



Sequential and network analyses to describe multiple signal use in captive mangabeys

Juliette Aychet, Catherine Blois-Heulin, Alban Lemasson

► To cite this version:

Juliette Aychet, Catherine Blois-Heulin, Alban Lemasson. Sequential and network analyses to describe multiple signal use in captive mangabeys. *Animal Behaviour*, 2021, 182, pp.203-226. 10.1016/j.anbehav.2021.09.005 . hal-03480471

HAL Id: hal-03480471

<https://hal.science/hal-03480471>

Submitted on 15 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Sequential and network analyses to describe multiple signal use in captive mangabeys

Juliette Aychet ^a, Catherine Blois-Heulin ^{*a}, Alban Lemasson ^{*a,b}

* same contribution

^a Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35380 Paimpont, France

^b Institut Universitaire de France

Corresponding author

Juliette Aychet

Juliette.aychet@gmail.com

+33(0)6 67 46 56 84

UMR 6552, Université de Rennes 1, Station Biologique de Paimpont, 35380 Paimpont, France

Declaration of interest

The authors declare no conflict of interest.

Acknowledgements

We are thankful to Arnaud Rossard and Philippe Bec for taking care of the monkeys and for their flexibility regarding our observation schedule, and to Véronique Guyot for her advice on statistical analyses.

ABSTRACT

Although studies on primate communication have mainly focused on single communication systems (gestural, facial or vocal), there is an increasing interest for a more integrative approach to describe the communication of our closest relatives, particularly with the aim of investigating the evolutionary roots of human language. The understanding of the ultimate functions of multimodal communication (involving signals of different sensory modalities) and multicomponentiality (association of different signal types) needs more systematic description of multiple signal use in primate species, with details on the circumstances leading to such complex signalling. In the present study, we describe the sequential use of communicative signals of different types and sensory modalities in a captive population of catarrhine monkeys, the red-capped mangabeys (*Cercocebus torquatus*). We applied existing analysis methods from other disciplines to define, systematically describe and quantify the production of signals among communicative sequences. We notably used sequence analysis tools (based on dissimilarity measures) to identify the typical signal sequences produced by mangabeys, and network analysis to describe dyadic signal associations among these sequences. The focal observation of five social groups of mangabeys allowed us to identify 424 communicative sequences, which could be grouped in eight main categories, and whose complexity, multimodality and multicomponentiality, were depending on social context and signaller characteristics. Overall, captive mangabeys frequently associated communication signals of all types (body, facial and vocal signals) and modalities (visual, audible and tactile), in a flexible way. Our results complete previous description of red-capped mangabey signalling, and highlight the need for a multimodal and multicomponent approach to understand the complexity of primate communication. Moreover, we propose the method we used as a way to enhance primate communication analysis, in the frame of comparative research.

Keywords

Gestures, Facial expressions, Flexibility, Multicomponent communication, Multimodal communication, Monkeys, Network analysis, Sequence analysis, Signal combination, Vocalisations

Language is an intrinsically multidimensional communication system, both regarding the production of different types of signals and their perception on different sensory modalities (Levinson & Holler, 2014; Vigliocco et al., 2014). Humans (*Homo sapiens*) frequently combine gestures, body movements and facial expressions with speech (Ekman, 1979, 2004; Goldin-Meadow, 1999; Iverson & Goldin-Meadow, 1998; Levinson & Holler, 2014; Özyürek & Woll, 2019). Co-speech gestures occur even in the absence of learning (Iverson & Goldin-Meadow, 1998), and pre-verbal children communicate with combinations of sounds, gestures and facial expressions (Tronick et al., 1978). Co-speech signals are presumably a way to enhance communication (Wu & Coulson, 2007): for instance, gestures might permit to disambiguate speech (Holle & Gunter, 2007), as receivers rely on both visual and vocal modalities to interpret the message conveyed (Cassell et al., 1999; Goldin-Meadow et al., 1992; Thompson & Massaro, 1986). Thus, it is often claimed that research on language evolutionary origins, which is commonly based on comparisons between human and non-human animal communication systems (e.g. Arbib et al., 2008; Liebal et al., 2014; Pika, 2008; Pollick & De Waal, 2007; Prieur et al., 2018, 2019; Scott-Phillips, 2015; Sievers & Gruber, 2020; Vauclair, 2013; Zuberbühler, 2005), should account for this multimodal characteristic of language, especially by adopting a more integrative approach to study the communication of our closest relatives, non-human primates (Fröhlich et al., 2019; Fröhlich & van Schaik, 2018; Liebal et al., 2014; Waller et al., 2013). As a matter of fact, non-human primate communication can also involve different types and modalities of signals produced together (e.g. in chimpanzees, *Pan troglodytes*: Hobaiter et al., 2017; Pollick & De Waal, 2007; Tagliatela et al., 2015; Wilke et al., 2017; bonobos, *Pan paniscus*: Genty, 2019; Genty et al., 2014; Pollick & De Waal, 2007; howler monkey, *Alouatta palliata*: Jones & van Cantfort, 2007; crested macaques, *Macaca nigra*: Micheletta et al., 2013), but communication systems are generally investigated separately, gestural communication being mainly studied in great apes, and vocal and facial communication in other species (Slocombe et al., 2011). This approach is sometimes inherent to the research question, given that different types of signalling such as gestures and facial expressions may have different underlying cognitive processes (Waller et al., 2013), or it can be constrained by methodological limitations (Liebal et al., 2014a). However, it is increasingly acknowledged that focusing on a unique type or modality of signalling prevents researchers from apprehending the whole complexity of primate communication (Liebal et al., 2014; Waller et al., 2013).

In the animal communication literature, “multimodal communication” sometimes refers to the use of signals simultaneously perceived through more than one sensory modality (for instance, auditory and visual: Partan & Marler, 1999; Partan & Marler, 2005), but can also be defined as the use of multiple signals from different production channels (for instance, gestures associated with facial expressions or vocalisations: e.g. Genty et al., 2014; Liebal et al., 2014; Pollick & De Waal, 2007; Tagliatela et al., 2015). Following Higham & Hebets (2013), we differentiate here “multicomponent communication” from “multimodal communication”. “Multicomponent communication” refers to communication that

implies different types of signals (notably gestures or body postures, facial expressions, vocalisations) delivered on one or more sensory channels, while “multimodal communication” refers to the production of several signals perceived on more than one sensory channel (notably visual, acoustic or tactile). We note that some signals can be perceived on different sensory channels in their inherent form, in a way that one modality is inevitably associated to the other (“fixed” multimodality, Partan & Marler, 2005). For instance, some vocal signals, even though they are perceived on the acoustic modality, are necessarily associated to specific facial movements that are perceived on the visual channel (Ghazanfar et al., 2005; Ghazanfar & Logothetis, 2003; Izumi & Kojima, 2004). In contrast, we will focus here on “free” multimodality (Fröhlich & van Schaik, 2018; Higham & Hebets, 2013; Partan & Marler, 2005; Wilke et al., 2017), which implies signal modalities that are not obligatory coupled, and whose associations may reflect a flexible use of communication signals, depending on social and environmental factors (Fröhlich & van Schaik, 2018). The “free” association of signal types and modalities, in simultaneous combinations or in sequences, have been mainly investigated in animals other than primates (e.g. in the courtship displays of jumping spiders, *Habronattus coecatus*: Elias et al. 2012; in numerous songbird species such as brown-headed cowbirds, *Molothrus ater*, Barbary dove, *Streptopelia risoria*, carib grackles, *Quiscalus lugubris*, and zebra finches, *Taeniopygia guttata*: Cooper & Goller, 2004; Hutchison et al., 1997; Wiley, 1975; Williams, 2001; in the alarm behaviour of Eastern gray squirrels, *Sciurus carolinensis*: Partan et al., 2010), yet reports of multimodal and multicomponent communication in primates are still rare and concern few species (i.e., captive and wild chimpanzees: Hobaiter et al., 2017; Leavens et al., 2004, 2010; Pollick & De Waal, 2007; Tagliabata et al., 2015; Wilke et al., 2017; bonobos: Genty, 2019; Genty et al., 2014, 2015; Pollick & De Waal, 2007; orangutans, *Pongo pygmaeus*: Knox et al., 2019; siamangs, *Symphalangus syndactylus*: Liebal et al., 2004a; crested and Rhesus macaques, *M. nigra* and *mulatta*: Micheletta et al., 2013; Partan, 2002; howler monkeys: Briseño-Jaramillo et al., 2015; Jones & van Cantfort, 2007).

Different hypotheses have been formulated on the ultimate functions of producing multiple signals of different types and modalities, based on selection pressures which could act at the content-level or efficacy-level of signal production (e.g. Guilford & Dawkins, 1991; Hebets & Papaj, 2005; Partan & Marler, 2005; reviewed in Fröhlich & van Schaik, 2018). To investigate the function of signal associations and understand the evolution constraints that shaped primate multimodal and multicomponent communication (including human language), a first necessary step is to formally describe these communication events. Systematic descriptions of multiple signal use are thus needed, including all signal types and modalities (Fröhlich & van Schaik, 2018; Liebal et al., 2014; Waller et al., 2013), together with reports on the circumstances leading to such complex communication (social contexts, physical environment, developmental stage, etc.). However, to our knowledge, the description of multimodal and multicomponent communication in primates is generally centred on one specific type of association (e.g. vocal and gestural signals: Genty, 2019; Pollick & De Waal, 2007) or on different

uses of one specific signal (e.g. lipsmacks in crested macaques: Micheletta et al., 2013; “pout-moan” and “contest hoot” calls in bonobos: Genty, 2019; Genty et al., 2014), but rarely addresses the entire repertoire of signal associations, i.e. of simultaneous combinations and/or signal sequences (but see the description of multimodal signals in chimpanzees by Wilke et al., 2017). Moreover, in order to draw distinctions between the effect of sociality and environment on the production of multimodal and multicomponent communication, this has to be investigated in a sufficient diversity of species, both in captive and wild environments (Fröhlich & van Schaik, 2018).

