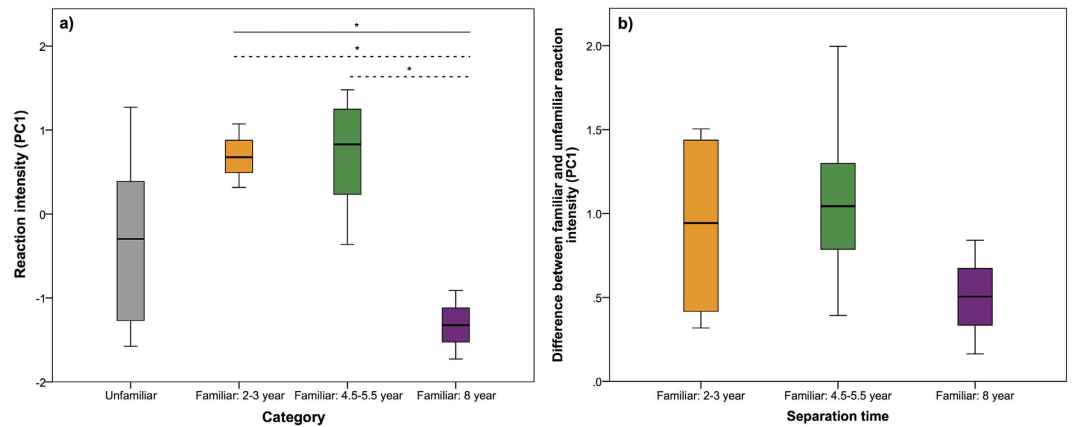
Figure 3. Effect of separation time on recognition. (**a**) Model 1 investigated the effects of the time of separation on the responses of subjects to a previous group member by comparing the reaction intensity (measured by the first principle component–PC1) between the three different separation categories (separated by 2–3 years, 4.5–5.5 years, or 8 years) (*$p < 0.001$; dashed lines are results of post-hoc comparisons). The PC1 scores for the unfamiliar condition are also presented on the graph as a reference. (**b**) Model 2 also investigated the effects of separation time by using a more conservative model comparing the absolute difference between the PC1 score for the unfamiliar and familiar conditions between the three separation categories. Despite showing the same pattern as model 1, the result of model 2 was not significant.

measuring the absolute difference between each subjects' response to familiar and unfamiliar individuals. While this second model shows the same trend as the first model (Fig. 3b), the difference between the three separation time groups was not significant ($n = 15$; LMM, $t = -1.240$, $P = 0.226$, Methods - Table 2c, Fig. 3b).

Discussion

Here we provide the first experimental evidence that an ape species is capable of long-term vocal recognition of former social partners even after five years of separation. Our results demonstrate the importance of individualised vocal signalling for bonobos, which helps them navigate a complex fission-fusion society by maintaining communication between community members with whom they are out of physical range. In the current experiment we utilised more short-range calls than are typically displayed in these wild inter- and intra-community interactions[7]. However, within the bonobo vocal repertoire many call types are known to carry a variety of information–for instance, call sequences can carry information about different food types[21] and copulation calls show individual vocal signatures[22]. Our results thus further demonstrate that loud long-calls are not alone in transmitting identity information, and that bonobos are able to use more short-ranged vocalisations to recognise familiar individuals. Additionally, in bonobos, inter-community encounters regularly begin with vocal exchanges, which appear to set the tone for the following interaction and generally result in either avoidance or peaceful interactions in which the two groups may even feed together[19–20]. In such a scenario, long-term vocal recognition enables bonobos to identify individuals without visual access, allowing them to favour meetings with affiliative individuals and avoid individuals with whom they have a conflictive relationship.

Despite the evidence of vocal recognition displayed by the experiment, our results suggest decreased reactions to former group members after eight years of separation, supported by statistical model 1 and underlying a possible limit to long-term vocal recognition (Fig. 3a). It is possible that bonobos are unable to recognise past social partners after a long period without contact, either because they cannot retain the memory of individual vocal signatures for longer than 6 to 8 years without reinforcement, or because a bonobo's voice significantly changes over time as they age. As bonobos are very visual creatures, it could also be that long-term social recognition is strongest in the visual modality. Even a high degree of relatedness between the pairs in the eight-year separation category did not appear to secure long-term recognition (See Supplemental Table S2), despite the fact that mother-son bonds are particularly strong in bonobos[19].

Alternatively, bonobos may in fact recognize the identity of the caller but are not motivated to react, as their social networks are highly dynamic, bonding and dominance between individuals may need to be reset after a significant period of separation.

However, we obtained contrasting support from the two statistical models testing the effects of long-term recognition. Despite a clear trend in the raw data (Fig. 3b), the more conservative model 2 did not support an upper limit of vocal recognition. Indeed, as bonobos have demonstrated advanced cognitive abilities and social skills[17,23,24] and have performed well in memory tasks[25], a social memory longer than 5 years was hypothesised. While long-term vocal memory has not been tested in any ape species or humans, both have demonstrated at least a decade long memory for the faces of former group mates, suggesting the potential for long-term social memory in the vocal modality[26,27]. A recent study showed that dolphins, another species with highly developed social and cognitive skills that also live in fission-fusion societies, could retain a vocal memory of conspecifics for decades[14] and anecdotal evidence from another study suggests a similar vocal memory in African elephants[12]. Importantly, as the bonobos' reactions were highly idiosyncratic we cannot exclude the possibility that this result is due to a

low sample size ($n = 3$) for individuals separated for more than 8 years. This discrepancy between both models may be explained by our sample size; therefore, to conclusively demonstrate the upper limit of bonobo vocal recognition further investigation is required.

Methods

Subjects. In order to test the long-term vocal recognition of bonobos we benefited from the transfers of captive individuals between zoos for population management and breeding programs–these movements can, in a way, mimic their migration patterns in the wild. After examining the life histories of the 34 individuals housed in multi-female, multi-male groups between three European zoos (Apenheul, Netherlands; Planckendael, Belgium; La Vallée des Singes, France), we found 15 individuals who had been housed together in the past (either at one of the three zoos or at other zoos throughout Europe). We only included individuals who had been housed together over the age of seven. At the time of the experiment the youngest tested subject was 10 years old and the oldest 45 (mean age = 21; median age = 19).

Each pair had been housed together for long periods of time (from 4 to 17 years) but had since been housed separately for 2 to 9 years. The 15 subjects, aged 10 to 45, were equally balanced across sex (male $n = 7$; female $n = 8$), rank (High rank $n = 5$; mid-rank $n = 6$; low rank $n = 4$) and zoo (Apenheul $n = 5$; Planckendael $n = 6$; La Vallée des Singes $n = 4$). All individuals were subadults or adults (≥ 7 years-old) when housed with their previous group member and aged 10 or more during the experiment (Supplemental Table S2). At each zoo, an individual's rank was assigned categorically based on the following: during agonistic encounters, whether mild such as food competition or severe such as conflicts resulting in serious injury, if an individual was most likely to be the aggressor they were classified as high-ranking, if an individual could be an aggressor or a victim, depending on the identity of the social partner during an event, they were classified as mid-ranking, and if an individual was most likely to the victim they were classified as low-ranking. Each rank assignment was then discussed and verified with a minimum of two keepers at each zoo.

Playback Stimuli. Calls used for the playback stimuli were taken from a databank of vocal recordings from all individuals at the three zoos amassed by SK in 2013. Calls were selected on the basis of an acoustic similarity to vocalisations recorded during an actual transfer event–where two individuals were transferred together to La Vallée des Singes in 2012–and can be described as peep-yelps[28] (See Supplemental Table S1 for acoustic details). Individual call sequences contained 4–6 calls and had decreasing intercall intervals along the progression of the sequence, as in the call sequence heard during the real transfer, with a mean total sequence time of 5.70 s (range of 3.79 s–7.882 s) for Apenheul and Planckendael. Because the enclosure at La Vallée des Singes is more than two-times larger than the enclosures at the other two zoos, the call sequences played there were followed by two additional calls 10 seconds later, to ensure all bonobos heard the playback sequences. All sequences were broadcast between 65–80 dB, SPL measured at 1 meter from the loudspeaker.

Playback Experiments. At each zoo the experiment began with a mock transfer of new bonobos, following their standard transfer procedure, respectively. Everything was done as if real bonobos were being moved–a truck/tractor was driven up to the building carrying a crate, the crate was placed to the slide opening to the separation cage, and the slides were opened and closed. As all of the subjects had previously experienced a real transfer event in similar conditions, we expected that the mimicked transfer would have bonobos believe that other individuals were arriving. During this time a loudspeaker (Juster Elite Speaker for Apenheul and Planckendael, and Western Rivers Nite Stalker Pro for La Vallée des Singes) was placed in a separation enclosure where new bonobos are normally kept upon arrival. All physical and visual access to this enclosure was blocked during the experiment and for at least 12 hours before the mock transfer. Beginning from 10–15 minutes prior to the mock transfer three to four observers set-up around the cages where bonobos are usually housed, and where the experiment was to occur. The bonobos at all three zoos are regularly observed by researchers and did not show any visible signs of disturbance. At each zoo the experiment consisted of a single mock transfer followed by a total of five playback trials. After the mock transfer we waited until the group returned to baseline behaviours-resting, foraging or grooming-before broadcasting the first playback. Before proceeding from the first to subsequent playbacks we again waited until the group returned to baseline behaviours–therefore time between playbacks varied from 10 minutes to 37 minutes (mean = 25 minutes) for Apenheul and Planckendael. Due to environmental conditions at La Vallée des Singes the first playback occurred four hours before the proceeding four trials (which then averaged 44 minutes between broadcasts).

Each playback, whether familiar or unfamiliar, contained a unique, acoustically distinct, set of calls. For all three zoos the voice of each familiar individual was broadcast only once (number of past group members used for the familiar trials: Planckendael, $n = 4$; Apenheul, n = 3; La Vallée des Singes, $n = 2$). The call sequences used for the unfamiliar trials at Apenheul came from a single female unknown to all individuals in the group, this was also the case for La Vallée des Singes. At Planckendael, there was not a single individual within our database that was unknown to everyone; therefore at this zoo each broadcast individual was familiar to some and unfamiliar to others (see Supplementary Table S4).

In total each subject was tested once in each of the two experimental conditions. As the enclosures at each zoo were of different sizes, shapes and contained different climbing structures, a precise distance from the speaker across all subjects during playback trials could not be set. To control for this variation we able to test the majority of individuals at roughly the same distance from the speaker for both of their playback trials (Within 0–2 metres of same position for both trials, $n = 12$; within 2–4 metres, $n = 2$; > 4 metres, $n = 1$; See Supplemental Table S3). Additionally, the order in which each bonobo heard the stimulus for each condition was counterbalanced (8 individuals heard the familiar first while 7 heard the unfamiliar first).

Due to limitations in the number of researchers able to observe during the experiments we were not able to test all individuals in the first two playback trials. As such, some individuals were exposed to familiar and/or unfamiliar voices before their responses to either or both conditions could be recorded (See Supplemental Table S4). We controlled for stimuli exposure (we did not differentiate between familiar and unfamiliar playbacks) prior to their own trials by including trial number as a random factor in the model.

To be able to test all 15 subjects in a realistic situation, we found that not disturbing the groups' normal composition was the best option. While this choice led to the potential for pseudoreplication, it avoided the stress caused to individuals by separating their normal group and allowed us to test multiple individuals at once, which helped to avoid habituation. This possible non-independence of the reactions of tested individuals, together with the fact that we took multiple individuals' reactions from the same playback trial, was controlled for by entering trial number as a random factor in statistical models (see statistical analysis of behavioural reactions below).

Measurements of behavioural responses. Each observer filmed (Handheld camera models: Canon Legria FS406 and Canon Legria HF200; Stable camera models : GoPro Hero3 and JVC GC-XA1 Adixxion HD) one focal subject that was randomly assigned. Subjects were video recorded for 10 minutes before and 10 minutes after each playback trial. SK coded all videos. To ensure unbiased coding, all videos were given numbers and coded blind to the condition a minimum four weeks after the experiment occurred. To inform which behavioural measures would be included we relied on reported behaviours when bonobos encounter neighbouring groups in the wild[20], observations by SK during a transfer of a female into the Apenheul group and on previous studies investigating vocal recognition in a variety of species[2,4,11,16,29]. As stated in the introduction, bonobos are known for their relatively tolerant nature towards foreign individuals, therefore we expected a mild intensity reaction to the playback broadcasts. We assessed a variety of measures on body and head movements, in relation to the speaker and in general. Social interactions and vocal responses were also coded; however, our playbacks elicited no interactions between individuals (neither aggressive or affiliative) and only one vocalisation (a single call by an unidentified individual). Therefore, the following eight behavioural variables, measured in the 60 seconds following the playback, were included:

- latency after the start of the playback to the first behaviour displayed. It could have been any behaviour or the cessation of a behaviour–for instance if they were eating and stopped;
- latency to the first locomotion after the start of the playback broadcast;
- total duration of locomotion in any direction;
- total duration of locomotion toward the speaker;
- number of separate locomotion occurrences;
- duration of time spent looking toward the speaker;
- number of times an individual looked toward the speaker;
- total number of head movements (each change of head direction was counted as one movement event);

To conduct inter-observer reliability FL coded 67% of the videos as above and results were compared for each variable separately. The intraclass correlation coefficient (ICC) was above 0.860 for all variables (Latency to first behaviour = 0.868; latency to first locomotion = 0.908; duration of locomotion (in any direction) = 0.984; duration of locomotion toward speaker = 0.969; number of locomotion events = 0.878; duration looking toward speaker = 0.895; number of head movements oriented toward speaker = 0.920; total number of all head movements = 0.874).

Instead of separately analysing the 8 dependent behavioural measures, we performed a principal component analysis (PCA) and retained a single composite score. By using a PCA, we approached a Gaussian distribution, built an integrated measure of the behavioural response and demonstrated which behavioural variables were important. As shown in Table 1, latency to the first locomotion, duration of locomotion, the number of movements, and the propensity to move towards the loudspeaker were the main factors that loaded on the first PC score (PC1). PC1 was thus chosen as a unique composite score representing the strength of an individual's behavioural response to a playback (with positive scores indicating a stronger behavioural response and negative scores representing a reduced reaction).

Statistical analysis of behavioural reactions. Each individual ($n = 15$) was tested with one familiar voice and one unfamiliar voice. The eight behavioural variables were measured in the 60 seconds following the playback and used in a principle component analysis, resulting in one principle component (PC1), which explained 42.2% of the variation in the data. To test for an effect of familiarity versus non-familiarity of the playback stimuli on the bonobos' behavioural response, we used a linear mixed effect model with PC1 as the dependent measure (R package lme4), after checking the distribution of the residuals with respect to normality and homoscedasticity (fixed effects: subject rank, subject sex, subject age; random effects: individual identity, playback trial number, zoo location). P values were obtained with likelihood-ratio tests comparing the fit of the full model with reduced models lacking fixed effects. To test for the effect of separation time, we used two different models. The first model was restricted to experiments with calls from past (familiar) partners. This analysis was followed by post-hoc multiple comparison tests (function glht in multcomp R package). Given individual differences in reactivity to playbacks and the small number of individuals having experienced the same separation time, we used a second, more conservative, model that took into account the relative difference in response to familiar and unfamiliar calls (for each individual we calculated the absolute difference in their PC1 score between the two conditions). The degree of relatedness between the pairs was initially considered, however it was not balanced across separation time conditions, and therefore could not be accurately tested (all three of the eight-year separation category were 1st degree related, while only one of the other 12 pairs was–Supplemental Table S2).

Ethics Statement.   All work was performed in accordance with the relevant guidelines and regulations, and all experimental protocols were approved by the Institutional Animal Ethical Committee of the University of Lyon/Saint-Etienne, under the authorization no. 42-218-0901-38 SV 09 (Lab ENES).

References

1. Cheney, D. L. & Seyfarth, R. M. Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour* **28,** 362–364 (1980).
2. Rendall, D., Rodman, P. S. & Emond, R. E. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour* **51,** 1007–1015 (1996).
3. Kojima, S., Izumi, A. & Ceugniet, M. Identification of vocalizers by pant hoots, pant grunts and screams in a chimpanzee. *Primates* **44,** 225–230 (2003).
4. Ramos-Fernández, G. Vocal Communication in a Fission-Fusion Society: Do Spider Monkeys Stay in Touch With Close Associates? *International Journal of Primatology*, **26** 1077–1092 (2005).
5. Miller, C. T. & Thomas, A. W. Individual recognition during bouts of antiphonal calling in common marmosets. *Journal of Comparative Physiology* **198,** 337–346 (2012).
6. Candiotti, A., Zuberbuhler, K. & Lemasson, A. Voice discrimination in four primates. *Behavioural Processes* **99,** 67–72 (2013).
7. Hohmann, G. & Fruth, B. Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology* **15,** 767–782 (1994).
8. White, F. J. In *Great Ape Societies* (eds McGrew, W. C., Marchant, L. F. & Nishida, T.) 29–41 (Cambridge Univ. Press, Cambridge, 1996).
9. Pusey, A. E. & Packer, C. In *Primate Societies* (eds Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W. & Struhsaker, T. T.) 250–266 (University of Chicago Press, Chicago, 1987).
10. Godard, R. Long-term memory of individual neighbours in a migratory songbird. *Nature* **350,** 228–229 (1991).
11. Insley, S. J. Long-term vocal recognition in the northern fur seal. *Nature* **406,** 404–405 (2000).
12. McComb, K., Moss, C., Sayialel, S. & Baker, L. Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour* **59,** 1103–1109 (2000).
13. Boeckle M. & Bugnyar T. Long-term memory for affiliates in ravens. *Current Biology* **22,** 801–806 (2012).
14. Bruck, J. N. Decades-long social memory in bottlenose dolphins. *Proceedings of the Royal Society B* **280,** 1–6 (2013).
15. Lemasson, A. & Hausberger, M. Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *The Journal of the Acoustical Society of America* **129,** 3341–3352 (2011).
16. Matthews, S. & Snowdon, C. T. Long-term memory for calls of relatives in cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology* **125,** 366–369 (2011).
17. Kanō, T. *The last ape: Pygmy chimpanzee behavior and ecology*. (Stanford University Press, Stanford, California, 1992).
18. Martin-Ordas, G., Berntsen, D. & Call, J. Memory for distant past events in chimpanzees and orangutans. *Current Biology* **23,** 1438–1441 (2013).
19. Furuchi, T. Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthrophology* **20,** 131–142 (2011).
20. Hohmann, G. & Fruth, B. In *Behavioural Diversity in Chimpanzees and Bonobos* (eds Boesch, C., Hohmann, G. & Marchant, L. F.) 138–150 (Cambridge Univ. Press, Cambridge, 2002).
21. Clay, Z. & Zuberbuhler, K. Bonobos extract meaning from call sequences. *PLoS One* **6,** e18786 (2011).
22. Clay, Z. & Zuberbuhler, K. The Structure of Bonobo Copulation Calls During Reproductive and Non-Reproductive Sex. *Ethology* **117,** 1158–1169 (2011).
23. Hare, B., Melis, A. P., Woods, V., Hastings, S. & Wrangham, R. Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology* **17,** 619–623 (2007).
24. Herrmann E, Hare B, Call J & Tomasello M. Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLoS One* **5,** e12438 (2010).
25. Martin-Ordas, G., Haun, D., Colmenares, F. & Call, J. Keeping track of time: evidence for episodic-like memory in great apes. *Animal Cognition* **13,** 331–340 (2010).
26. Bahrick, H. P., Bahrick, P. O. & Wittlinger, R. P. Fifty years of memory for names and faces: a cross sectional approach. *Journal of Experimental Psychology: General* **104,** 54–75 (1975).
27. Hanazuka, Y., Shimahara, N., Tokuda, Y. & Midorikawa, A. Orangutans (Pongo pygmaeus) remember old acquaintances. *PLoS One* **8,** e82073 (2013).
28. deWaal, F. B. M. The communicative repertoire of captive bonobos (Pan paniscus) compared to that of chimpanzees. *Behaviour* **27,** 183–251 (1988).
29. Bergman, T. J. Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proceedings of the Royal Society B* **277,** 3045–3053 (2010).

Acknowledgements

Author Contributions

F.L. proposed the project. S.K., F.L., J.M.G.S. and K.Z. developed the experimental protocol and designed the set-up. S.K. recorded bonobo vocalisations used as broadcast sounds in the experiment and edited all audio files with the help of F.L. and J.M.G.S. facilitated approval to conduct the experiment at Planckendael and Apenheul. J.P.G. facilitated approval to conduct the experiment at Vallée des Singes. S.K. and J.M.G.S. conducted the experiments at Planckendael and Apenheul, and S.K. conducted the experiment at Vallée des Singes. N.M., F.L. and S.K. performed the analysis. S.K. wrote the manuscript with editing from N.M., F.L. and J.M.G.S.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Keenan, S. *et al.* Enduring voice recognition in bonobos. *Sci. Rep.* **6**, 22046; doi: 10.1038/srep22046 (2016).

SUPPLEMENTAL TABLES

(For Manuscript 3 – Enduring voice recognition in bonobos)

Supplemental Table S1. Acoustic features of the calls recorded during a real transfer event and the calls used for the broadcast sequences in the playback experiments.

		Calls from real transfer event (*n*= 8)	Calls used for playbacks (*n*=50)
Start Frequency (Hz)	Mean	1683	1542
	S.D.	120	218
	Min	1497	1138
	Max	1821	2309
End Frequency (Hz)	Mean	1772	1580
	S.D.	133	209
	Min	1596	1227
	Max	1896	2021
Frequency modulation (Hz)	Mean	90	37
	S.D.	47	129
	Min	0	-288
	Max	150	316
Duration (s)	Mean	.231	.224
	S.D.	.070	.075
	Min	.12	.14
	Max	.32	.51

Supplemental Table S2. Tested subjects (age, sex, zoo of residence, rank) and their relationship with the individuals whose voices were broadcast in the familiar condition during playback experiments.

Subject	Age	Sex	Residence	Rank	Familiar Call Conditions			
					Separation Time (months)	Name of familiar	Sex of familiar	Subject and familiar relation
Daniela	45	F	VDS	High	54	Lina	F	Unrelated
Kirembo	21	M	VDS	Low	54	Lina	F	Unrelated
Diwani	17	M	VDS	Low	54	Louisoko	M	Paternal half-brother
David	12	M	VDS	Mid	54	Louisoko	M	Unrelated
Lina	28	F	Planck	High	54	Khaya	F	Unrelated
Louisoko	16	M	Planck	Mid	54	Khaya	F	Unrelated
Lucuma	11	M	Planck	Low	54	Khaya	F	Unrelated
Djanoa	19	F	Planck	Mid	65	Hortense	F	Unrelated
Vifijo	19	M	Planck	Low	100	Hortense	F	Mother-son
Busira	10	F	Planck	Low	32	Lisala	F	Unrelated
Jill	29	F	Apen	High	33	Lingala	F	Mother-daughter
Hortense	36	F	Apen	Mid	100	Vifijo	M	Mother-son
Zuani	22	F	Apen	High	33	Lingala	F	Unrelated
Kumbuka	14	F	Apen	Mid	33	Lingala	F	Unrelated
Zamba	16	M	Apen	Mid	100	Vifijo	M	Full Brothers

Supplemental Table S3. Playback experimental details for each subject's familiar and unfamiliar conditions, including the principal component scores (PC1) calculated from their behavioural reactions to the playback stimuli.

Subject	Zoo	Familiar Condition				Unfamiliar Condition			
		Play-back Trial No.	Distance from speaker (in metres)	Individuals familiar to PB voice / Total number of individuals in group*	PC1 Score	Play-back Trial No.	Distance from speaker (in metres)	Individuals familiar to PB voice / Total number of individuals in group*	PC1 Score
Hortense	Apen**	3	6	0/5	-1.727	1	6	0/5	-1.563
Jill	Apen**	2	8	3/5	0.683	1	8	0/5	1.001
Kumbuka	Apen**	2	4	1/2	1.072	4	2	0/2	-0.298
Zamba	Apen**	3	2	0/2	-1.324	1	2	0/2	-0.483
Zuani	Apen**	2	5	3/5	0.317	3	6	0/5	-1.187
Busira	Planck	3	8	0/8	0.669	1	8	3/8	0.153
Djanoa	Planck	1	5	2/8	0.974	2	12	4/8	-0.134
Lina	Planck	4	8	3/8	-0.363	3	8	1/8	-1.350
Louisoko	Planck	4	12	3/8	1.479	3	8	1/8	0.390
Lucuma	Planck	4	8	3/8	1.381	3	8	1/8	0.383
Vifijo	Planck	1	4	2/8	-0.912	2	3	4/8	-0.407
Daniela	VDS	1	7	4/14	1.116	2	7	0/14	0.723
David	VDS	3	9	4/14	0.046	4	9	0/14	-1.442
Diwani	VDS	3	14	4/14	0.422	4	18	0/14	-1.575
Kirembo	VDS	1	9	4/14	0.684	2	9	0/14	1.271

*Number of individuals in each group familiar to the voice in the playback broadcast, excluding subject / total number of individuals in the group, including subjects and non-subjects but excluding individuals under the age of four. **The Apenheul group have lived in two sub-groups since 2013, during the playback experiment Zamba and Kumbuka were completely separate from the other individuals - therefore, total number of familiar individuals listed in this column only includes individuals within his or her sub-group and not across the whole Apenheul population.