In order to permit comparisons between species and environmental conditions, primate multimodal communication studies have to be based on similar methodologies, notably regarding their working definitions of multimodality and multicomponentiality. The differential use of the word “multimodal” explained above could be easily identified and is not a major obstacle while comparing study results. However, one constraining methodological inconsistency is the temporal definition used to identify multiple signal use. Some studies consider signals as “combined” only if the signals produced by a same individual temporally overlap (e.g. Wilke et al., 2017), while some others will consider both overlapping and sequences of signals, i.e. signals separated by a pause up to an arbitrary value (e.g. 1 sec: Genty, 2019; Genty et al., 2015; Hobaiter et al., 2017; Knox et al., 2019; 2 sec: Taglialatela et al., 2015; and in gestural studies 1 sec: Genty & Byrne, 2010; Hobaiter & Byrne, 2011a; 5 sec: Liebal et al., 2004b, 2006; Tempelmann & Liebal, 2012), or signals which initiations are separated by an arbitrary latency (e.g. 10 sec: Pollick & De Waal, 2007). This may lead to identify more or less instances of multimodal or multicomponent signal use, which prevent reliable comparisons on their frequencies in different species and under different conditions. In order to prevent any misunderstanding regarding the description of multiple signal use, we will differentiate in the present paper the terms “signal combinations” (i.e. the simultaneous production of several signals), “signal sequences” (i.e. series of signals which are initiated within a defined latency, and which partially overlap or not), and “signal associations” (i.e. signals produced together, in combinations and/or sequences). In addition, one methodological problem that can arise when describing multimodal and/or multicomponent communication in a population is whether we should account for slight variations in sequential signal production. The less frequent signal associations could either be random variations of more frequent ones, or could be biologically relevant by themselves, hence it is important to analyse them in their entirety without transformation (Baraud et al., 2016). One approach to acknowledge this variability in multiple signal production, without missing their similarities, can be to focus on one unique signal and report all possible use, with all possible associations and their contexts (for instance, see the description of “pout-call” use in young bonobos: Genty, 2019). However, this approach needs a preliminary selection of one signal of interest, and is likely to be time-consuming if applied to the entire communication repertoire of a species. Thus, another method is to use sequence analysis tools to

describe the structure of such signal use (e.g. Baraud et al., 2016), taking into account all slight variations observed in the population while grouping similar events.

Although rare studies have addressed the description of gesture sequence structures in primates (e.g. exhaustive reports of sequence compositions and their frequencies in chimpanzees: Liebal et al., 2004b; or use of Markov diagrams to represent the most frequent gesture associations in gorillas: Genty & Byrne, 2010), most methodological development to describe animal communication sequence structure arises from studies on vocal communication. Above various methodologies used to describe the variability in communication sequences (reviewed in Kershenbaum et al., 2016), one performant tool is the measure of between-sequence dissimilarities, such as the Levenshtein distance, based on the number of operations needed to transform one sequence into another (i.e. insertions, deletions or substitutions of units) (Abbott & Forrest, 1986; Levenshtein, 1966; Needleman & Wunsch, 1970; Studer & Ritschard, 2016; see Kershenbaum & Garland, 2015, for comparisons with other metrics). While such measures are still rarely used in the field of animal behaviour (but applied in other disciplines such as social sciences or bioinformatics, e.g. Abbott & Forrest, 1986; Abbott & Tsay, 2000; Li & Becchi, 2012), some studies have used it to describe animal vocal sequences (e.g. in humpback whales, *Megaptera novaeangliae*: Garland et al., 2012, 2013; in rock hyraxes, *Procavia capensis*: Kershenbaum et al., 2012; in Australian gannets, *Morus serrator*: Krull et al., 2012). Another promising approach, which could complete the analysis of sequence structure, is the use of networks to describe signal associations and assess their relative importance in complex communicative events (Kershenbaum et al., 2016). The relevance of network statistics and diagrams for the study of animal communication has notably been demonstrated by their application to describe vocal units composing birdsongs (e.g.. in house wrens, *Troglodytes musculus*: Deslandes et al., 2014; in California thrasher, *Toxostoma redivivum*: Sasahara et al., 2012; in common nightingales, *Luscinia megarhynchos*: Weiss et al., 2014). More recently, network analyses have been applied in the field of primate communication, to describe the association of vocal units in chimpanzee (Girard-Buttoz et al., 2021), and to develop an analytical tool for facial expression data resulting from the Facial Action Coding System (“NetFACS”: Mielke et al., 2020; FACS: Ekman et al., 1997). These methods could also be appropriate to describe primate signal associations with a multimodal and multicomponent approach, and therefore constitute a possible way for the development of a rigorous and widely usable methodology currently needed in this area.

In the present study, we adopted such an approach to define and systematically describe the use of multiple communicative signals in a captive population of catarrhine monkeys, the red-capped mangabeys (*Cercocebus torquatus*). Red-capped mangabeys are semi-terrestrial monkeys (Mitani et al., 2012) originating from West African rainforest coastal regions (Cooke, 2012; Gautier-Hion et al., 1999; Jones & Sabater-Pi, 1968; Orimaye, 2017), and naturally living in large multi-male, multi-female groups from 10 to 25 individuals (Gautier-Hion et al., 1999; Jones & Sabater-Pi, 1968). In captivity, they are

organised in steep linear hierarchy (Ruth Dolado & Beltran, 2011) and exhibit aggression and affiliation patterns observed in both despotic and more relaxed macaque societies (Dolado & Beltran, 2012; e.g. unidirectionality of aggressions, high rate of aggressions, low aggression intensity, reciprocal grooming behaviour. Thierry, 2007; Thierry et al., 2000). Communicative signals of captive red-capped mangabeys have been previously described, bringing to light a rich set of spontaneous gestures and body signals (produced either on the visual, audible or tactile channel, Schel et al., under review), as well as facial (Aychet et al. 2021) and vocal signals (Bouchet et al., 2010). Moreover, a first description of their multimodal interactions, centred on their use of vocal signals, suggest flexible multimodal and multicomponent signalling in this species (Baraud et al., 2016).

Our objective here was threefold. First, we attempted to complete the communicative repertoire of captive red-capped mangabeys, by describing the typical sequences of signals they spontaneously produce in their intraspecific dyadic interactions, and identifying the most commonly associated signals among these sequences. Second, our aim was to describe the social circumstances leading to the production of complex signalling. For this purpose, we tested the effect of social context and characteristics of signallers on signal sequence production. Considering that such effects have been identified on the production of facial and vocal signals in red-capped mangabeys (Aychet et al., 2021; Bouchet et al., 2010), we expected that different typical sequences of signals would be produced in specific social contexts and depending on the sex and age of the signaller. Moreover, we made the hypothesis that social context would affect the characteristics of the communicative sequences produced by mangabeys, such as the number and diversity of associated signals, as well as the sequence complexity, multimodality and multicomponentiality. These hypotheses were based on the assumption that the production of multiple signals would have particular functions in communication and would depend on different motivations from the signaller. Our third objective was to address the above-cited methodological issues regarding the description of signal associations. In that aim, we applied methods that we propose as a way to enhance future comparisons between descriptions of primate species multimodal and multicomponent communication. We notably propose to base the temporal definition of signal sequence on a less arbitrary criterion, but also adapted to the study species. Based on the method used in vocal exchange studies (Alban Lemasson et al., 2010, 2018; Levréro et al., 2019), we established the sequence definition time-window on the distribution of latencies between two signals that was actually observed in the studied population. Additionally, we used a sequence analysis methodology based on dissimilarity measures (Gabadinho et al., 2011) to describe red-capped mangabey sequences of signals, and network analysis (Whitehead, 2008) to describe signal associations among these communicative sequences, and particularly investigate how signal types (body, facial, vocal signals) and modalities (visual, audible, tactile) were integrated in complex communication events.

METHODS

Ethical note

This purely observational study was conducted without modifying the living conditions of the captive mangabeys housed at the Station Biologique de Paimpont (University Rennes 1, France), where animal facilities and animal care procedures are regularly monitored by the responsible local authorities (Housing agreement for research D35-211-18, delivered by the “Direction Départementale de la Cohésion Sociale et de la Protection des Populations” (DDCSPP)).

Subjects

We studied twenty-five captive red-capped mangabeys (*Cercocebus torquatus*), housed at the Biological Station of Paimpont (University of Rennes 1, France) and living in social groups composed of two to 13 individuals (**Table 1**), whose compositions have been stable for at least 8 months. Individuals were considered as adults when they were older than 4 years old for females and older than 7 years old for males, based on demographic data on a closely related species, grey-cheeked mangabeys (*Lophocebus albigena*) (Chalmers, 1968; Deputte, 1992; Gautier-Hion & Gautier, 1976). Old adults were differentiated from middle-aged adults if they were more than 15 years old, corresponding to the median lifespan of mangabeys in captivity (based on 45 years data from our breeding facility on grey-cheeked and red-capped mangabeys).