Supplemental Table S4: Details on each playback trial at each zoo, including when each individual was recorded for the familiar and unfamiliar conditions and their familiarity to each playback. Trials marked as **'Observed'** are each individuals' trials that were included in the analysis for the current study, along with the trial's condition. Five playbacks occurred at each zoo, however only the first four were needed for analysis.

a) Apenheul

Playback Trial Number	Playback Stimulus	Hortense	Jill	Kumbuka	Zamba	Zuani
1	Lisala #1	**Observed - Unfamiliar**	**Observed - Unfamiliar**	Present - Unfamiliar	**Observed - Unfamiliar**	Present - Unfamiliar
2	Lingala #1	Present - Familiar	**Observed - Familiar**	**Observed - Familiar**	Present - Familiar	**Observed - Familiar**
3	Vifijo #1	**Observed - Familiar**	Present - Unfamiliar	Present - Unfamiliar	**Observed - Familiar**	**Observed - Unfamiliar**
4	Lisala #2	Present - Unfamiliar	Present - Unfamiliar	**Observed - Unfamiliar**	Present - Unfamiliar	Present - Unfamiliar

b) Planckendael

Playback Trial Number	Playback Stimulus	Busira	Djanoa	Lina	Louisoko	Lucuma	Vifijo
1	Hortense #1	**Observed - Unfamiliar**	**Observed - Familiar**	Present - Unfamiliar	Present - Unfamiliar	Present - Unfamiliar	**Observed - Familiar**
2	Daniela #1	Present - Unfamiliar	**Observed - Unfamiliar**	Present - Familiar	Present - Familiar	Present - Familiar	**Observed - Unfamiliar**
3	Lisala #1	**Observed - Familiar**	Present - Unfamiliar	**Observed - Unfamiliar**	**Observed - Unfamiliar**	**Observed - Unfamiliar**	Present - Unfamiliar
4	Khaya #1	Present - Unfamiliar	Present - Unfamiliar	**Observed - Familiar**	**Observed - Familiar**	**Observed - Familiar**	Present - Unfamiliar

c) La Vallée des Singes

Playback Trial Number	Playback Stimulus	Daniela	David	Diwani	Kirembo
1	Lina #1	**Observed - Familiar**	Present - Familiar	Present - Familiar	**Observed - Familiar**
2	Djanoa #1	**Observed - Unfamiliar**	Present - Unfamiliar	Present - Unfamiliar	**Observed - Unfamiliar**
3	Louisoko #1	Present - Familiar	**Observed - Familiar**	**Observed - Familiar**	Present - Familiar
4	Djanoa #2	Present - Unfamiliar	**Observed - Unfamiliar**	**Observed - Unfamiliar**	Present - Unfamiliar

150

CONCLUSION AND PERSPECTIVES

IV. CONCLUSION AND PERSPECTIVES

A. Individual vocal signatures and vocal recognition

The acoustic and statistical analyses of five common call types of the bonobo vocal repertoire all showed significant evidence of individual vocal signatures (manuscript 1). Given the importance of individual recognition, the common use of the vocal modality in bonobo communication, and the extensive evidence of individual vocal signatures from across the animal kingdom, this result was predicted. Interestingly though, the strength of the vocal signatures was not consistent across the five call types investigated. Loud call types used in the wild for distance communication (Bermejo & Omedes, 1999; Hohmann & Fruth, 1994, 1995; Mori, 1983) and during situations of high excitement or stress, contained stronger identity information than the more soft, close calls used mostly during foraging, contact or grooming. These results provide us with insight into the function and importance of certain call types and clues to what selective pressures shaped the bonobo vocal repertoire.

Within the limited number of studies that have investigated identity information in multiple call types of a species repertoire, the majority have found that, like in our results, the strength of identity information varies between call types (Bouchet et al., 2012, 2013; Charrier, et al., 2001; Leliveld et al., 2011; Lemasson & Hausberger, 2011). However, an exception to this is a single study on another ape species, the gorilla (Salmi et al., 2014). This study found no significant differences in the strength of identity information between a large number of call types, despite the presence of some minor differences in the potential for individual coding values between some of the call types (Salmi et al., 2014).

Across studies little consensus has emerged as to what evolutionary factors shaped these differences in acoustically coded identity information. Studies with two monkey species have shown that the highest levels of identity information are found in close call types used in direct social interaction (Bouchet et al., 2012; Lemasson & Hausberger, 2011). Additionally, the loudest call types, which largely function as alarm calls, show reduced levels of individual distinctiveness, arguably because the most important encoded information is predator type and urgency but not necessarily the vocaliser's identity (Bouchet et al., 2012; Lemasson & Hausberger, 2011). This pattern has also been found in a colonial living bird species, the South Polar skua, whose courtship and contact calls show higher levels of individual distinctiveness than their loud alarm calls (Charrier et al., 2001).

While these differences perhaps provide clues as to how identity information within species' repertoire evolves, due to the vast differences between habitat, social structure and morphology between these species it highly unlikely that a single rule applies. For example, for many monkey species, including Campbell's monkeys and Red-capped managbeys, the loudest call types, which show the lowest levels of individual distinctiveness, function as alarm calls; while for bonobos the loudest call types, which show the highest levels of individual distinctiveness, are primarily used for direct social function. Therefore, what is most reasonable to conclude is that identity coding is strongest in call types where it is most functional for a given species, which is in turn reflected by a species' habitat and social structure, including general group cohesiveness, the regularity that individuals need to communicate with one another over long distances, if the species is territorial, and importantly the species position in the food chain (i.e. whether they are regularly preyed upon).

What is clear from our study is that calls used for distance communication in bonobos contain the strongest identity signatures, enabling us to conclude that the successful function of these calls depends on listeners to be able to identify

the signaller. Additionally, we could conclude that either: 1) loud calls used in distance communication also have a strong social function, and individual recognition is paramount when compared to the close call types; or 2) that individual recognition is important for all call types, and that loud calls encode stronger levels of identity information to ensure its propagation over distances and through dense forest.

If we consider these two hypotheses in conjunction with the results from our playback experiments (manuscript 3) the conclusions become clearer. The playback experiments demonstrated that bonobos are capable of individual vocal recognition using the peep-yelp call type. The peep-yelp had one of the lowest individual distinctiveness in the call types explored in manuscript 1. As such, it appears that call types with statistical low levels of identity information, contain enough distinctiveness to be individually recognisable to familiar conspecifics, at least over short, unobstructed distances. This leads to the conclusion that, as individuals can utilise even low levels of identity information for recognition, the higher levels of identity information found in the loud calls is likely redundant information, encoded to ensure successful information transfer across a challenging physical environment.

In addition to the clear transfer of identity information, even in call types with statistically low individual vocal signatures, our results also demonstrate that this recognition ability is stable over long periods. The results of manuscript 3 suggested that there may be a loss of vocal recognition after six years of separation from a social partner, however due to a low sample size these results could not be conclusively confirmed. What was clear, however, was that bonobos are capable of remembering and recognising past social partners for at least six years after their last encounter. These results suggest long-term memory of past social partners is important for individual and group social interactions. Bonobo society is characterised by the fluid and fluctuating structure of fission-fusion dynamics, where individuals in a social group separate into sub-groups for

154

varying periods of time and with varying social cohesiveness. Interestingly, two other highly social, intelligent species, elephants and dolphins that have fission-fusion societies also demonstrate long-term vocal recognition. This suggests that the fluidity of the fission-fusion social structure may have a distinct, adaptive influence on the cognitive and communicative processes of these species. Vocal recognition allows for individual recognition between individuals of the same group who are physically distant enough from one another to render all other communicative modalities non-functional. This is particularly important in fission fusion groups, where small parties separate to forage for food but who often still stay within auditory range of one another and will communicate if a food source is located (Hohmann & Fruth, 1994; van Krunkelsven et al., 1996). However, the presence of long-term vocal recognition suggests that in these species, recognising past familiar individuals is also important. For bonobos in particular, this long-term vocal recognition likely helps identify: 1) past female group members who have emigrated to other groups at sexual maturity, and 2) individuals from neighbouring communities with whom an individual has had previous social interactions but with whom they do not regularly associate. Unlike chimpanzees, who are notoriously territorial and aggressive to out-group conspecifics, sometimes even fatally (Mitani et al., 2010; Wilson & Wrangham, 2003; Wilson et al., 2014), bonobos behave variably and often tolerably to individuals from outside their own community. Inter-community encounters regularly begin with vocal exchanges when individuals are still at a distance from one another, which likely provide identity information and allow for individual recognition (Furuchi, 2011; Idani, 1990). These exchanges may influence the selective social responses of each group, which range from complete inter-group avoidance, temporary aggressive or affiliative social interactions and even temporary feeding and nesting associations (Furuchi, 2011; Idani, 1990). Furuchi (2011) described that, if two communities proceed to interact after vocal exchanges, the females from both groups engage in genital rubbing or grooming as if they have a met old associates or relatives.

B. The effect of relatedness and social group on acoustic features

In the second manuscript of this thesis I investigated if two factors, relatedness and familiarity, affect the acoustic features that encode identity information in the bark call type. The results demonstrated that individuals had increasingly similar vocal features (as described by the composite measurement - Euclidean acoustic distance) if they are related and/or live in the same group. When investigating each acoustic feature separately, relatedness had a slightly stronger effect with increased vocal similarity between related dyads in seven of the twelve parameters measured, while five of the twelve parameters showed acoustic convergence between individuals living in the same group.

1. Vocal similarity between relatives

The source-filter theory describes how anatomical features of the vocal cords (source) and vocal tract (filter) shape the acoustic features of an individual's voice (Fitch & Hauser, 1998). Given that an individual's vocal morphology is shaped by his/her genes, and that genes are shared between related individuals, it is not surprising that related individuals show greater similarity in the acoustic features of their calls – independently of sex and age – than unrelated individuals. However, it is unclear what, if any, function this has in bonobo communication and social interaction.

One possible function is that vocal similarity between related individuals aids in kin recognition. Kin recognition is a necessary mechanism in kin selection, by which individuals selectively bias their social behaviour toward related conspecifics, either by acting altruistically toward relatives or to inhibit inbreeding (Silk, 2002; Widdig, 2007). The majority of studies on kin recognition in non-human primates have focused on phenotype matching, when an individual can identify kin based on features similar to his/her own, in the olfactory and visual

modalities (Langergraber, 2012). A few studies have focused on the vocal modality and have found that some primate species demonstrate significant vocal similarity between related individuals (kin signatures) (e.g. pigtail macaques – Gouzoules & Gouzoules, 1990; mouse lemurs - Kessler et al., 2012, 2014). A limited number of studies with macaques (Pfefferle et al., 2014, 2015; Rendall et al., 1996), grey mouse lemurs (Kessler et al., 2012), and mandrills (Levréro et al., 2015) have experimentally tested, and shown, that individuals are capable of using vocalisations to recognise kin from non-kin. Notably, however, studies of kin selection and vocal recognition in non-human primates have almost entirely focused on species whose social structures are centred around matrilineal networks of related individuals and in which males disperse at sexual maturity (Gouzoules & Gouzoules, 1987; Melnick & Pearl, 1987). Evidence from macaques, as well as baboons and vervet monkeys, have demonstrated that species with these social structures spend more time with kin than non-kin and display nepotism during grooming, coalition support, and reconciliation (Silk, 2002), making recognising kin important in social functioning.

For male-philopatric species like bonobos and chimpanzees, whose social structure is dependent upon cooperation, coordination and coalition between kin and non-kin, there is considerably less evidence of kin selection in social behaviour, excluding between mother and offspring (Goldberg & Wrangham, 1997; Hohmann, et al., 1999; Langergraber et al., 2007). This is particularly true for bonobos, whose exceptional relationships between non-related females violate the kin selection model of social relationships (Tokuyama & Furuichi, 2016). In such social environments, the importance of kin recognition within a community is debatable.

Based on the results of manuscript 3, vocal recognition could serve to mediate inter-community encounters and allow for selective and differential responses to out-group conspecifics based on individual recognition and social knowledge. It is possible that in addition to this, kin selection and recognition are functioning

during inter-community vocal exchanges, allowing individuals to selectively react to other communities based on the number of kin recognised in other group. Additionally, as suggested in manuscript 2, it is possible that kin recognition is utilised by females when choosing a new community after migration from their natal group. As female natal dispersal is generally poorly understood in bonobos, this is an intriguing hypothesis. However, as with the other hypotheses, further research is necessary to test if the vocal similarity found between related individuals is utilised in any functional way or if it is simply a vestigial artefact of similar morphology due to genetics.

2. Vocal similarity between group members

In manuscript 2, I found distinct evidence of vocal similarity between individuals living in the same group. In the analysis, multiple factors were controlled for, importantly including relatedness, ensuring that the observed vocal convergence was an effect of familiarity between current group members.

Familiarity between group members is a common mechanism that enables individual identification (Bradbury & Vehrencamp, 2011). However, a number of studies have argued that the vocal dialects (or group signatures, which are created when vocal convergence occurs between group members) found in many species foster recognition and aid in distinguishing group members from non-members (Tyack, 2007). Group signatures utilised in this way are likely essential in very large groups of individuals, where knowing the identity of each individual would pose an extreme cognitive challenge. Studies with some bat and bird species have suggested that in large grouped species group dialects serve as a 'password' to allow access to shared resources (Wilkinson and Boughman; 1998) or to distinguish one's own group from a distance (Boughman 1997; Brown & Farabaugh, 1997; Catchpole & Slater, 2008).

For non-human primates, vocal convergence has been hypothesised to create these vocal dialects, which may promote group cohesion as well as recognition (Henry et al, 2015). However, arguments have been made that vocal convergence also has a similar function to that in humans, as described by speech accommodation theory, to reinforce bonding between established group members and to facilitate social integration of new group members (Giles, 1991; Hammerschmidt & Fischer, 2008; Locke & Snow, 1997). Support for both hypotheses have been found in a number of primate species. Barbary macaques (Fischer et al., 1998), Japanese macaques (Tanaka et al., 2006), and pygmy marmosets (de la Torre & Snowden, 2009) have demonstrated consistent acoustic differences between distinct social groups, and Diana monkeys (Candiott et al., 2012), pygmy marmosets (Snowden & Elowson, 1999), Campbell's monkeys (Lemasson et al., 2011) and Japanese macaques (Suguira, 1998) have demonstrated active vocal convergence during chorusing or with closely affiliated group members.

Evidence from chimpanzees supports the argument that both hypotheses are functioning concurrently. Chimpanzees have shown that the acoustic figures of their pant-hoot call type converge during chorusing (Mitani & Brandt 1994; Mitani & Gros-Louis, 1998) and chorusing is significantly more likely to occur between individuals with strong affiliative social bonds (Fedurek et al., 2014). Together these evidences solidly suggest that choruses play a crucial role maintaining strong social bonds between chimpanzee males. On the other hand, strong evidence that vocal dialects function in group recognition comes from Crockford et al. (2004). Their study found evidence of vocal dialects in three wild chimpanzee groups, two that were neighbouring and a third that was distant to the other two. While all three groups showed vocal convergence within the groups, the two neighbouring groups showed notable acoustic divergence from one another that neither demonstrated with the third, distant group. These results show that as the acoustic call features were converging within a group they were also diverging between groups, suggesting that it may be just as important to

sound similar to group members as it is dissimilar to non-group conspecifics, which would enhance individual recognition by vocal dialect.

The evidence of vocal convergence presented in manuscript 2 allows us to compellingly hypothesise that vocal convergence in bonobos, like in their closest relative the chimpanzee, functions dually for both social bonding and in creating vocal dialects which can aid in distinguishing group members from non-group individuals. However, further studies are necessary to explicitly test this hypothesis, which will be discussed in the next section (in Future directions).

The historical lack of evidence for vocal learning in non-human primates has been considered a paradox in the evolutionary models of human language (Egnor & Hauser, 2004). However, in the last decade a limited but consistent body of evidence has accumulated suggesting that vocal production learning is indeed occurring in non-human primates. The evidence of vocal convergence between non-related group members presented in this thesis provides further support for vocal production learning in a primate species. While the scope and functionality of vocal learning may be limited when compared to some bird and cetacean species, it is increasingly clear that vocal learning is an integral aspect of non-human primate vocal communication and should be more broadly recognised as such.

C. Future Directions

1. Testing the limits of vocal recognition

The results of manuscript 3 strongly indicated that after long periods of separation bonobos may no longer be able to recognise past familiar individuals by vocalisations alone. However, due to low sample size we were unable to conclusively confirm this. As such, further playback experiments would be necessary to establish if there is indeed an upper limit to long-term vocal recognition.

Secondly, we demonstrated that vocal recognition is possible even with a call type that shows low levels of identity information. However, we did not test vocal recognition with the call type that had the *lowest* level of individually, the peep, or the highest, the high-hoot. It would be interesting to test vocal recognition in multiple call types as it would: 1) give us a clearer indication as to the function of each call type and 2) allow us to assess the reliability of statistical models measuring the coding of identity information.

2. Vocal recognition and identity information in specific acoustic features

Studies with species such as seals (Charrier et al., 2003), penguins (Jouventin et al., 1999) and zebra finches (Vignal et al., 2008) have demonstrated that individual recognition is still possible after artificial distortion of certain acoustic features but not others. The results of Manuscript 1 demonstrate that, in bonobo vocalisations, certain acoustic parameters contain higher potential for individuality than others. By experimentally manipulating these different acoustic features and performing playback experiments on individual recognition we could test: 1) the limits of vocal recognition, i.e. how many acoustic parameters carrying identity information can be modified before recognition is no longer

possible and 2) which specific acoustic features are particularly vital for recognition. Again, these experiments would also allow us to test if the statistical measures of identity information reflect if and how bonobos utilise this information in real world scenarios.

3. Further testing on vocal convergence

a. Is vocal convergence creating group dialects?

This thesis tested and found vocal convergence between bonobo group members, which strongly suggest the presence of group vocal signatures. However, I was unable to explicitly test for group signatures, as the studied zoo populations comprised varying and unbalanced numbers of related individuals. Any analysis would have therefore been skewed, as we have shown that related individuals are acoustically more similar than unrelated individuals, and attempts to control for the imbalance were unsuccessful. Within bonobos it has yet to be precisely determined whether groups demonstrate vocal dialects, and as such directly testing this would be valuable.

b. What is the function of vocal convergence – social bonding, recognition or both?

Social bonding hypothesis
There has been much evidence from within non-human primates that vocal convergence serves to facilitate social bonding between group members. To test this in bonobos, we could investigate the levels of vocal similarity between closely affiliated individuals within a group compared to individuals who have poor or neutral social relationships. If vocal convergence facilitates social bonding, we would expect to see higher levels of vocal similarity between closely affiliated individuals. Another way to test the social bonding hypothesis would be to investigate the temporary convergence of acoustic features between

individuals during vocal chorusing, as has been found in chimpanzees. The advantage of this latter method is that it allows for direct comparison with multiple other studies that have used it to investigate vocal convergence and social bonding.

A final and intriguing way to approach the social bonding hypothesis would be to investigate multiple call types and compare the levels of vocal convergence to the function and use of each call type. For this thesis, I assessed vocal convergence only in the bark call type, which is used for both long and short distance communication but largely in times of high excitement or stress. Vocal convergence in these contexts could serve to decrease tension between strongly bonded individuals. On the other hand, vocal convergence in call types that are specifically used in close contact interactions, such as grooming, could function to maintain established relationships or help build new relationships between group members.

Recognition hypothesis

Our results demonstrating robust individual vocal recognition in bonobos, even after years of separation (manuscript 3), strongly suggest that recognition of group members occurs on an individualistic basis. This calls into question whether vocal dialects in bonobos function in identifying group members from non-group members. However, that being said, it is possible that group dialects serve to maintain cohesion with a group, which could be particularly important in bonobo intra-group communication when a community fuses into smaller foraging parties during the day and are separated by distance and dense forest. This hypothesis could be tested by comparing the strength of group identity information versus individual identity information encoded in call types used for distance communication. If group signatures were stronger than individual signatures, suggesting more successfully propagation through a difficult acoustic environment, one could theorise that they function for group cohesion.

Dual function hypothesis

It is important to highlight that the aforementioned hypotheses are not mutually exclusive. Positive results for the experimentation and analyses described in the above two sub-sections could potentially support a third hypothesis. Namely, that vocal convergence is functioning for both recognition and social bonding in bonobo society.

D. Conclusion

Overall the results presented in this thesis have provided a number of new insights into the form and function of identity information in the bonobo vocal repertoire. From the signaller perspective, we have seen that identity information is coded in all of the most common bonobo call types, but that calls used for distance communication contain the strongest individual vocal signatures. We have also found that social life can influence the call structures that encode identity information, as the vocal features of familiar individuals converge within a bonobo community. Finally, we found that bonobos can use this identity information to recognise familiar individuals, even after years of separation from one another.

Despite the challenges, bonobos have been a fascinating species to work with and I am thankful to have been fortunate enough to study them for the last four years. I hope that this thesis helps to progress our understanding of bonobo communication and society and inspires other researchers to study this remarkable species.

REFERENCES

V. REFERENCES

Aubin, T., & Jouventin, P. (2002). How to vocally identify kin in a crowd: the penguin model. *Advances in the Study of Behavior, 31*, 243–277.

Baker, M.C., & Becker, A.M. (2002). Mobbing calls of black–capped chickadees: Effects of urgency on call production. *The Wilson Bulletin, 114,* 510–516.

Bass, A.H., & Clark, C.W. (2003). The physical acoustics of underwater sound communication. In A. M. Simmons, A.N. Popper & R.R. Fay (Eds.), *Acoustic Communication* (pp. 15–64). New York, New York: Springer–Verlag.

Bauers, K. A., (1993). A functional analysis of staccato grunt vocalizations in the stumptailed macaque (*Macaca arctoides*). *Ethology, 94,* 147–161.

Bee, M. A., Kozich, C. E., Blackwell, K. J., & Gerhardt, H. C. (2001). Individual variation in advertisement calls of territorial male Green frogs, *Rana clamitans*: Implications for individual discrimination. *Ethology*, *107*, 65–84.

Beecher, M. D. (1989). Signalling system for individual recognition: an information theory approach. *Animal Behaviour*, *38*, 248–261.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology and Evolution*, *20*, 143–149.

Behringer, V., Deschner, T., Deimel, C., Stevens, J. M. G., & Hohmann, G. (2014). Age–related changes in urinary testosterone levels suggest differences in puberty onset and divergent life history strategies in bonobos and chimpanzees. *Hormones and Behavior*, *66*, 525–533.

Bergman, T. J., Beehner, J. C., Cheng, L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, *302,* 1234–1236.

Bergman, T. J. (2010). Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proceedings of the Royal Society B – Biological Sciences*, 277, 3045–3053.

Bergman, T. J., & Sheehan, M. J. (2013). Social knowledge and signals in primates. *American Journal of Primatology*, 75, 683–694.

Bermejo, M., Illera, G., & Pí, J. S. (1995). Animals and mushrooms consumed by bonobos (*Pan paniscus*): New records from Lilungu (Ikela), Zaire. *International Journal of Primatology*, *16*, 879–898.

Bermejo, M., & Omedes, A. (1999). Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, *70*, 328–357.

Blumstein, D. T., Nguyen, K. T., & Martin, J. G. A. (2013). Ontogenetic variation of heritability and maternal effects in yellow–bellied marmot alarm calls. *Proceedings of the Royal Society B – Biological Sciences*, *280*, 20130176.

Boeckle, M., & Bugnyar, T. (2012). Long-term memory for affiliates in ravens. *Current Biology*, *22*, 801–6.

Boesch, C. (2002). Behavioural diversity in *Pan*. In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp.1–8). Cambridge, UK: Cambridge University Press.

Boesch, C., Hohmann, G. & Marchant, L. (2002). *Behavioural Diversity in Chimpanzees and Bonobos*. Cambridge, UK: Cambridge University Press.

Bouchet, H., Blois–Heulin, C., Pellier, A. S., Zuberbühler, K., & Lemasson, A. (2012). Acoustic variability and individual distinctiveness in the vocal repertoire of red–capped mangabeys (*Cercocebus torquatus*). *Journal of Comparative Psychology*, *126,* 45–56.

Bouchet, H., Blois–Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non–human primate species. *Frontiers in Psychology, 4,* 1–15.