Mangabeys were housed in outdoor-indoor enclosures of different sizes (from 8 to 26.4 m² for indoor enclosures, 14.7 to 37.2 m² for outdoor enclosures, and height from 2.5 m to 4.4 m). Individuals were free to move in and out at all time, using connecting tunnels. Indoor enclosure temperature was maintained at 22°C. Each enclosure was enriched with wood and metal perches, completed with chains or hessian ribbons. The floor of the indoor enclosures was covered with straw and sawdust, while the outdoor enclosure floor was covered with cement or bark. Mangabeys were fed twice a day with fresh fruits and vegetables in the morning and with monkey chows in the afternoon. Water was available *ad libitum* in their indoor enclosures.

Data collection

Observations

We collected 52.22 hours of observations, from 29 January 2018 to 27 June 2018. Applying the individual focal sampling method (Altmann 1974), we recorded focal videos of about 15 minutes per session (mean session duration \pm SE = 15.67 \pm 0.10 min) using a JVC Full HD GZ-RX615 camcorder. Each subject was observed during 8 sessions, for a total of more than 2 hours (125.33 \pm 0.75 min). These sessions were scheduled semi-randomly, so that all individuals were observed at all times (from 9 am to 6 pm) and all periods of day (before or during first feeding / between two feedings / after or during second feeding). We transferred the videos to a computer for further video analyses. Software BORIS v.6.0.6 (Friard and Gamba 2016) was used to play videos and to code each event of communicative signals and each social interaction involving the focal individual. For each signal we then coded associated parameters and behaviours (i.e. signaller and receiver identities, signal characteristics, social context).

Communicative signals

We recorded and identified every occurrence of communicative signals produced by or in the direction of the focal individual. Signals were categorized in three types: “body signals”, “facial signals” and “vocal signals”. The modalities on which they could be perceived by the receiver were classified as: “visual” if signals were silent distant movements of the face or the body; “tactile” if they included physical contact between the signaller and the receiver; “audible” if they included vocal or nonvocal sound production (Fröhlich et al., 2016; Liebal, Call, et al., 2004; Pika, 2008; Pika et al., 2003; Schel et al., under review).

Body and facial signals were defined as movements of the head, limbs, body or part of the face that were (i) mechanically ineffective, i.e. not designed to act as direct physical agents on receiver, so that such signals could be differentiate from non-communicative social actions, though not excluding signals that implied a contact with a substrate or the receiver (e.g. Pollick & De Waal, 2007; Pollick et al., 2008; Gupta & Sinha, 2019; Schel et al., under review); (ii) physically directed toward a receiver, i.e., if at least one of the following markers of directionality was observed during signal production: signaller’s head or body was oriented towards the receiver; the signal was a movement produced in the direction of the receiver; the signaller was approaching the receiver; the signaller and the receiver were in physical contact (Aychet et al., 2021; Call & Tomasello, 2007; Leavens & Hopkins, 1998; Liebal et al., 2004a;

Tomasello et al., 1985, 1994; Schel et al., under review) and/or the signal was preceded by “audience checking” within the 5 seconds before (e.g. Aychet et al., 2021; Hobaiter & Byrne, 2011b); (iii) adapted to the receiver attention, i.e. for visual only signals, produced in front of a visually attentive recipient (the signaler being more or less 45° in front of the receiver’s face) (e.g. Aychet et al., 2021; Pika et al., 2003; Liebal et al., 2004a; Pollick & De Waal, 2007; Genty et al., 2009; Schel et al., under review). Moreover, to check for the communicative value of these movements, we kept in our analyses only body and facial signals that could induce a voluntary response from the receiver, i.e. for which we observed at least once a change in the receiver behaviour within the 5 seconds following the signal (e.g. Aychet et al., 2021; Liebal et al., 2004a, 2006; Roberts et al., 2014; Schel et al., under review) (See appendix **Figure A1** for details on the receiver’s responses observed). Body and facial signals were identified based on previous repertoire descriptions (Aychet et al., 2021; Schel et al. under review), that we completed when necessary. Vocal signals were identified following the repertoire described by Bouchet et al. (2010). Unlike body and facials signals, vocalisations cannot be confused with non-communicative physical actions, hence we only verified for each vocal signal occurrence that it was part of a dyadic communication event, i.e. directed to a specific receiver (see (ii)).

Contexts of signal production

We coded the social contexts in which signals were produced according to the signaller and recipient’s behaviours observed within the 5 seconds before and the 5 seconds after the signal production (Aychet et al., 2021). Six social context categories were distinguished: affiliative, grooming, sexual, playful, aggressive, and submissive (**Table 2**). Moreover, for the purpose of weighting contextual data in our analysis according to the overall importance of each of these categories, we also established a population-level time budget. Thus, we also coded every 10 sec (scan sampling method, Altmann, 1974) the social and non-social contexts in which the focal individual was involved, for one random session per individual (totalizing 2250 scans for the population).

Sequence analysis

Sequence definition and characteristics

We defined as “sequence of signals” a series of signals that were directed by a same signaller towards a same receiver, and whose initiations were separated by latencies from 0 to 8 seconds. We determined this time window graphically, based on the observed frequency distribution of latencies between two

consecutive signal initiations by a same signaller toward a same receiver in our red-capped mangabeys (**Figure 1**). The threshold latency for the sequence definition corresponded to the maximum value beyond which the frequencies were below a graphically determined baseline, representing apparent “background noise” (i.e. below which the observed frequencies seemed randomly distributed, going up and down. Lemasson et al., 2010, 2018; Levréro et al., 2019).

We extracted all the so-defined communicative sequences from our data set, and used the package {TraMineR} (Gabadinho et al., 2011) from the software R v.4.0.4 (R Core Team, 2021) to visualize and analyse them. Communicative sequences were characterized based on their length (number of signals within the sequence), diversity (number of different signals within the sequence), and complexity. The complexity index, ranging from 0 to 1, is based on longitudinal entropy and number of transitions within the sequence (Gabadinho et al., 2010, 2011). Thus, sequences composed of few different signals, consecutively repeated, will have low complexity indices; in contrast, the index will rise if signals are less repeated among the sequence and more diverse. Moreover, we defined as “multicomponent”, as opposed to “unicomponent”, the sequences that were composed of more than one type of signal (for instance those that were composed of both body and facial signals). We also defined as “multimodal”, as opposed to “unimodal”, the sequences that were composed of more than one modality of signal (i.e. “free” multimodality, Fröhlich & van Schaik, 2018; Higham & Hebets, 2013; Partan & Marler, 2005; Wilke et al., 2017: for instance visual and tactile signals, or visual and audible signals).

Sequence clustering and representative sequences

In order to describe the typical communicative sequences produced by red-capped mangabeys, we grouped them based on their composition similarities and extracted “representative sequences” for each group. To do so, we used the package {TraMineR} (Gabadinho et al., 2011) and first computed dissimilarities between observed sequences, based on the “optimal matching method” (Abbott & Forrest, 1986; Abbott & Tsay, 2000; Levenshtein, 1966; Needleman & Wunsch, 1970; Studer & Ritschard, 2016). A distance was measured for each sequence dyad of our dataset, corresponding to the minimum cost needed to transform one sequence into another, considering that each insertion, deletion or substitution of signal has a particular cost. The insertion / deletion cost was set at a constant value (indel = 1). The substitution costs were derived from the observed transition rates between signals, so that the cost of substituting between signals that were frequently produced consecutively was lower than substituting between signals that were rarely produced one after the other (mean cost $\pm S.E.$ = 1.917 ± 0.003). Moreover, as we analysed sequences of different lengths, we applied Abbott’s normalization (Gabadinho et al., 2011) to the “optimal matching” distances, by dividing each distance by the length of the longest of the two corresponding sequences. Based on these normalized distances, we made a

hierarchical clustering of the sequences using Ward method (Ward, 1963), and the number of sequence groups was determined graphically based on the generated dendrogram plot.

For each sequence group, we extracted a set of representative sequences based on the “neighbourhood density” criterion (Gabadinho et al., 2011; Gabadinho & Ritschard, 2013). This method permits to define an as small as possible set of non-redundant representatives which cover a minimum percentage of the sample sequences in their “neighbourhood”. Here, we chose to set this minimum coverage at 50%, but we dropped this threshold to 25% for the most heterogeneous groups, i.e. when the needed number of representatives to cover 50% of the group was about a quarter of the group size. Were considered as “neighbours” the sequences for which the distance was under 10% of the maximum theoretical distance (radius = 0.10).

Sequence context

The contexts of production of the communicative sequences were determined based on the context category already assigned to the signal occurrences they were composed of (see above, **Table 2**). When more than one context were identified, we categorized the overall sequence context based on the one that was attributed to the first signal of the sequence.

Signal associations

In order to represent and quantify the associations of different signals, signal types and signal modalities among the communicative sequences of red-capped mangabeys, we applied graphical and statistical methods of network analysis. We discarded here the communicative sequences composed of a unique repeated signal, given that they did not imply different signal associations. We used SOCPROG 2.9 software (Whitehead, 2009) to compute association indices for each signal dyad (simple ratio, Ginsberg & Young, 1992). In order to test whether this signal association network was not random, observed associations were compared to expected random data, generated using the “permutation of groups within sample” method in SOCPROG (Whitehead, 1999, 2008, 2009; Bejdrrer et al., 1998; Manly, 1995). The “sampling period” was parameterised as the signaller identity, in order to permute data at the individual level (Mielke et al., 2020), and 5000 permutations of 1000 trials were performed. Permutations were done so that the total number and the lengths of sequences were kept constant, while randomising which signal appear in which sequence. To test for preferred associations of signals across individuals (“long-term preferred associations”, Whitehead, 2008), we compared observed and expected association indices’ coefficients of variation (*C.V.*). A one-sided *P*-value was computed corresponding to the

proportion of expected association indices $C.V.$ that were higher than the observed value (significance level: $P \leq 0.025$. Whitehead, 2008). To identify the significant associations of signals, a P -value was computed for each dyad, which corresponded to the proportion of expected association indices that were lower than the observed value (preferred association significance level: $P \geq 0.975$. Whitehead, 2008; Rose & Croft, 2020).