Boughman, J.W. (1997). Greater spear–nosed bats give group–distinctive calls. *Behavioral Ecology and Sociobiology*, *40*, 61–70.

Boughman, J.W. (1998). Vocal learning by greater spear–nosed bats. *Proceedings of the Royal Society B – Biological Sciences*, 265, 227–233.

Bradbury, J.W., & Vehrencamp, S.L. (2011). *Principles of Animal Communication,* 2nd Edition. Sunderland, Massachusetts: Sinauer Associates.

Bradbury, J. W., & Balsby, T. J. S. (2016). The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology*, *70,* 293–312.

Briefer, E., Rybak, F., & Aubin, T. (2010). Are unfamiliar neighbours considered to be dear–enemies? *PLoS ONE*, *5*, 1–7.

Briefer, E. F., Vannoni, E., & McElligott, A. G. (2010). Quality prevails over identity in the sexually selected vocalisations of an ageing mammal. *BMC Biology*, *8,* 35.

Briefer, E. F., Padilla de la Torre, M., & McElligott, A. G. (2012). Mother goats do not forget their kids' calls. *Proceedings of the Royal Society B – Biological Sciences*, *279*, 3749–3755.

Briseño–Jaramillo, M., Estrada, A., & Lemasson, A. (2014). Individual voice recognition and an auditory map of neighbours in free–ranging black howler monkeys (*Alouatta pigra*). *Behavioral Ecology and Sociobiology*, *69*, 13–25.

Brown, C. H., Gomez, R., & Waser, P. M. (1995). Old-world monkey vocalizations: adaption to the local habitat. *Animal Behaviour*, *50*, 945–961.

Brown, E.D., & Farabaugh, S.M. (1997). What birds with complex social relationships can tell us about vocal learning: vocal sharing in avian groups. In C.T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 98–127). Cambridge, UK: Cambridge University Press.

Bruck, J. N. (2013). Decades-long social memory in bottlenose dolphins. *Proceedings of the Royal Society B – Biological Sciences*, *280*, 20131726.

Bugnyar, T., Kijne, M., & Kotrschal, K. (2001). Food calling in ravens: Are yells referential signals? *Animal Behaviour*, *61*, 949–958.

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, *8*, 382–5.

Catchpole, C., & Slater, P. (2008). *Bird Song: Biological Themes and Variations*. Cambridge, UK: Cambridge University Press.

Chapman, C. A., & Weary, D. M. (1990). Variability in spider monkeys' vocalizations may provide basis for individual recognition. *American Journal of Primatology*, *22,* 279–284.

Charlton, B.D., Reby, D., Ellis, W.A.H., Brumm, J., & Fitch, W.T. (2012) Estimating the active space of male koala bellows: Propagation of cues to size and identity in a eucalyptus forest. *PLoS ONE*, *7*, e45420.

Charrier, I., Jouventin, P., Mathevon, N., & Aubin, T. (2001). Individual identity coding depends on call type in the South Polar skua (*Catharacta maccormicki*). *Polar Biology*, *24*, 378–382.

Charrier, I., Mathevon, N., & Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Animal Behaviour*, *65*, 543–550.

Charrier, I., Aubin, T., & Mathevon, N. (2010). Mother–calf vocal communication in Atlantic walrus: a first field experimental study. *Animal Cognition*, *13*, 471–482.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Animal Behaviour*, *50*, 249–257.

Clay, Z., & Zuberbühler, K. (2009). Food–associated calling sequences in bonobos. *Animal Behaviour*, *77*, 1387–1396.

Clay, Z., & Zuberbühler, K. (2011a). Bonobos extract meaning from call sequences. *PloS One*, *6*, e18786.

Clay, Z., & Zuberbühler, K. (2011b). The Structure of Bonobo Copulation Calls During Reproductive and Non-Reproductive Sex. *Ethology*, *117*, 1158–1169.

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, *3*: e1124.

Coolidge, H.J. (1933). Pan paniscus, pigmy chimpanzee from south of the Congo River. *American Journal of Physical Anthropology*, *18*, 254–264.

Coolidge, H.J., & Shea, B.T. (1982). External body dimensions of *Pan paniscus* and *Pan troglodytes* chimpanzees. *Primates, 23*, 245–251.

Cowlishaw, G. (1996). Sexual selection and information content in gibbon song bouts. *Ethology*, *102*, 272–284.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls : a case for vocal learning? *Ethology*, *243*, 221–243.

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., & Suki, B. (2002). Only male fin whales sing loud songs. *Nature*, *417,* 809-810.

Dall, S. R. X., Giraldeau, L.A., Olsson, O., McNamara, J. M., & Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology & Evolution*, *20,* 187–93.

Deecke, V.B., Ford, J.K.B., & Spong, P. (2000). Dialect change in resident killer whales (*Orcinus orca*): Implications for vocal learning and cultural transmission. *Animal Behaviour*, 60, 629–638.

Deecke, V. B., Barrett–Lennard, L. G., Spong, P., & Ford, J. K. B. (2010). The structure of stereotyped calls reflects kinship and social affiliation in resident killer whales (*Orcinus orca*). *Naturwissenschaften*, *97,* 513–518.

De Lathouwers, M., & Van Elsacker, L. (2006). Comparing infant and juvenile behavior in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): a preliminary study. *Primates*, *47*, 287–293.

De La Torre, S., & Snowdon, C. T. (2009). Dialects in pygmy marmosets? population variation in call structure. *American Journal of Primatology*, *71*, 333–342.

de Waal, F. B. M. (1988). The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour*, *27*, 183–251.

Dentressangle, F., Aubin, T., & Mathevon, N. (2012). Males use time whereas females prefer harmony: Individual call recognition in the dimorphic blue–footed booby. *Animal Behaviour*, *84*, 413–420.

Dornhaus, A., & Chittka, L. (2004). Why do honey bees dance? *Behavioral Ecology and Sociobiology*, *55*, 395–401.

Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*, 649–54.

Elowson, A. M., & Snowdon, C. T. (1994). Pygmy marmosets, Cebuella pygmaea, modify vocal structure in response to changed social environment. *Animal Behaviour*, 47, 1267–1277.

Evans, C. S., Evans, L., & Marler, P. (1993). On the meaning of alarm calls: functional reference in an avian vocal system. *Animal Behaviour*, *46*, 23-38.

Evans, C. S., & Evans, L. (1999). Chicken food calls are functionally referential. *Animal Behaviour*, *58*, 307–319.

Evans, C. S., & Evans, L. (2007). Representational signalling in birds. *Biology Letters*, *3,* 8–11.

Ey, E., Pfefferle, D., & Fischer, J. (2007). Do age– and sex–related variations reliably reflect body size in non–human primate vocalizations? A review. *Primates*, *48,* 253–267.

Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour*, *86,* 189–196.

Field, A. (2009). *Discover Statistics Using SPSS (3rd Edition).* London, Thousand Oaks, New Delhi, Singapore: Sage Publications Ltd.

Fischer, J., Hammerschmidt, K., & Todt, D. (1998). Local variation in Barbary macaque shrill barks. *Animal Behaviour*, *56*, 623–629.

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

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

Fitch, W. T., & Hauser, M. D. (1995). Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on " Honest " advertisement. *American Journal of Primatology*, *37*, 191–219.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*, 1213–1222.

Fitch, W. T., & Hauser, M. D. (1998). Unpacking " Honesty ": Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, A.N. Popper & R.R. Fay (Eds.), *Acoustic Communication* (pp. 15–64). New York, New York: Springer–Verlag.

Fitch, W.T. (2010). *The evolution of language*. Cambridge, U.K.: Cambridge University Press.

Foote, A. D., Griffin, R.M., Howitt, D., Larsson, L., Miller, P.J.O., & Hoelzel, A.R. (2006). Killer whales are capable of vocal learning. *Biology Letters*, 2, 509–512.

Forrest, T. G. (1994). From sender to receiver: Propagation and environmental effects on acoustic signals. *American Zoologist*, *34*, 644–654.

Fruth, B. & Hohmann, G. (2002). How bonobos handle hunts and harvests: why share food? In C. Boesch, G. Hohmann & L. Marchant (Eds.) *Behavioural Diversity in Chimpanzees and Bonobos* (pp.231–243). Cambridge, UK: Cambridge University Press.

Fruth, B, & Hohmann, G. (2006). Social grease for females? Same–sex genital contacts in wild bonobos. In V. Sommer & P. Vasey (Eds.), *Homosexual behaviour in animals: An evolutionary perspective* (pp. 294–315). Cambridge, UK: Cambridge University Press.

Furuichi, T. (1989). Social interactions and the life history of female *Pan paniscus* in Wamba, Zaire. *International Journal of Primatology*, *10*, 173–197.

Furuichi, T., & Ihobe, H. (1994). Variation in male relationships in bonobos and chimpanzees. *Behaviour*, *130*, 211–228.

Furuichi, T. (1997). Agonistic interactions and matrifocal dominance rank of wild bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, *18,* 855–875.

Furuichi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthropology*, *20*, 131–142.

Garstang, M., Larom, D., Raspet, R. & Lindeque, M. (1995). Atmospheric controls on elephant communication. *Journal of Experimental Biology*, *198,* 939–951.

Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi–modal use of a socially directed call in bonobos. *PloS One*, *9*, 1–12.

Genty, E., & Zuberbühler, K. (2014). Spatial reference in a bonobo gesture. *Current Biology, 24*, 1–5.

Genty, E., Neumann, C., & Zuberbühler, K. (2015). Complex patterns of signalling to convey different social goals of sex in bonobos, *Pan paniscus*. *Scientific Reports*, *5*, 16135.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology*, *18*, 457–460.

Giles, H., Coupland, N., & Coupland, J. (1991). Accomodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accomodation: Developments in Applied Sociolinguistics* (pp. 1-68). Cambridge, U.K.; Cambridge University Press.

Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird. *Nature, 350,* 228–229.

Goldberg, T. I., & Wrangham, R. W. (1997). Genetic correlates of social behaviour in wild chimpanzees: evidence from mitochondrial DNA. *Animal Behaviour*, *54*, 559–570.

Gouzoules, S., & Gouzoules, H. (1987). Kinship. In B.B. Smuts, D.L. Cheney, R. M. Seyfarth, R.W. Wrangham & T.T. Struhsaker (Eds.) *Primate Societies* (pp. 299-305). Chicago: University of Chicago Press.

Gouzoules, H., & Gouzoules, S. (1990). Matrilineal Signatures in the Recruitment Screams of Pigtail Macaques , Macaca nemestrina. *Behaviour*, *115*(3/4), 327–347.

Griesser, M. (2008). Referential calls signal predator behavior in a group–living bird species. *Current Biology*, *18*, 69–73.

Grinnell, J., & McComb, K. (2001). Roaring and social communication in African lions: the limitations imposed by listeners. *Animal Behaviour*, *62*, 93–98.

Halina, M., Rossano, F., & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal Cognition*, *16*, 653–666.

Hammerschmidt, K., Freudenstein, T. & Jürgens U. (2001) Vocal development in squirrel monkeys. *Behaviour*, *138*, 1179–1120.

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In Oller, D.K. & U. Griebel (Eds.) *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication* (pp. 367-395). Cambridge, Mass.: MIT Press.

Hashimoto, C. (1997). Context and Development of Sexual Behavior of Wild Bonobos (*Pan paniscus*) at Wamba, Zaire. *International Journal of Primatology*, *18*, 1–21.

Hashimoto, C., Tashiro, Y., Kimura, D., Enomoto, T., Ingmanson, E. J., Idani, G., & Furuichi, T. (1998). Habitat use and ranging of wild bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, *19*, 1045–1060.

Hayes, C. (1951). The ape in our house. New York: Harper.

Henry, L., Barbu, S., Lemasson, A., & Hausberger, M. (2015). Dialects in animals : Evidence , development and potential functions. *Animal Behavior and Cognition*, *2,* 132–155.

Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology*, *24*, 1–5.

Hohmann, G., & Fruth, B. (1994). Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology*, *15*, 767–782.

Hohmann, G., & Fruth, B. (1995). Loud calls in great apes: sex differences and social correlates. *Current Topics in Primate Vocal Communication*, 161–184.

Hohmann, G., Gerloff, U., & Tautz, D. (1999). Social bonds and genetic ties: kinship, association and affiliation in a community of bonobos (*Pan paniscus*). *Behaviour*, *136*, 1219–1235.

Hohmann, G., & Fruth, B. (2000). Use and function of genital contacts among female bonobos. *Animal Behaviour*, *60,* 107–120.

Hohmann, G., & Fruth, B. (2002). Dynamics in social organization of bonobos (*Pan paniscus*). In C. Boesch, G. Hohmann & L. Marchant (Eds.). *Behavioural Diversity in Chimpanzees and Bonobos* (pp.61–70). Cambridge, UK: Cambridge University Press.

Hohmann, G., & Fruth, B. (2003). Intra– and inter–sexual aggression by bonobos in the context of mating. *Behaviour, 140*, 1389–1413.

Hohmann, G., & Fruth, B. (2008). New records on prey capture and meat eating by bonobos at Lui Kotale, Salonga National Park, Democratic Republic of Congo. *Folia Primatologica*, *79*, 103–110.

Idani, G. (1990). Relations between unit–groups of bonobos at Wamba, Zaire: Encounters and temporary fusions. *African Study Monographs*, *11*, 153–186.

Idani, G. (1991). Social relationships between immigrant and resident bonobo (*Pan paniscus*) females at Wamba. *Folia Primatologica, 57*, 83–95.

Idani, G., Kuroda, S., Kano, T., & Asato, R. (1994). Flora and vegetation of Wamba forest, Central Zaire with reference to bonobo (*Pan paniscus*) foods. *Tropics*, *3*, 309–332.

Idani, G., Mwanza, N., Ihobe, H., Hashimoto, C., Tashiro, Y. & Furuichi, T. (2008) Changes in the status of bonobos, their habitat, and the situation of humans at Wamba in the Luo Scientific Reserve, Democratic Republic of Congo. In: Furuichi, T. & Thompson, J. (Eds.) *The bonobos: behavior, ecology, and conservation*. (pp 121–134). New York: Springer.

Insley, S. J. (2000). Long–term vocal recognition in the northern fur seal. *Nature*, *406*, 404–5.

Janicke, T., Hahn, S., Ritz, M. S., & Peter, H. U. (2008). Vocal performance reflects individual quality in a nonpasserine. *Animal Behaviour*, *75*, 91–98.

Janik, V.M., & Slater, P.J.B. (1997). Vocal learning in mammals. *Advances in the Study Behavior*, 26, 59–99.

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour, 60*, 1–11.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28,* 60–65.

Jones, G., & Holderied, M. W. (2007). Bat echolocation calls: adaptation and convergent evolution. *Proceedings of the Royal Society B – Biological Sciences*, *274*, 905–912.

Jouventin, P., Aubin, T., & Lengagne, T. (1999). Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Animal Behaviour*, *57*, 1175–1183.

Jungers, W.L., & Susman, R.L. (1984). Body size and skeletal allometry in African apes. In: R.L. Susman (Ed.) *The Pygmy Chimpanzee*, (pp. 131–178). New York, Plenum Press.

Kano, T. (1992). *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford, California: Stanford University Press.

Kellogg, W.N. & Kellogg, L.A. (1933). *The Ape and the Child*. New York, New York: McGraw–Hill.

Kessler, S. E., Scheumann, M., Nash, L. T., & Zimmermann, E. (2012). Paternal kin recognition in the high frequency / ultrasonic range in a solitary foraging mammal. BMC Ecology, 12, 26.

Kessler, S. E., Radespiel, U., Hasiniaina, A. I. F., Leliveld, L. M. C., Nash, L. T., & Zimmermann, E. (2014). Modeling the origins of mammalian sociality: moderate evidence for matrilineal signatures in mouse lemur vocalizations. *Frontiers in Zoology*, *11*, 14.

King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W., & Janik, V. M. (2013). Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proceedings of the Royal Society B – Biological Sciences, 280*, 20130053.

Kiriazis, J., & Slobodchikoff, C. N. (2006). Perceptual specificity in the alarm calls of Gunnison's prairie dogs. *Behavioural Processes*, *73*, 29–35.

Kitzmann, C. D., & Caine, N. G. (2009). Marmoset (*Callithrix geoffroyi*) food–associated calls are functionally referential. *Ethology*, *115*, 439–448.

Knornschild, M., & von Helversen, O. (2008). Nonmutual vocal mother–pup recognition in the greater sac–winged bat. *Animal Behaviour, 76*, 1001–1009.

Kondo, N., Izawa, E., & Watanabe, S. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): Contact call signature and discrimination. *Behaviour*, *147,* 1051–1072.

Kuroda S. 1989. Developmental retardation and behavioral characteristics of pygmy chimpanzees. In P.G. Heltne & L.A Marquardt (Eds.), *Understanding Chimpanzees* (pp. 184–93). Cambridge, Mass.: Harvard University Press.

Langergraber, K. E., Mitani, J. C., & Vigilant, L. (2007). The limited impact of kinship on cooperation in wild chimpanzees. *Proceedings of the National Academy of Sciences*, *104*, 7786–7790.

Langergraber, K.E. (2012) Cooperation among kin. In J.C. Mitani, J. Call, P.M. Kappeler, R.A. Palombit, J.B. Silk (Eds.), *The evolution of primate societies*. (pp 491–551). Chicago: University of Chicago Press.

Larom, D., Garstang, M., Payne, K., Raspet, R., & Lindeque, M. (1997). The influence of surface atmospheric conditions on the range and area reached by animal vocalizations. *The Journal of Experimental Biology*, *200*, 421–431.

Leliveld, L. M. C., Scheumann, M., & Zimmermann, E. (2011). Acoustic correlates of individuality in the vocal repertoire of a nocturnal primate (*Microcebus murinus*). *The Journal of the Acoustical Society of America*, *129*(4), 2278–2288.

Lemasson, A., Boutin, A., Boivin, S., Blois–Heulin, C., & Hausberger, M. (2009). Horse (*Equus caballus*) whinnies: a source of social information. *Animal Cognition*, 12, 693–704.

Lemasson, A., & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Journal of the Acoustical Society of America*, *129*, 3341–3352.

Lemasson, A., Ouattara, K., Petit, E. J., & Zuberbühler, K. (2011). Social learning of vocal structure in a nonhuman primate? *BMC Evolutionary Biology*, *11*, 362.

Levréro, F., & Mathevon, N. (2013). Vocal signature in wild infant chimpanzees. *American Journal of Primatology*, *75*, 324–332.

Levréro, F., Carrete–Vega, G., Herbert, A., Lawabi, I., Courtiol, A., Willaume, E., Kappler, P.E. & Charpentier, M. J. E. (2015). Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communication*, *6*, 7609.

Locke, J.L., & Snow, C. (1997). Vocal learning in humans and nonhuman primates. In C.T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 234–248). Cambridge, UK: Cambridge University Press.

Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B – Biological Sciences, 268,* 2315–2324.

Manser, M. B., Bell, M. B., & Fletcher, L. B. (2001). The information that receivers extract from alarm calls in suricates. *Proceedings of the Royal Society B – Biological Sciences*, *268*, 2485–2491.

Mateo, J. M. (2004). Recognition systems and biological organization: The perception component of social recognition. *Annales Zoologici Fennici*, *41,* 729–745.

Mathevon, N., Charrier, I., & Jouventin, P. (2003). Potential for individual recognition in acoustic signals: a comparative study of two gulls with different nesting patterns. *Comptes Rendus Biologies*, *326*, 329–337.

Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E., & Theunissen, F. E. (2010). What the hyena's laugh tells: sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecology*, *10*, 9.

Matthews, S., & Snowdon, C. T. (2011). Long–term memory for calls of relatives in cotton–top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 125,* 366–369.

Matrosova, V. A., Blumstein, D. T., Volodin, I. A., & Volodina, E. V. (2011). The potential to encode sex, age, and individual identity in the alarm calls of three species of Marmotinae. *Naturwissenschaften*, *98*, 181–192.

Marten, K., & Marler, P. (1977). Sound transmission and its significance for animal vocalization: I. Temperate habitats. *Behavioral Ecology and Sociobiology, 2*, 271-290.

Maynard Smith, J., & Harper, D. (2003). *Animal Signals.* Oxford, U.K.: Oxford University Press.

McComb, K., Moss, C., Sayialel, S., & Baker, L. (2000). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour, 59*, 1103–1109.

McComb, K., Reby, D., Baker, L., Moss, C., & Sayialel, S. (2003). Long–distance communication of acoustic cues to social identity in African elephants. *Animal Behaviour*, *65*, 317–329.

Melnick, D.J., & Pearl, M.C. (1987). Cercopithecines in multimale groups: Genetic diversity and population structure. In B.B. Smuts, D.L. Cheney, R. M. Seyfarth, R.W. Wrangham & T.T. Struhsaker (Eds.), *Primate Societies* (pp. 121–134). Chicago: University of Chicago Press.

Mercado, E., & Frazer, L. N. (1999). Environmental constraints on sound transmission by humpback whales. *The Journal of the Acoustical Society of America*, *106,* 3004–3016.

Mitani, J. C. (1990). Experimental field studies of Asian ape social–systems. *International Journal of Primatology*, *11*, 103–126.

Mitani, J. C., & Brandt, K. L. (1994). Social factors influence the acoustic variability in the long–distance calls of male chimpanzees. *Ethology*, *96*, 233–252.

Mitani, J., Gros–Louis, J., & Macedonia, J. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology*, *17*, 569–583.

Mitani, J. C., & Gros–Louis, J. (1998). Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour*, *135*, 1041–1064.

Mitani, J. C., & Stuht, J. (1998). The evolution of nonhuman primate loud calls : Acoustic adaptation for long–distance transmission. *Primates*, *39,* 171–182.

Mitani, J. C., Watts, D. P., & Amsler, S. J. (2010). Lethal intergroup aggression leads to territorial expansion in wild chimpanzees. *Current Biology*, *20*, 507–508.

Mori, A. (1983). Comparison of the Communicative Vocalizations and Behaviors of Group Ranging in Eastern Gorillas , Chimpanzees and Pygmy Chimpanzees *. *Primates*, *24*(October), 486–500.

Morton, E. S. (1975). The American Society of Naturalists ecological sources of selection on avian sounds. *The American Naturalist*, *109*, 17–34.

Mouterde, S. C., Theunissen, F. E., Elie, J. E., Vignal, C., & Mathevon, N. (2014). Acoustic communication and sound degradation: How do the individual signatures of male and female zebra finch calls transmit over distance? *PLoS ONE*, *9*, e102842.

Mulavwa, M., Furuichi, T., Yangozene, K., Yamba–Yamba, M., Motema–Salo, B., Idani, G., Ihobe, H., Hashimoto, C., Tashiro, Y., & Mwanza, N. (2008) Seasonal changes in fruit production and party size of bonobos at Wamba. In Furuichi, T., & Thompson, J. (Eds.) *The bonobos: behavior, ecology, and conservation* (pp 121–134). New York: Springer.

Mundry, R., & Sommer, C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour*, *74*(4), 965–976.

Myers–Thompson, J. A. (2002). Bonobos of the Luruku Wildlife Research Project. In C. Boesch, G. Hohmann, & L. Marchant (Eds.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp.61–70). Cambridge, UK: Cambridge University Press.

Neuweiler, G., (1983). Echolocation and adaptivity to ecological constraints. In F. Huber, & H. Markl (Eds.), *Neuroethology and Behavioral Physiology* (pp. 280–302). Berlin: Springer.

Newman, J. D. & Symmes, D. (1982). Inheritance and experience in the acquisition of primate acoustic behavior. In C. T. Snowdon, C. H. Brown, & M. R. Petersen (Eds.), *Primate Communication* (pp. 259–278). Cambridge: Cambridge University Press.

Newton-Fisher, N.E. (2002). Relationships of male chimpanzee in the Budongo Forest, Uganda. In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp.125–137). Cambridge, UK: Cambridge University Press.

Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1992). "Food" calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally–raised offspring, and offspring cross–fostered between species. *Behaviour, 120*, 218–231.

Owren, M. J., Seyfarth, R. M., & Cheney, D. L. (1997). The acoustic features of vowel–like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions. *The Journal of the Acoustical Society of America*, *101*, 2951–2963.

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, *10*, 58–71.

Owren, M. J., Rendall, D., & Ryan, M. J. (2010). Redefining animal signaling: Influence versus information in communication. *Biology and Philosophy*, *25*, 755–780.

Palagi, E., & Dapporto, L. (2006). Beyond odor discrimination: Demonstrating individual recognition by scent in *Lemur catta*. *Chemical Senses, 31*, 437–443.

Paoli, T., Palagi, E., & Tarli, S. M. B. (2006). Reevaluation of dominance hierarchy in bonobos (*Pan paniscus*). *American Journal of Physical Anthropology*, *130*, 116–122.