We then computed four different network statistics for each of the studied signal, in order to test for possible differences of association potential between signal types and modalities. For each signal we computed: (i) the degree, which corresponds to the number of different associates; (ii) the strength, which is the sum of all signal's association index (high values thus indicating frequent and/or numerous associations with other signals); (iii) the affinity, which corresponds to the average strength of the signal's associates, weighted by their association indices with the signal (Whitehead, 2008, 2009); and (iv) the clustering coefficient, which represents the extent to which a signal's associates were themselves associated within the signal association network (Holme et al., 2007; Sosa, Sueur, et al., 2020; Whitehead, 2008, 2009).

Finally, we used GEPHI 0.9.2 (Bastian et al., 2009) to represent graphically the network of the most common signal associations we observed. We represented signal associations produced by at least two signallers (discarding idiosyncratic associations), and for which association indices were greater or equal to the third quartile value.

Statistical analyses

Besides the descriptive approach explained above, statistical tests were run to assess the effect of context and signaller characteristics on communicative sequence production (type of sequences produced, characteristics of sequences), and to compare signal types and modalities' integration in the "association network" by comparing their network statistics. We used R v. 4.0.4 (R Core Team, 2021), all tests were two-tailed and alpha-level was set at 0.05. Because of our sample sizes, we performed non-parametrical tests. For all multiple comparisons, P -values were adjusted using "False Discovery Rate" (FDR) method (Noble, 2009).

Firstly, we analysed the effect of social contexts on sequence production. For each sequence group, the frequencies of sequences produced by focal signallers were compared between social contexts (i.e., affiliative, grooming, sexual, playful, aggressive, submissive, and unclear contexts) using Friedman rank sum tests and post-hoc Wilcoxon signed rank tests. These frequencies were weighted based on the general time budget (affiliative context typically representing 2.40% of total time; grooming: 5.82%; sexual: 0.09%; playful: 0.62%; aggressive: 0.53%; submissive: 2.13%; unclear: 4.89%; other –non

social - contexts: 83.51%), so that they corresponded to the number of communicative sequences produced per minute spent in each context. Moreover, we tested the effect of social contexts on sequence characteristics. Because in such analysis the statistical units were the communicative sequences and not the individuals producing them, we used mixed models instead of non-parametrical tests, so that the signaller and receiver identities were taken into account as random effects and the social context was included as a fixed effect. We used Generalised Linear Mixed Models (GLMM) of the Gamma family, Poisson family and a Linear Mixed Model (LMM) to analyse this effect on sequence length, diversity and complexity respectively, as well as GLMM of the Binomial family for multimodality and multicomponentiality (see appendix **Table A1** for details). In order to make all two-by-two comparisons between the different social contexts, each model was run six times making vary the reference level of the fixed effect. Model quality was assessed by graphically verifying the independence of the residuals (Hervé, 2021), as well as checking for the normality of the residuals for the LMM and test for the absence of overdispersion for the GLMMs (Harrison et al., 2018; Hervé, 2021).

Secondly, we evaluated the effect of focal signallers' age and sex on sequence production, i.e., on the number of sequences of each group produced were tested using Mann-Whitney tests.

Finally, we tested whether different types and modalities of signal had different potentialities of being associated with other signals. To do so, we compared signal network statistics (degree, strength, affinity and clustering coefficient) between types and modalities using Mann-Whitney tests.

RESULTS

Communicative signals and sequences

We recorded 2338 signals produced by or in the direction of focal red-capped mangabeys. The cumulative number of described signals at the population-level reached an asymptote (Molesti et al., 2019; Pika et al., 2005) at the 101st session of observation (see appendix **Figure A2**). We identified 45 different signals (**Table 3**), notably completing the previously described repertoire of body and facial signals. Among the 2338 signals we observed, 1699 were part of communicative sequences (69% of the body signals, 88% of the facial signals and 68% of the vocalisations). We identified a total of 424 communicative sequences, produced by 24 different signallers and involving 43 of the 45 different signals of the repertoire. Sequences were composed of up to 23 signals (mean length $\pm S.E.$ = 4.00 ± 0.11), and up to 8 different signals (mean diversity $\pm S.E.$ = 2.50 ± 0.07 ; mean complexity index $\pm S.E.$ = 0.35 ± 0.01). 112 of the 424 sequences were repetitions of one unique signal, 227 were multicomponent (i.e.,

composed of more than one signal type), and 218 were multimodal (i.e., involved more than one sensory modality).

Sequence groups and representative sequences of each group

We grouped the communicative sequences according to their composition, in order to describe the typical communicative sequences of red-capped mangabeys. Based on the dendrogram of the hierarchical clustering of the sequence dissimilarities (**Figure 2**, agglomerative coefficient = 0.97), we found that the best solution was to classify the 424 sequences in 8 groups (represented in **Figure 3**). Representative sequences of each group, selected using “neighbourhood density” criterion, are presented in **Table 4**. For half sequence groups we could extract a set of representatives covering at least 50% of their sequences, however groups 1, 4, 7, 8 were too heterogeneous and we had to choose representatives covering 27.3%, 25%, 25.2% and 26.7% of these groups respectively.

Social contexts of sequence production

Communicative sequences of most different groups were produced by red-capped mangabeys in specific social contexts (**Figure 4**, Friedman rank sum tests: group 1: $X^2_6 = 31.50$, $P < 0.001$; group 4: $X^2_6 = 12.72$, $P = 0.048$; group 5: $X^2_6 = 36.00$, $P < 0.001$; group 7: $X^2_6 = 17.87$, $P = 0.007$; group 8: $X^2_6 = 14.19$, $P = 0.028$), except for groups 2, 3 and 5 for which contexts of production were unclear (group 2: $X^2_6 = 10.59$, $P = 0.102$; group 3: $X^2_6 = 5.56$, $P = 0.474$; group 5: $X^2_6 = 11.66$, $P = 0.070$. See appendix **Table A2** for detailed Wilcoxon signed-rank test results). Main social contexts of the different sequence groups are summarized in **Table 4**.

Moreover, communicative sequence characteristics were dependent on the production context. Sequences produced in playful context were longer than in other contexts (**Figure 5.a**, see appendix **Table A1** for detailed model results) and involved more different signals (**Figure 5.b**), followed by sequences produced in contexts of aggression. Communicative sequences were more complex in playful context (**Figure 5.c**), followed by aggression and submission contexts, and the less complex sequences were produced in sexual context. The proportions of multimodal sequences during play and aggressions were higher than in other contexts (**Figure 5.d**). Indeed, playful sequences were mainly associations of visual and tactile signals or of visual, tactile and audible signals (55 and 27/121 sequences respectively), and aggression sequences were mostly associations of visual signals with audible, tactile signals, or both (21, 20 and 11/91 sequences respectively). On the contrary, communicative sequences produced in sexual, submissive, grooming and affiliative contexts were more unimodal (**Figure 5.d**), with mainly

visual or tactile signals in sexual context (10 and 16/38 sequences), and visual in others (15/22, 27/59 and 27/59 sequences produced in submissive, grooming and affiliative contexts). Sequence multimodality also depended on social contexts, play inducing significantly more multicomponent signal associations (**Figure 5.e.**), i.e. involving body and facial signals or body, facial and vocal signals (76 and 28/121 sequences), while grooming and sexual contexts were mainly unicomponent, involving mostly body signals (41/59 and 29/38 sequences respectively).

Age and sex effect on sequence production for each sequence group

Age and sex of subjects did not have any effect on the overall frequency of communicative sequences produced, however these characteristics affected the production of some types of sequences, as summarised in **Table 4** (see appendix **Table A3** for detailed Mann-Whitney test results).

Signal associations

We found that 312 of the 424 sequences were composed of at least two different signals. They involved 43 different signals of the repertoire, and on the 903 theoretically possible associations, we observed 136 different associations that were produced by at least two signallers (association network density = 15%). The main associations of signals that were observed, i.e. produced by at least 2 different signallers and for which association indices were above the third quartile value (median association index $\pm IQR = 0.13 \pm 0.16$; third quartile = 0.23), involved 23 different signals of the repertoire and are represented in **Figure 6**. The association network differed significantly from what would be expected if signals were randomly associated (“within samples” permutations: observed association $C.V. = 2.017$, mean expected $C.V. = 1.824$, one-sided $P < 0.001$).