Paoli, T., & Palagi, E. (2008). What does agonistic dominance imply in bonobos? In Furuichi, T., & Thompson, J. (Eds.) *The bonobos: behavior, ecology, and conservation* (pp. 39–54). New York: Springer.

Parish, A.R. (1994). Sex and food control in the "uncommon chimpanzee": How bonobo females overcome a phylogenetic legacy of male dominance. *Ethology and Sociobiolgy*, *15*, 157–179.

Parish, A. R. (1996). Female relationships in bonobos (*Pan paniscus*) – Evidence for bonding, cooperation, and female dominance in a male–philopatric species. *Human Nature – An Interdisciplinary Biosocial Perspective*, *7*, 61–96.

Payne, R. & Webb, D. (1971). Orientation by means of long range acoustic signaling in baleen whales. *Annals of the New York Academy of Science*, *188*, 110–141.

Pennec, F., Krief, S., Hladik, A., Ayingweu, C. L., Bortolamiol, S., Ngawolo, J. B., & Narat, V. (2016). Floristic and structural vegetation typology of bonobo habitats in a forest–savanna mosaic (Bolobo Territory, D.R. Congo). *Plant Ecology and Evolution*, *149*, 199–215.

Pfefferle, D., West, P. M., Grinnell, J., Packer, C., & Fischer, J. (2007). Do acoustic features of lion, *Panthera leo*, roars reflect sex and male condition? *The Journal of the Acoustical Society of America*, *121,* 3947–3953.

Pfefferle, D., Ruiz–lambides, A. V, & Widdig, A. (2014). Female rhesus macaques discriminate unfamiliar paternal sisters in playback experiments: support for acoustic phenotype matching. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20131628.

Pfefferle, D., Ruiz–Lambides, A. V., & Widdig, A. (2015). Male rhesus macaques use vocalizations to distinguish female maternal, but not paternal, kin from non–kin. *Behavioral Ecology and Sociobiology*, *69*, 1677–1686.

Pika, S., Liebal, K., & Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): Gestural repertoire and use. *American Journal of Primatology*, *65,* 39–51.

Pika, S., & Zuberbühler, K. (2008). Social games between bonobos and humans: Evidence for shared intentionality? *American Journal of Primatology*, *70*, 207–210.

Pollard, K. A., & Blumstein, D. T. (2011). Social group size predicts the evolution of individuality. *Current Biology*, *21*, 413–417.

Podos, J., & Warren, P. S. (2007). The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, *37*, 403–444.

Pollick, A. S., & de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 8184–8189.

Pollick, A. S., Jeneson, A., & Waal, F. B. M. De. (2008). Gestures and Multimodal Signaling in Bonobos. In T. Furuichi & J. Thompson (Eds.), *The Bonobos: Behavior, Ecology, and Conservation* (pp. 75–94). New York: Springer.

Poole, J. H., Payne, K., Langbauer, W. R., & Moss, C. J. (1988). The social contexts of some very low–frequency calls of African elephants. *Behavioral Ecology and Sociobiology*, *22*, 385–392.

Poole, J. H., Tyack, P. L., Stoeger–Horwath, A. S., & Watwood, S. (2005). Elephants are capable of vocal learning. *Nature*, *434,* 455–456.

Premack, A. J., & Premack, D. (1972). Teaching language to an ape. Scientific American, 227, 92-99.

Price, T., Arnold, K., Zuberbühler, K., & Semple, S. (2009). Pyow but not hack calls of the male putty–nosed monkey (*Cercopithcus nictitans*) convey information about caller identity. *Behaviour*, *146,* 871–888.

Prufer, K., Munch, K., Hellmann, I., Akagi, K., Miller, J. R., Walenz, B., Koren, S., Sutton, G., Kodira, C., Winer, R., Knight, J., Mullikin, J. C, Meader, S. J., Ponting, C. P., Lunter, G., Higashino, S., Hobolth, A., Dutheil, J., Karakoc, E., Alkan, C., Sajjadian, S., Catacchio, C. R., Ventura, M., Marques–Bonet, T., Eichler, E. E., Andre, C., Atencia, R., Mugisha, L., Junhold, J., Patterson, N., Siebauer, M., Good, J, M., Fischer, A., Ptak, S. E., Lachmann, M., Symer, D. E., Mailund, T., Schierup, M. H., Andres, A. M., Kelso, J., & Paabo, S. (2012). The bonobo genome compared with the chimpanzee and human genomes. *Nature, 486*, 527–531.

Ramos-Fernandez, G. (2005). Vocal communication in a fission–fusion society: Do spider monkeys stay in touch with close associates? *International Journal of Primatology*, *26*, 1077–1092.

Rasoloharijaona, S., Randrianambinina, B., Braune, P., & Zimmermann, E. (2006). Loud calling, spacing and cohesiveness in a nocturnal primate, the Milne Edwards' Sportive Lemur (*Lepilemur edwardsi*). *American Journal of Physical Anthropology*, *129*, 591–600.

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, *65*, 519–530.

Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology*, *28*, 66–71.

Rek, P., & Osiejuk, T.S. (2011). No male identity information loss during call propagation through dense vegetation: the case of the corncrake. Behavioral Processes, *86,* 323–328.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free–ranging Rhesus Monkeys. *Animal Behaviour*, *51*, 1007–1015.

Rendall, D., Owren, M. J., & Rodman, P. S. (1998). The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *Journal of Acoustic Society of American*, *103*, 602–614.

Rendall, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel–like grunt vocalizations of baboons. *The Journal of the Acoustical Society of America*, *113*, 3390–3402.

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel–like baboon grunts: The role of vocalizer body size and voice–acoustic allometry. *The Journal of the Acoustical Society of America*, *117*, 944–955.

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*, 233–240.

Rendell, L., & Whitehead, H. (2001). Culture in whales and dolphins. *The Behavioral and Brain Sciences*, *24*, 309–324.

Richards, D. G., & Wiley, R. H. (1980). Reverberations and amplitude fluctuations in the propagation of sound in a forest : Implications for animal communication. *The American Naturalist*, *115*, 381–399.

Riede, T., Owren, M. J., & Arcadi, A. C. (2004). Nonlinear acoustics in pant hoots of common chimpanzees (*Pan troglodytes*): Frequency jumps, subharmonics, biphonation, and deterministic chaos. *American Journal of Primatology*, *64*, 277–291.

Robisson, P., Aubin, T., & Bremond, J.-C. (1993). Individuality in the Voice of the Emperor Penguin Aptenodytes forsteri: Adaptation to a Noisy Environment. *Ethology*, *94*, 279–290.

Salmi, R., Hammerschmidt, K., & Doran-Sheehy, D. M. (2014). Individual distinctiveness in call types of wild western female gorillas. *PLoS ONE, 9*(7), e101940.

Santorelli, C. J., Aureli, F., Ramos-Fernandez, G., & Schaffner, C. M. (2013). Individual variation of whinnies reflects differences in membership between Spider Monkey (*Ateles geoffroyi*) communities. *International Journal of Primatology. 34*, 1172–1189.

Scarantino, A. (2010). Animal communication between information and influence. *Animal Behaviour, 79*, e1–e5.

Scarantino, A., & Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour, 100*, e1–e8.

Schnitzler, H. U., Moss, C. F., & Denzinger, A. (2003). From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology and Evolution, 18*, 386–394.

Scott–Phillips, T. C. (2008). Defining biological communication. *Journal of Evolutionary Biology, 21*, 387–395.

Searby, A., & Jouventin, P. (2003). Mother–lamb acoustic recognition in sheep: a frequency coding. *Proceedings of the Royal Society B – Biological Sciences, 270*, 1765–1771.

Searcy, W. A. & Nowicki, S. (2010). *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems.* Princeton: Princeton University Press.

Searcy, W. A., Akçay, C., Nowicki, S., & Beecher, M. D. (2014). Aggressive signaling in song sparrows and other songbirds. *Advances in the Study of Behavior, 46*, 89–125.

Semple, S., McComb, K., Alberts, S., & Altmann, J. (2002). Information content of female copulation calls in yellow baboons. *American Journal of Primatology, 56*, 43–56.

Serckx, A., Kuhl, H. S., Beudels–Jamar, R. C., Poncin, P., Bastin, J. F., & Huynen, M. C. (2015). Feeding ecology of bonobos living in forest–savannah mosaics : Diet seasonal variation and importance of fallback foods. *American Journal of Primatology, 77*, 948–962.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free–ranging primate. *Animal Behaviour, 28*, 1070–1094.

Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173.

Seyfarth, R. M., Cheney, D. L., Bergman, T. J., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, *80*, 3–8.

Shapiro, A. D. (2010). Recognition of individuals within the social group: signature vocalizations. In: Brudzynski, S.M. (Ed.) *Handbook of mammalian vocalization—an integrative neuroscience approach* (pp 495–503). Oxford: Academic Press.

Shannon, C.E., & Weaver, W. (1949). *The Mathematical Theory of Information.* Urbana, IL: University of Illinois Press.

Sheehan, M. J., & Tibbetts, E. A. (2009). Evolution of identity signals: Frequency–dependent benefits of distinctive phenotypes used for individual recognition. *Evolution*, *63*, 3106–3113.

Sheehan, M. J., & Tibbetts, E. A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science, 334*, 1272–1275.

Siddiqi, A., Cronin, T. W., Loew, E. R., Vorobyev, M., & Summers, K. (2004). Interspecific and intraspecific views of color signals in the strawberry poison frog, *Dendrobates pumilio*. *The Journal of Experimental Biology*, *207*, 2471–2485.

Silk, J. B. (2002). Kin Seletion in Primate Groups. *International Journal of Primatology*, *23*, 849–875.

Sillero–Zubiri, C., & Macdonald, D. W. (1998). Scent–marking and territorial behaviour of Ethiopian wolves, *Canis simensis*. *Journal of Zoology*, *245*, 351–361.

Slocombe, K. E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, *15*, 1779–1784.

Slocombe, K. E., & Zuberbühler, K. (2006). Food–associated calls in chimpanzees: responses to food types or food preferences? *Animal Behaviour*, *72*, 989–999.

Slocombe, K. E., Kaller, T., Call, J., & Zuberbühler, K. (2010). Chimpanzees Extract Social Information from Agonistic Screams. *PloS One*, *5*(7), 3–8.

Snowdon, C.T., Elowson, A.M. & Roush, R. (1997). Social influences on vocal development in New World primates. In C.T. Snowdon & M. Hausberger (Eds.) *Social influences on vocal development* (pp. 234–248). Cambridge, UK: Cambridge University Press.

Snowdon, C. T., & Elowson, A. M. (1999). Pygmy Marmosets Modify Call Structure When Paired. *Ethology*, *105*, 893–908.

Snowdon, C. T. (2009). Plasticity of Communication in Nonhuman Primates. *Advances in the Study of Behavior, 40,* 239–276.

Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication II: Rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour*, *70*, 589–599.

Spillmann, B., Dunkel, L. P., van Noordwijk, M. A., Amda, R. N. A., Lameira, A. R., Wich, S. A., & van Schaik, C. P. (2010). Acoustic properties of long calls given by flanged male orangutans (*Pongo pygmaeus wurmbii*) reflect both individual identity and context. *Ethology*, *116*, 385–395.

Stevens, J. M. G., Vervaecke, H., de Vries, H., & Van Elsacker, L. (2005). The influence of the steepness of dominance hierarchies on reciprocity and interchange in captive groups of bonobos (*Pan paniscus*). *Behaviour*, *142*, 941–960.

Stevens, J. M. G., Vervaecke, H., de Vries, H., & van Elsacker, L. (2006). Social structures in *Pan paniscus*: Testing the female bonding hypothesis. *Primates*, *47*, 210–217.

Stevens, J. M. G., Vervaecke, H., De Vries, H., & Van Elsacker, L. (2007). Sex differences in the steepness of dominance hierarchies in captive bonobo groups. *International Journal of Primatology*, *28*, 1417–1430.

Stevens, J. M. G., Vervaecke, H., & Elsacker, L. (2008). The bonobo's adaptive potential: social relations under captive conditions. In T. Furuichi & J. Thompson (Eds.) *The bonobos: Behaviour, Ecology and Conservation* (pp. 19– 38). New York: Springer.

Sueur, J., Thierry, A., & Simonis, C. (2008). Equipment review: Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics 18*, 213–226.

Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, *55*, 673–687.

Surbeck, M., & Hohmann, G. (2008). Primate hunting by bonobos at LuiKotale, Salonga National Park. *Current Biology*, *18*, 906–907.

Surbeck, M., & Hohmann, G. (2013). Intersexual dominance relationships and the influence of leverage on the outcome of conflicts in wild bonobos (*Pan paniscus*). *Behavioral Ecology and Sociobiology*, *67*, 1767–1780.