Among signal associations produced by at least two signallers, twelve had significantly higher association indices than expected random ones (i.e. one-sided $P > 0.975$): “Grab body part” and “Open mouth” (index = 0.60, $P = 0.991$); “Bite” and “Grab body part” (index = 0.45, $P = 0.984$); “Grab body part” and “Slap receiver” (index = 0.45, $P = 0.996$); “Present body part” and “Present rear” (index = 0.43, $P = 0.980$); “Slap object” and “Un+” (index = 0.39, $P = 1.000$); “Raise eyebrows” and “Slap receiver” (index = 0.33, $P = 0.996$); “Slap object” and “Throw head” (index = 0.31, $P = 0.992$); “Pull body part” and “Push” (index = 0.30, $P = 0.992$); “Grabbing movement” and “Throw arm” (index = 0.29, $P = 0.995$); “Grabbing movement” and “Throw body” (index = 0.27, $P = 0.999$); “Open mouth” and “ND” (index = 0.29, $P = 0.988$); “Throw arm” and “Throw body” (index = 0.21, $P = 0.986$). Moreover, the signals that were associated with the most others were the body signals “grab body part”

(degree (D) = 32, strength (S) = 7.31), “throw arm” (D = 30, S = 5.68), and “present rear” (D = 25, S = 3.56), the facial signal “open mouth” (D = 31, S = 6.15) and the vocal signals “Un+” (D = 27, S = 3.64) and “Ro+” (D = 24, S = 4.4).

We then tested whether signals were more or less likely to be associated with others amongst communicative sequences depending on their types or modalities, by comparing their network measures. We found no significant differences between signal types and modalities regarding these network statistics (degree strength, affinity and clustering coefficient, **Table 5**), except that tactile signals had higher strengths than visual signals (Mann-Whitney test: $N_{tactile} = 8$, $N_{visual} = 26$, $W = 166$, $P = 0.038$) and tended to have higher strengths than audible signals ($N_{audible} = 9$, $W = 15$, $P = 0.073$) and higher degrees than visual ones ($W = 161$, $P = 0.065$).

DISCUSSION

We applied sequential and network analyses to systematically describe the use of multiple signals in the spontaneous communication of captive red-capped mangabeys. A method based on sequence dissimilarity measures permitted to rigorously identify the typical sequences of communicative signals produced by mangabeys, and network analyses permitted to visualize and identify the main signal associations among these sequences. We found that captive mangabeys frequently produce communication signals in sequences, particularly relying on different signal types and modalities in a flexible way. Our results brought to light contextual effects on such complex communication, which gives perspectives to test for current hypotheses on multimodal and multicomponent communication functions. We discuss the benefits of the present methodology to formally describe signal associations with a multimodal and multicomponent approach, as a first step to explore their potential function in communication. Particularly, we suggest that its application to other species and in different environments should permit reliable comparisons on complex signalling.

We adopted an integrative approach to describe the intraspecific communication of twenty-five captive red-capped mangabeys, taking into account the production of body, facial and vocal signals of their repertoire (completing previous description from Aychet et al., 2021; Bouchet et al., 2010; and Schel et al., under review), that were delivered on the visual, audible and/or tactile modality. Given that the cumulative number of described signals reached a maximum at our 101st session of observations, we assume that our analysis included almost all the repertoire of this red-capped mangabey population (Molesti et al., 2019; Pika et al., 2005). However, our observation time was quite low compared to what may be needed to describe the entirety of gestural repertoires at the individual level in other primate species (e.g. Hobaiter & Byrne, 2011b), thus it is possible that some new body signals may be described

through supplementary observations. Moreover, we took into account here three sensory modalities of signals that are the most investigated in primate communication (i.e., visual only, audible and tactile signals. Liebal et al., 2014), but we note that red-capped mangabey communication may even encompass other sensory channels, which would need specific methods to be measured. For instance, although this has rarely been investigated in catarrhine species, “presenting rear” in sexual context may be a way for females to deliver olfactory signals (Dixson, 2015; Liebal et al., 2014; Reichert et al., 2002), and some other mangabey body signals are likely to be perceived on the seismic modality, such as “bang object”, which typically provoke tremors of substrates, in addition of noise.

Applying the same graphical method as in studies on vocal exchanges (Lemasson et al., 2010, 2018; Levréro et al., 2019), we found that consecutive signals produced by a same signaller toward a same receiver were typically initiated within a latency up to 8 seconds, thus we temporally defined signal sequences based on this time-window. We found that the major part of mangabey signal instances was produced in sequences, i.e., 69% of their body signals, 88% of facial ones and 71% of vocalisations, and identified 424 communicative sequences in our sample, most of which were multicomponent (i.e. involved different types of signals) and/or multimodal (i.e. involved different sensory modalities of signals). The fact that such signal use was so frequent in the spontaneous communication of mangabeys may reflect their complementary roles in communication, a key argument for the “multimodal theory” of language origins, which advocates that human language emerged from the co-evolution of gestural, vocal and oro-facial communication in the primate lineage under different ecological and social constraints (e.g. Fröhlich et al., 2019; Levinson & Holler, 2014; Masataka, 2008; Meguerditchian & Vauclair, 2014; Prieur et al., 2019). Moreover, it emphasises the importance of adopting a multimodal and multicomponent approach to describe primate communication, as focussing on a unique type of signal might prevent us to grab the actual complexity of intraspecific communication events (Liebal et al. 2014; Waller et al. 2013). We identified here 8 main categories of signal sequences produced by red-capped mangabeys, and a set of typical “representative” sequences for each of them. In addition to the sequence description, network analyses permitted us to represent the main signal associations in mangabey communication, and identify twelve significant dyadic associations. Therefore, the present analyses permit to refine the description of captive red-capped mangabey communication, by completing single-type signal repertoires (i.e., Aychet et al., 2021; Bouchet et al., 2010; and Schel et al., under review) with a first description of their signal sequence repertoire. This formal characterisation of multiple signal use is the prior step to then investigate the potential biological significance of these events (e.g. typical analytical procedure to examine acoustic sequences: Kershenbaum et al., 2016).

Our second aim here was to describe the social circumstances leading to the production of complex signalling, focussing particularly on multimodal and multicomponent signal associations. We found some contextual effects on sequence production which raises questions that future studies may

address, regarding the potential functions of signal sequences on the information conveyed (i.e., their meaning) or concerning the efficiency of communication (i.e., the probability to receive a satisfactory response from the receiver) (e.g. Guilford & Dawkins, 1991; Heberts & Papaj, 2005; Partan & Marler, 2005; reviewed in Fröhlich & van Schaik, 2018). Some of the typical communicative sequences of red-capped mangabeys were produced under specific social contexts. For instance, the repetition of “Un+” vocalisations or the sequences “Slap object; Un+” and “Throw head; Slap object” were produced in aggression contexts, whereas the repetitions of “Mouth-to-mouth” or the sequences “Mouth-to-mouth; Present rear” were typically produced during affiliative interactions. Moreover, although captive mangabeys of all sex and age categories produced communication signals in sequences, it affected the types of sequence produced. These effects of social context and signaller’s characteristics suggest that some signal sequences may have specific meanings, which may be investigated in future studies by looking at receiver responses (e.g. Cartmill & Byrne, 2010; Hobaiter & Byrne, 2014, 2017). Moreover, it could be examined whether syntactical rules underpin these meanings, i.e., whether sequence meaning differ from the ones of the independent signals of which it is composed, and whether it depends on the sequence structure (Suzuki & Zuberbühler, 2019; Zuberbühler, 2020). In addition, it could be examined whether the multimodality and multicomponentiality of signal sequences have specific functions regarding the information conveyed. In red-capped mangabeys, vocalisations in grooming and socio-sexual contexts have been hypothesized to add informative values to other sensory signals, increasing the efficiency of interactions and decision-making process (Baraud et al., 2016). Thus, despite being rare, multicomponent sequences in these contexts may be in favour of the “multiple message hypothesis”, which propose that associated signals from multiple types and modalities convey different information that complement each other’s (Fröhlich & van Schaik, 2018; Johnstone, 1996; Moller & Pomiankowski, 1993; Partan & Marler, 1999). This is the case in human multimodal communication, in which meaning can arise from the integration of signals produced on different channels (Vigliocco et al., 2014). In contrast, it could be hypothesised that the signals produced within sequences all have the same meanings (“redundant signal hypothesis”, Moller & Pomiankowski, 1993; Partan & Marler, 1999), and that sequential use actually reflect communication persistence or elaboration in case of first-attempt failure (e.g. in apes: Byrne et al., 2017; Cartmill & Byrne, 2007; Genty & Byrne, 2010; Leavens et al., 2005; Liebal et al., 2004b; Tomasello & Call, 2018).

Aside from the context-specificity of some sequence types, we found that different social contexts affected the characteristics of communicative sequences. Particularly, playful contexts elicited longer, more diverse and complex communicative sequences, and favoured both multimodal and multicomponent communication. Play elicited a wide use of body and facial signal associations, produced on the visual and tactile modalities. Playful interactions are known to elicit a wider use of communication signals in primates than other contexts, with more diverse signals (e.g. gesture use in apes: Call & Tomasello, 2007; Liebal et al., 2006; Pika et al., 2003), production of idiosyncratic gestures

(e.g. in gorillas: Pika et al., 2003) and frequent use of signal associations (e.g. in orangutans and chimpanzees: Liebal et al., 2004b, 2006). Thus, more complex signalling in red-capped mangabeys during play compared to other contexts, particularly for juveniles, corroborates theories proposing social play as a way to develop behavioural flexibility, notably by exploring and exercising communicative actions (e.g. Fagen, 1982; Tomasello et al., 1989). The length of communicative sequences in play and aggression events may also result from the high level of arousal in those contexts, compared to affiliation or grooming, as hypothesized for orangutans gestural sequences (Tempelmann & Liebal, 2012). Another explanation for the observed communication complexity and multimodal/multicomponent use of signals could be a “meaning refinement” function of signal sequences, one signal permitting the first one to be accurately understood by the receiver (“refinement” hypothesis, Fröhlich & van Schaik, 2018; as suggested for vocal-gesture combinations in wild bonobos and chimpanzees: Genty et al., 2014; Wilke et al., 2017). The high proportion of body and facial signal associations during play might reflect the use of “open mouth” having the role of metacommunicative signals (Altmann, 1967; Bateson, 1955; Bekoff, 1975), to disambiguate the meaning of agonistic-like gestures (Demuru et al., 2015; Palagi, 2008; Waller & Cherry, 2012; Waller & Dunbar, 2005). This hypothesis is also supported by the fact that tactile signals had higher strengths than visual ones in the association network. Investigation at the signal-level, using the same approach as proposed by Genty (2019), should permit to test for the “refinement” hypothesis, comparing receiver response to the body signals of red-capped mangabey repertoire that are the most associated to others (e.g. “Grab body part” or “Throw arm”), when produced alone or in sequences with different other signals. Moreover, “match-to-sample” paradigm may permit to test for this hypothesis with an experimental approach, looking at whether mangabeys associate single or sequential use of these body signals to different social interactions between a signaller and a receiver (Waller et al., 2016).

Furthermore, we believe that the methodology we used here may represent a useful tool to reliably test for hypotheses about the evolutionary constraints which shaped the complexity of primate communication, by facilitating reliable comparisons between different species and populations (Fröhlich & van Schaik, 2018). Indeed, it has been hypothesised that environmental factors may influence the use of multimodal signalling, which seems particularly appropriate for short-distance communication and when animals live in dense habitats (see “efficacy-backup hypothesis”, Fröhlich & van Schaik, 2018; Heberts & Papaj, 2005; Partan & Marler, 2005), or that captivity could either favour or constrain this complex communication (see for review: Fröhlich & van Schaik, 2018). Moreover, species characteristics such as arborealism (Call & Tomasello, 2007; Lemasson, 2011; Liebal et al., 2004) or sociality (“social complexity hypothesis”: Dunbar, 2009; Freeberg et al., 2012) may induce more or less frequent use of multimodal and multicomponent signalling. One first aspect that could permit to enhance the reliability of these future comparisons is the temporal definition used to identify signal associations. Instead of using different arbitrary criterion, looking at signal sequences basing the temporal definition

on observed between-signal latencies allow to both use a relevant analysis time-window in regard of the species biology, while using the same method from one study to another. Thus, the frequencies of so-defined signal sequences could be compared between different species, for instance between primates with different social styles to test whether species organised in larger groups, with more complex social interactions, might be more likely to use multiple signals than others (Dunbar, 2009; Freeberg et al., 2012). Moreover, in the same way as repertoire sizes and variability can be compared between species to test for the “social complexity hypothesis” (e.g. Bouchet et al., 2013; Dobson, 2012), sequence clustering and network analysis could permit to compare the number of typical sequences identified in different species, the number of significant signal associations, as well as the variability in the use of shared signals (e.g. gestures shared by several ape species, Byrne et al., 2017). In addition to this gain in description objectivity, the sequence index plots and the use of network diagrams permit explicit visual representation of communicative sequence variability and of common signal associations. These methods have also the advantage of being widely available, such as the the toolbox {TraMineR} on R (Gabadinho et al., 2011). Regarding network analyses, numerous free softwares exist, such as SOCPROG 2.9 software that we used (Whitehead, 2009), and a new package was recently developed in R permitting both statistical measures and network visualisation (Sosa, Puga-Gonzalez, et al., 2020), which could even facilitate such approach.

Several perspectives may be considered to further improve the present descriptive method, starting with the way signal themselves are described. Although the acoustic measures on which vocalisation repertoires are based allow reliable comparisons between studies, the accuracy with which discrete body and facial signals are morphologically described in primates may vary from one study to another (Liebal et al., 2014), as well as their criteria of definition (Bourjade et al., 2020). The objectivity of facial signal description could be improved by the use of FACS (Ekman et al., 1997), as already developed in several primate clades (i.e., chimpanzees, gibbons, orangutans and macaques: Parr et al., 2007; Waller et al., 2012; Caeiro et al., 2013; Julle-Danière et al., 2015). To our knowledge, no such anatomically-based tool yet exist to describe body signals, although some studies explored more objective approaches to describe primate gestures (e.g. in chimpanzees: Roberts et al., 2012). The present analysis of signal associations may benefit from methodological developments in this area, by taking into account signal of all types and modalities that would have been described on more objective, structure-based methods. Finally, another dimension on which the present approach could be improved would be the analysis of the potential overlapping of signals among the sequences, i.e. signal combinations. This supplementary description would be a valuable advance for the objective study of primate complex communication, given that the simultaneous production or perception of signals on multiple modalities could have different functions at the content or efficacy-level of communication (e.g. Guilford & Dawkins, 1991; Hebets & Papaj, 2005; Partan & Marler, 2005; reviewed in Fröhlich & van Schaik, 2018).

Conclusion

Captive red-capped mangabeys frequently produce their communicative signals in multimodal and/or multicomponent sequences, under specific contextual conditions. Completing previous findings from studies focussing on unique signal types, our results highlight the need for a multimodal and multicomponent approach to understand the complexity of mangabey communication. The study of primate multimodal and multicomponent communication is at its beginning, and we believe that the present methodology would enhance future comparisons between species, regarding their use of multiple signals in different contexts, particularly with the perspective of investigating the evolutionary roots of complex communication and human language.

REFERENCES

- Abbott, A., & Forrest, J. (1986). Optimal Matching Methods for Historical Sequences. *Journal of Interdisciplinary History*, 16(3), 471. <https://doi.org/10.2307/204500>
- Abbott, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology: *Review and Prospect. Sociological Methods & Research*, 29(1), 3–33. <https://doi.org/10.1177/0049124100029001001>
- Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, 49(3), 227–266. <https://doi-org.passerelle.univ-rennes1.fr/10.1163/156853974X00534>
- Altmann, S. A. (1967). The structure of primate social communication, Social communication among primates. In *Social communication among primates* (University of Chicago Press, pp. 325–362).
- Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate Vocalization, Gesture, and the Evolution of Human Language. *Current Anthropology*, 49(6), 1053–1076. <https://doi.org/10.1086/593015>
- Aychet, J., Blois-Heulin, C., Palagi, E., Lemasson, A. (2021). Facial displays in red-capped mangabeys (*Cercocebus torquatus*): repertoire, social context and potential intentionality. *Journal of Comparative Psychology*, 135(1), 98–113. <https://doi.org/10.1037/com0000252>
- Baraud, I., Deputte, B. L., Pierre, J.-S., & Blois-Heulin, C. (2016). Informative Value of Vocalizations during Multimodal Interactions in Red-Capped Mangabeys. In *Discovering Hidden Temporal Patterns in Behavior and Interaction* (pp. 255–277). Springer, New York, NY. https://doi.org/10.1007/978-1-4939-3249-8_14
- Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: an open source software for exploring and manipulating networks*. International AAAI Conference on weblogs and social media.
- Bateson, G. (1955). A theory of play and fantasy. *Psychiat. Res. Rept. A.*, 2, 39–51.
- Bekoff, M. (1975). The communication of play intention: are play signals functional? *Semiotica*, 15(3). <https://doi.org/10.1515/semi.1975.15.3.231>
- Bejder, L., Fletcher, D., & Bräger, S. (1998). A method for testing association patterns of social animals. *Animal Behaviour*, 56(3), 719–725. <https://doi.org/10.1006/anbe.1998.0802>
- Bouchet, H., Pellier, A.-S., Blois-Heulin, C., & Lemasson, A. (2010). Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): a multi-level acoustic analysis. *American Journal of Primatology*, 72(4), 360–375. <https://doi.org/10.1002/ajp.20791>

- Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, 4, 390. <https://doi.org/10.3389/fpsyg.2013.00390>
- Bourjade, M., Cochet, H., Molesti, S., & Guidetti, M. (2020). Is Conceptual Diversity an Advantage for Scientific Inquiry? A Case Study on the Concept of 'Gesture' in Comparative Psychology. *Integrative Psychological and Behavioral Science*, 54(4), 805–832. <https://doi.org/10.1007/s12124-020-09516-5>
- Briseño-Jaramillo, M., Estrada, A., & Lemasson, A. (2015). Behavioural innovation and cultural transmission of communication signal in black howler monkeys. *Scientific Reports*, 5(1), 1–10. <https://doi.org/10.1038/srep13400>
- Byrne, R. W., Cartmill, E. A., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. E. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition*, 20, 755–769. <https://doi.org/10.1007/s10071-017-1127-1>
- Cacero, C. C., Waller, B. M., Zimmermann, E., Burrows, A. M., & Davila-Ross, M. (2013). OrangFACS: A Muscle-Based Facial Movement Coding System for Orangutans (*Pongo* spp.). *International Journal of Primatology*, 34(1), 115–129. <https://doi.org/10.1007/s10764-012-9652-x>
- Call, J., & Tomasello, M. (Eds.). (2007). *The gestural communication of apes and monkeys*. Psychology Press, New York, NY.
- Cartmill, E. A., & Byrne, R. W. (2007). Orangutans Modify Their Gestural Signaling According to Their Audience's Comprehension. *Current Biology*, 17(15), 1345–1348. <https://doi.org/10.1016/j.cub.2007.06.069>
- Cartmill, E. A., & Byrne, R. W. (2010). Semantics of primate gestures: Intentional meanings of orangutan gestures. *Animal Cognition*, 13(6), 793–804. <https://doi.org/10.1007/s10071-010-0328-7>
- Cassell, J., McNeill, D., & McCullough, K.-E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & Cognition*, 7(1), 1–34. <https://doi.org/10.1075/pc.7.1.03cas>
- Chalmers, N. R. (1968). Group composition, ecology and daily activities of free living mangabeys in Uganda. *Folia Primatologica*, 8, 247–262. <https://doi.org/10.1159/000155157>
- Cooke, C. A. (2012). *The feeding, ranging and positional behaviors of Cercopithecus torquatus, the red-capped mangabey, in Sette Cama Gabon: a phylogenetic perspective (Doctoral dissertation)*. Ohio State University.
- Cooper, B. G., & Goller, F. (2004). Multimodal Signals: Enhancement and Constraint of Song Motor Patterns by Visual Display. *Science*, 303(5657), 544–546. <https://doi.org/10.1126/science.1091099>
- Demuru, E., Ferrari, P. F., & Palagi, E. (2015). Emotionality and intentionality in bonobo playful communication. *Animal Cognition*, 18(1), 333–344. <https://doi.org/10.1007/s10071-014-0804-6>
- Deputte, B. L. (1992). Life history of captive gray-cheeked mangabeys: physical and sexual development. *International Journal of Primatology*, 13(5), 509–531. <https://doi.org/10.1007/BF02547830>
- Deslandes, V., Faria, L. R. R., Borges, M. E., & Pie, M. R. (2014). The structure of an avian syllable syntax network. *Behavioural Processes*, 106, 53–59. <https://doi.org/10.1016/j.beproc.2014.04.010>
- Dixson, A. (2015). Primate sexuality. In Bolin, A., Whelehan, P. (Eds.), *The International Encyclopedia of Human Sexuality* (pp. 861–1042). John Wiley & Sons, Ltd, Oxford, UK. <https://doi.org/10.1002/9781118896877.wbiehs375>
- Dobson, S. D. (2012). Coevolution of Facial Expression and Social Tolerance in Macaques: Facial Expression and Social Tolerance. *American Journal of Primatology*, 74(3), 229–235. <https://doi.org/10.1002/ajp.21991>
- Dolado, R., & Beltran, F. S. (2012). Emergent patterns of social organization in captive *Cercopithecus torquatus*: Testing the GrooFiWorld agent-based model. *Journal of Biosciences*, 37(4), 777–784. <https://doi.org/10.1007/s12038-012-9231-5>
- Dolado, R., & Beltran, F. S. (2011). Dominance hierarchy and spatial distribution in captive red-capped mangabeys (*Cercopithecus torquatus torquatus*): Testing

- Hemelrijk's agent-based model. *Interaction Studies*, 12(3), 461–473. <https://doi.org/10.1075/is.12.3.05dol>
- Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36(5), 562–572. <https://doi.org/10.1080/03014460902960289>
- Ekman, P. (1979). About brows: emotional and conversational signals. In M. von Cranach, F. Koppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: claims and limits of a new discipline*. Cambridge University Press, New York, NY.
- Ekman, P. (2004). Emotional and Conversational Nonverbal Signals. In J. M. Larrazabal & L. A. P. Miranda (Eds.), *Language, Knowledge, and Representation* (pp. 39–50). Springer Netherlands, Dordrecht, Netherlands. https://doi.org/10.1007/978-1-4020-2783-3_3
- Ekman, P., D. P. of P. P. E., PH, Rosenberg, E. L., & Rosenberg, L. D. of P. E. L. (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press, New York, NY.
- Elias, D. O., Maddison, W. P., Peckmezian, C., Girard, M. B., & Mason, A. C. (2012). Orchestrating the score: Complex multimodal courtship in the *Habronattus coecatus* group of *Habronattus* jumping spiders (Araneae: Salticidae). *Biological Journal of the Linnean Society*, 105(3), 522–547. <https://doi.org/10.1111/j.1095-8312.2011.01817.x>
- Fagen, R. (1982). Evolutionary Issues in Development of Behavioral Flexibility. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Ontogeny* (pp. 365–383). Springer US, Boston, MA. https://doi.org/10.1007/978-1-4615-7578-8_8
- Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1785–1801. <https://doi.org/10.1098/rstb.2011.0213>
- Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11), 1325–1330. <https://doi.org/10.1111/2041-210X.12584>
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, 6(1). <https://doi.org/10.1038/srep25887>
- Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & van Schaik, C. P. (2019). Multimodal communication and language origins: integrating gestures and vocalizations. *Biological Reviews*, 0(0). <https://doi.org/10.1111/brv.12535>
- Fröhlich, M., & van Schaik, C. P. (2018). The function of primate multimodal communication. *Animal Cognition*, 21(5), 619–629. <https://doi.org/10.1007/s10071-018-1197-8>
- Gabadinho, A., & Ritschard, G. (2013). Searching for typical life trajectories applied to childbirth histories. In R. Lévy & E. Widmer (Eds.), *Gendered life courses* (pp. 287–312).
- Gabadinho, A., Ritschard, G., Mueller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37. <https://doi.org/10.18637/jss.v040.i04>
- Gabadinho, A., Ritschard, G., & Studer, M. (2010). *Indice de complexité pour le tri et la comparaison de séquences catégorielles*. Extraction et gestion des connaissances (EGC'2010), Hammamet, Tunisie.
- Garland, E. C., Lilley, M., Goldizen, A., Rekdahl, M., Garrigue, C., & Noad, M. (2012). Improved versions of the Levenshtein distance method for comparing sequence information in animals' vocalisations: Tests using humpback whale song. *Behaviour*, 149, 1413–1441. <https://doi.org/10.2307/41720621>
- Garland, E. C., Noad, M., Goldizen, A., Lilley, M. S., Rekdahl, M. L., Garrigue, C., Constantine, R., Hauser, N. D., Poole, M., & Robbins, J. (2013). Quantifying humpback whale song sequences to understand the dynamics of song exchange at the ocean basin scale. *The Journal of the Acoustical Society of America*. <https://doi.org/10.1121/1.4770232>
- Gautier-Hion, A., Colyn, M., & Gautier, J.-P. (1999). *Histoire naturelle des primates d'Afrique centrale*. ECOFAC, Libreville, Gabon.

- Gautier-Hion, A., & Gautier, J.-P. (1976). Croissance, maturité sexuelle et sociale, reproduction chez les cercopithécins forestiers africains. *Folia Primatologica*, 26, 165–184. <https://doi.org/doi:10.1159/000155749>
- Genty, E. (2019). Vocal-gestural combinations in infants bonobos: new insights into signal functional specificity. *Animal Cognition*, 1–14. <https://doi.org/10.1007/s10071-019-01267-0>
- Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (Gorilla gorilla): Repertoire, intentionality and possible origins. *Animal Cognition*, 12(3), 527–546. <https://doi.org/10.1007/s10071-009-0213-4>
- Genty, E., & Byrne, R. W. (2010). Why do gorillas make sequences of gestures? *Animal Cognition*, 13(2), 287–301. <https://doi.org/10.1007/s10071-009-0266-4>
- Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-Modal Use of a Socially Directed Call in Bonobos. *PLOS ONE*, 9(1), e84738. <https://doi.org/10.1371/journal.pone.0084738>
- Genty, E., Neumann, C., & Zuberbühler, K. (2015). Complex patterns of signalling to convey different social goals of sex in bonobos, *Pan paniscus*. *Scientific Reports*, 5(1), 1–13. <https://doi.org/10.1038/srep16135>
- Ghazanfar, A. A., & Logothetis, N. K. (2003). Facial expressions linked to monkey calls. *Nature*, 423(6943), 937–938. <https://doi.org/10.1038/423937a>
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *Journal of Neuroscience*, 25(20), 5004–5012. <https://doi.org/10.1523/JNEUROSCI.0799-05.2005>
- Ginsberg, J. R., & Young, T. P. (1992). Measuring association between individuals or groups in behavioural studies. *Animal Behaviour*, 44, 377–379. [https://doi.org/10.1016/0003-3472\(92\)90042-8](https://doi.org/10.1016/0003-3472(92)90042-8)
- Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M., & Crockford, C. (2021). Chimpanzees use numerous flexible vocal sequences with more than two vocal units: A step towards language? (p. 429517). Cold Spring Harbor Laboratory. <https://doi.org/10.1101/2021.02.03.429517> (Preprint)
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3(11), 419–429. [https://doi.org/10.1016/S1364-6613\(99\)01397-2](https://doi.org/10.1016/S1364-6613(99)01397-2)
- Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing Knowledge Through Gesture: Using Children’s Hands to Read Their Minds. *Cognition and Instruction*, 9(3), 201–219. https://doi.org/10.1207/s1532690xc0903_2
- Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, 42(1), 1–14. [https://doi.org/10.1016/S0003-3472\(05\)80600-1](https://doi.org/10.1016/S0003-3472(05)80600-1)
- Gupta, S., & Sinha, A. (2019). Gestural communication of wild bonnet macaques in the Bandipur National Park, Southern India. *Behavioural Processes*, 168, 103956. <https://doi.org/10.1016/j.beproc.2019.103956>
- Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, 6, e4794. <https://doi.org/10.7717/peerj.4794>
- Hebets, E. A., & Papaj, D. R. (2005). Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*, 57(3), 197–214. <https://doi.org/10.1007/s00265-004-0865-7>
- Hervé, M. (2021). *RVAideMemoire: Testing and plotting procedures for Biostatistics* (R Package v. 0.9-79).
- Higham, J. P., & Hebets, E. A. (2013). An introduction to multimodal communication. *Behavioral Ecology and Sociobiology*, 67(9), 1381–1388. <https://doi.org/10.1007/s00265-013-1590-x>
- Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees’ use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, 71(6), 96. <https://doi.org/10.1007/s00265-017-2325-1>