Suzuki, T. N. (2013). Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. *Animal Behaviour*, *87*, 59–65.

Suzuki, T. N. (2015). Assessment of predation risk through referential communication in incubating birds. *Scientific Reports*, *5*, 10239.

Takemoto, H., Kawamoto, Y., & Furuichi, T. (2015). How did bonobos come to range south of the congo river? Reconsideration of the divergence of *Pan paniscus* from other *Pan* populations. *Evolutionary Anthropology: Issues, News, and Reviews*, *24*, 170–184.

Tanaka, T., Sugiura, H., & Masataka, N. (2006). Cross-sectional and longitudinal studies of the development of group differences in acoustic features of coo calls in two groups of Japanese macaques. *Ethology*, *112*, 7–21.

Taylor, A. M., & Reby, D. (2010). The contribution of source–filter theory to mammal vocal communication research. *Journal of Zoology*, *280*, 221–236.

Temeles, E. J. (1994). The role of neighbours in territorial systems: when are they "dear enemies"? *Animal Behaviour*, *47*, 339–350.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black–capped chickadees encode information about predator size. *Science*, *308*, 1934–1937.

Tibbetts, E. A., & Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, *22*, 529–37.

Tokuyama, N., & Furuichi, T. (2016). Do friends help each other? Patterns of female coalition formation in wild bonobos at Wamba. *Animal Behaviour*, *119*, 27–35.

Townsend, S. W., Hollen, L. I., & Manser, M. B. (2010). Meerkat close calls encode group–specific signatures, but receivers fail to discriminate. *Animal Behaviour*, *80*, 133–138.

Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: The Past, Present and the Future. *Ethology*, *119*, 1–9.

Tyack, P.L., & Clark, C.W. (2000). Communication and acoustic behavior of dolphins and whales. In W. Au, A.N. Popper, & R. Fay (Eds.) *Hearing by whales and dolphins – Springer Handbook of Auditory Research Series* (pp. 156–224). New York: Springer.

Tyack, P. L. (2007). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology, 122*, 319–331.

Vannoni, E., & McElligott, A. G. (2007). Individual acoustic variation in fallow deer (*Dama dama*) common and harsh groans: A source–filter theory perspective. *Ethology*, *113*, 223–234.

van Krunkelsven, E., Dupain, J., van Elsacker, L., & Verheyen, R. F. (1996). Food calling by captive bonobos (*Pan paniscus*). *International Journal of Primatology, 17*, 207–217.

Vervaecke, H., Vries, H. De, & van Elsacker, L. (2000a). Dominance and its behavioral measures in a captive group of bonobos (*Pan paniscus*). *International Journal of Primatology*, *21*, 47–68.

Vervaecke, H., de Vries, H., & van Elsacker, L. (2000b). Function and distribution of coalitions in captive bonobos *(Pan paniscus)*. *Primates, 41*, 249–265.

Vignal, C., Mathevon, N., & Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature*, *430*, 448–451.

Vignal, C., Mathevon, N., & Mottin, S. (2008). Mate recognition by female zebra finch: Analysis of individuality in male call and first investigations on female decoding process. *Behavioural Processes*, *77*, 191–198.

Voigt–Heucke, S. L., Taborsky, M., & Dechmann, D. K. N. (2010). A dual function of echolocation: bats use echolocation calls to identify familiar and unfamiliar individuals. *Animal Behaviour*, *80*, 59–67.

Waser, P.M. (1977). Individual recognition, intragroup cohesion and intergroup spacing: evidence from sound playback to forest monkeys. *Behaviour, 60*, 28–74.

Waser, P. M., & Brown, C. H. (1986). Habitat acoustics and primate communication. *American Journal of Primatology*, *10*, 135–154.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C.,Wallace, E.K., Cheng, L., West, V., & Slocombe, K. E. (2015). Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. *Current Biology*, 25, 495–499.

Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology*, *21*, 195–205.

Wheeler, B. C., & Fischer, J. (2015). The blurred boundaries of functional reference: a response to Scarantino & Clay. *Animal Behaviour*, *100*, e9–e13.

White, F.J. (1996). Comparative socio–ecology of *Pan paniscus*. In W. C. McGrew, L.F. Marchant & T. Nishida (Eds.), *Great Ape Societies* (pp. 29–41). Cambridge, UK: Cambridge University Press.

White, F. J. (1998). Seasonality and socioecology: the importance of variation in fruit abundance to bonobo sociality. *International Journal of Primatology*, *19*, 1013–1027.

Wich, S. A., & Nunn, C. L. (2002). Do male "long–distance calls" function in mate defense? A comparative study of long–distance calls in primates. *Behavioral Ecology and Sociobiology*, *52*, 474–484.

Wich, S. A., Koski, S., De Vries, H., & Van Schaik, C. P. (2003). Individual and contextual variation in Thomas langur male loud calls. *Ethology*, *109*, 1–13.

Widdig, A. (2007). Paternal kin discrimination: the evidence and likely mechanisms, *Biological Reviews*, *49*, 319–334.

Wildman, D. E., Uddin, M., Liu, G., Grossman, L. I., & Goodman, M. (2003). Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: Enlarging genus Homo. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 7181–7188.

Wilkinson, G. S. & Boughman, J. W. (1998). Social calls coordinate foraging in greater spear-nosed bats. *Animal Behaviour, 55*, 337–350.

Wilson, M. L., & Wrangham, R. W. (2003). Intergroup relations in chimpanzees. *Annual Review of Anthropology*, *32*, 363–392.

Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I.C., Hashimoto, C., Hobaiter, C.L., Hohmann, G., Itoh, N., Koops, K., Lloyd, J.N., Matsuzawa, T., Mitani, J. C., Mjungu, D.C., Morgan, D., Muller, M.N., Mundry, R., Nakamura, M., Pruetz, J., Pusey, A.E., Riedel, J., Sanz, C., Schel, A.M. Simmons, N., Waller, M., Watts, D.P., White, F., Wittig, R.M., Zuberbühler, K., & Wrangham, R.W. (2014). Lethal aggression in Pan is better explained by adaptive strategies than human impacts. *Nature*, *513*, 414–417.

Winter, P., Handley, P., Ploog, D., & Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour, 47,* 230–239.

Wyman, M. T., Mooring, M. S., McCowan, B., Penedo, M. C. T., Reby, D., & Hart, L. A. (2012). Acoustic cues to size and quality in the vocalizations of male North American bison, *Bison bison*. *Animal Behaviour*, *84*, 1381–1391.

Zimmerman, E. (1995). Loud calls in nocturnal prosimians: Structure, evolution and ontogeny. In E. Zimmerman, J. D. Newman, & U. Jurgens (Eds.) *Current Topics in Primate Vocal Communication* (pp. 311–330). New York, London: Plenum Press.

Zuberbühler, K. (2000). Referential labelling in wild Diana monkeys. *Animal Behaviour, 59*, 917–927.

Zuberbühler K. (2003). Referential signaling in non–human primates: cognitive precursors and limitations for the evolution of language. *Advances in the Study of Behaviour, 33*, 265–307.

Zuberbühler, K. (2005). Linguistic prerequisites in the primate lineage. In M. Tallerman (Ed.) *Language origins: Perspectives on evolution* (pp. 262-282). Oxford: Oxford University Press.

All photographs used in this thesis are copyright of Sumir Keenan, *except* : Figure 5b (of SK during data collection), Figure 9c and photographs on page 69 (three images on the top left, one on the bottom left) and 105 (single image on top right) are courtesy of Jeroen Stevens; and Figure 6 is courtesy of Gladez Shorland.

TITLE : IDENTITY INFORMATION IN BONOBO IN VOCAL COMMUNICATION : FROM SENDER TO RECEIVER

Identity information is vital for highly social species as it facilitates individual recognition and allows for differentiation between social partners in many contexts, such as dominance hierarchies, territorial defence, mating and parent-offspring identification, and group cohesion and coordination. For many species vocalisations can be the most effective communication channel in complex environments and over long-distances and are encoded with the stable features of an individual's voice. Associations between these individual vocal signatures and accumulated social knowledge about conspecifics can greatly increase an animal's fitness, as it facilitates adaptively constructive social decisions. This thesis investigates the encoding and decoding of identity information in the vocal communication system of the bonobo, *Pan paniscus.* We firstly investigated the stability of vocal signatures across the five most common call types in the bonobo vocal repertoire. Results showed that while all call types have the potential to code identity information, loud calls used during times of high arousal and for distance communication have the strongest individual vocal signatures. Following the first study, we investigated if social familiarity and relatedness affect the acoustic features that code individual information in the bark call type. Overall, we found strong evidence for vocal convergence, and specifically, that individuals who are related and familiar, independently from one another, are more vocally similar to one another than unrelated and unfamiliar individuals. In a final study, we tested if bonobos are capable of using the encoded identity information to recognise past group members that they no longer live with. Through a series playback experiments we demonstrated that bonobos are capable of recognising familiar individuals from vocalisations alone even after years of separation. Collectively, the results of this thesis show that the encoding and decoding of identity information in bonobo vocalisations is a dynamic system, subject to modification through social processes but robust enough to allow for individual recognition over time. In conclusion, these studies contribute to a better understanding of the vocal communication system of a non-human primate species with a unique and complex social network.

Key words: vocal communication, vocalisation, individual vocal signature, identity information, individual vocal recognition, vocal convergence, kin vocal signature, non-human primate, bonobo

TITRE : L'INFORMATION "IDENTITÉ INDIVIDUELLE" DANS LA COMMUNICATION VOCALE DU BONOBO: DE L'ÉMETTEUR AU RÉCEPTEUR

L'information "identité individuelle" est essentielle chez les espèces fortement sociales car elle permet la reconnaissance individuelle et la différenciation des partenaires sociaux dans de nombreux contextes tels que les relations de dominance, les relations mère-jeunes, la défense territoriale, ou encore participe à la cohésion et coordination de groupe. Chez de nombreuses espèces, le canal audio est l'une des voies les plus efficaces de communication dans des environnements complexes et à longue distance. Les vocalisations sont empreintes de caractéristiques acoustiques propres à la voix de chaque individu. La combinaison entre ces signatures vocales individuelles et la connaissance sociale accumulée sur les congénères peut grandement favoriser la valeur sélective des animaux, en facilitant notamment les prises de décisions sociales les plus adaptées. Le but de ma recherche est d'étudier le codage et décodage de l'information "identité individuelle" du système vocal de communication du bonobo, *Pan paniscus*. Premièrement, nous avons recherché la stabilité des signatures vocales des cinq types de cris les plus courants du répertoire du bonobo. Nous avons trouvé que, bien que ces cinq types de cris aient le potentiel de coder l'information individuelle, les cris les plus forts émis dans des contextes d'excitation intense et de communication à longue distance ont les signatures vocales individuelles les plus marquées. Deuxièmement, nous avons étudié l'effet de la familiarité sociale et des liens de parenté sur les caractéristiques acoustiques qui codent l'information individuelle dans un type de cri "*bark*". Nous avons mis en évidence l'existence d'une forte convergence vocale. Les individus apparentés et familiers, et indépendamment l'un de l'autre, présentent plus de similarités vocales qu'entre des individus non apparentés et non familiers. Enfin, dans une troisième étude, nous avons testé la capacité des bonobos à utiliser l'information "identité individuelle" codée dans les vocalisations pour discriminer la voix d'anciens partenaires sociaux avec qui ils ne vivent plus. Par une série d'expériences de repasse, nous avons démontré que les bonobos étaient capables de reconnaître la voix d'individus familiers sur la seule base de l'acoustique, et cela même après des années de séparation. L'ensemble de ce travail de thèse montre que le codage et décodage de l'information "identité individuelle" chez le bonobo est un système dynamique, sujet à modification avec l'environnement social mais suffisamment fiable pour permettre la reconnaissance individuelle au cours du temps. En conclusion cette étude participe à une meilleure compréhension du système de communication vocale chez un primate non-humain forestier, au réseau social unique et complexe.

Mots clés: communication vocale, vocalisation, signature vocale individuelle, information de l'identité, reconnaissance vocale individuelle, convergence vocale, signature vocale d'apparentement, primate non-humain, bonobo

Discipline : Bioacoustique, Primatologie.

Laboratoire : Equipe de Neuro-Ethologie Sensorielle, ENES-CNPS CNRS UMR 8195. Université de Lyon/Saint-Etienne, Saint-Etienne, France **&** Department of Psychology & Neuroscience, University of St. Andrews, St. Andrews, Scotland