- Hobaiter, C., & Byrne, R. W. (2011a). Serial gesturing by wild chimpanzees: its nature and function for communication. *Animal Cognition*, 14(6), 827–838. <https://doi.org/10.1007/s10071-011-0416-3>
- Hobaiter, C., & Byrne, R. W. (2011b). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, 14(5), 745–767. <https://doi.org/10.1007/s10071-011-0409-2>
- Hobaiter, C., & Byrne, R. W. (2014). The Meanings of Chimpanzee Gestures. *Current Biology*, 24(14), 1596–1600. <https://doi.org/10.1016/j.cub.2014.05.066>
- Hobaiter, C., & Byrne, R. W. (2017). What is a gesture? A meaning-based approach to defining gestural repertoires. *Neuroscience & Biobehavioral Reviews*, 82, 3–12. <https://doi.org/10.1016/j.neubiorev.2017.03.008>
- Holle, H., & Gunter, T. C. (2007). The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience*, 19(7), 1175–1192. <https://doi.org/10.1162/jocn.2007.19.7.1175>
- Holme, P., Park, S. M., Kim, B. J., & Edling, C. R. (2007). Korean university life in a network perspective: Dynamics of a large affiliation network. *Physica A: Statistical Mechanics and Its Applications*, 373, 821–830. <https://doi.org/10.1016/j.physa.2006.04.066>
- Hutchison, R. E., Hutchison, J. B., & Fusani, L. (1997). Vocal-Postural Co-Ordination of a Sexually Dimorphic Display in a Monomorphic Species: the Barbary Dove. *Behaviour*, 134(5–6), 321–335. <https://doi.org/10.1163/156853997X00566>
- Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, 396(6708), 228–228. <https://doi.org/10.1038/24300>
- Izumi, A., & Kojima, S. (2004). Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Animal Cognition*, 7(3), 179–184. <https://doi.org/10.1007/s10071-004-0212-4>
- Johnstone, R. A. (1996). Multiple displays in animal communication: ‘backup signals’ and ‘multiple messages.’ *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1337), 329–338. <https://doi.org/10.1098/rstb.1996.0026>
- Jones, C. B., & van Cantfort, T. E. V. (2007). Multimodal Communication by Male Mantled Howler Monkeys (*Alouatta palliata*) in Sexual Contexts: A Descriptive Analysis. *Folia Primatologica*, 78(3), 166–185. <https://doi.org/10.1159/000099138>
- Jones, C., & Sabater-Pi, J. (1968). Comparative ecology of *Cercocebus albigena* and *Cercocebus torquatus* in Rio Muni, West Africa. *Folia Primatologica*, 9, 99–113.
- Julle-Danière, É., Micheletta, J., Whitehouse, J., Joly, M., Gass, C., Burrows, A. M., & Waller, B. M. (2015). MaqFACS (Macaque Facial Action Coding System) can be used to document facial movements in Barbary macaques (*Macaca sylvanus*). *PeerJ*, 3, e1248. <https://doi.org/10.7717/peerj.1248>
- Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akcay, C., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., DeRuiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T. M., Garland, E. C., Gustison, M., Harley, H. E., ... Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews of the Cambridge Philosophical Society*, 91, 13–52. <https://doi.org/10.1111/brv.12160>
- Kershenbaum, A., & Garland, E. C. (2015). Quantifying similarity in animal vocal sequences: Which metric performs best? *Methods in Ecology and Evolution*, 6, 1452–1461. <https://doi.org/10.1111/2041-210X.12433>
- Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1740), 2974–2981. <https://doi.org/10.1098/rspb.2012.0322>
- Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J. F., & Morrogh-Bernard, H. (2019). Gesture use in communication between mothers and offspring in wild orangutans (*Pongo pygmaeus wurmbii*) from the Sabangau Peat-Swamp forest, Borneo. *International Journal of Primatology*. <https://doi.org/10.1007/s10764-019-00095-w>
- Krull, C. R., Ranjard, L., Landers, T. J., Ismar, S. M. H., Matthews, J. L., & Hauber, M. E. (2012). Analyses of sex and individual differences in vocalizations of Australasian gannets using a dynamic time warping algorithm. *The Journal of the Acoustical Society of*

- America, 132(2), 1189–1198. <https://doi.org/10.1121/1.4734237>
- Leavens, D. A., & Hopkins, W. D. (1998). Intentional Communication by Chimpanzees: A Cross-Sectional Study of the Use of Referential Gestures. *Dev Psychol.*, 34(5), 813–822.
- Leavens, D. A., Hostetter, A. B., Wesley, M. J., & Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Animal Behaviour*, 67(3), 467–476. <https://doi.org/10.1016/j.anbehav.2003.04.007>
- Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Development*, 76(1), 291–306. <https://doi.org/10.1111/j.1467-8624.2005.00845.x>
- Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2010). Multimodal communication by captive chimpanzees (*Pan troglodytes*). *Animal Cognition*, 13(1), 33–40. <https://doi.org/10.1007/s10071-009-0242-z>
- Lemasson, A. (2011). What can forest guenons « tell » us about the origin of language? In A. Vilain, J.-L. Schwartz, C. Abry, & J. Vauclair (Eds.), *Primate Communication and Human Language: Vocalisation, gestures, imitation and deixis in humans and non-humans*. (pp. 39-70). John Benjamins Publishing Company., Amsterdam, Netherlands / Philadelphia, PA.
- Lemasson, A., Gandon, E., & Hausberger, M. (2010). Attention to elders' voice in non-human primates. *Biology Letters*, 6(3), 325–328. <https://doi.org/10.1098/rsbl.2009.0875>
- Lemasson, A., Pereira, H., & Levréro, F. (2018). Social basis of vocal interactions in western lowland gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology*, 132(2), 141–151. <https://doi.org/10.1037/com0000105>
- Levenshtein, V. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady*, 10, 707–710.
- Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0302>
- Levréro, F., Touitou, S., Frédet, J., Nairaud, B., Guéry, J.-P., & Lemasson, A. (2019). Social bonding drives vocal exchanges in Bonobos. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-018-36024-9>
- Li, D., & Becchi, M. (2012). Multiple Pairwise Sequence Alignments with the Needleman-Wunsch Algorithm on GPU. *2012 SC Companion: High Performance Computing, Networking Storage and Analysis*, 1471–1472. <https://doi.org/10.1109/SC.Companion.2012.267>
- Liebal, K., Pika, S., & Tomasello, M. (2004a). Social communication in siamangs (*Symphalangus syndactylus*): use of gestures and facial expressions. *Primates*, 45(1), 41–57. <https://doi.org/10.1007/s10329-003-0063-7>
- Liebal, K., Call, J., & Tomasello, M. (2004b). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, 64(4), 377–396. <https://doi.org/10.1002/ajp.20087>
- Liebal, K., Pika, S., & Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture*, 6(1), 1–38. <https://doi.org/10.1075/gest.6.1.02lie>
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014a). *Primate communication: a multimodal approach*. Cambridge University Press.
- Manly, B. F. J. (1995). A Note on the Analysis of Species Co-Occurrences. *Ecology*, 76(4), 1109–1115. <https://doi.org/10.2307/1940919>
- Masataka, N. (2008). The Gestural Theory and the Vocal Theory of Language Origins Are Not Incompatible with One Another. In Masataka, N. (Ed.), *The Origins of Language* (pp. 1–10). Springer, Tokyo, Japan. https://doi.org/10.1007/978-4-431-79102-7_1
- Meguerditchian, A., & Vauclair, J. (2014). Communicative Signaling, Lateralization and Brain Substrate in Nonhuman Primates: Toward a Gestural or a Multimodal Origin of Language? *Humana Mente. Journal of Philosophical Studies*, 27, 135–160.
- Micheletta, J., Engelhardt, A., Matthews, L., Agil, M., & Waller, B. M. (2013). Multicomponent and

- Multimodal Lipsmacking in Crested Macaques (*Macaca nigra*). *American Journal of Primatology*, 75(7), 763–773. <https://doi.org/10.1002/ajp.22105>
- Mielke, A., Waller, B., Perez, C. J., Duboscq, J., & Micheletta, J. (2020). NetFACS: Using network science to understand facial communication systems. *PsyArXiv*. <https://doi.org/10.31234/osf.io/4vghk> (Preprint)
- Mitani, J. C., Call, J., Kappeler, P. M., Palombit, R. A., & Silk, J. B. (Eds.). (2012). *The evolution of primate societies*. The University of Chicago Press, Chicago, IL.
- Molesti, S., Meguerditchian, A., & Bourjade, M. (2019). Gestural communication in olive baboons (*Papio anubis*): Repertoire and intentionality. *Animal Cognition*. <https://doi.org/10.1007/s10071-019-01312-y>
- Moller, A. P., & Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology*, 32(3), 167–176. <https://doi.org/10.1007/BF00173774>
- Needleman, S. B., & Wunsch, C. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48, 443–453.
- Noble, W. S. (2009). How does multiple testing correction work? *Nature Biotechnology*, 27(12), 1135–1137. <https://doi.org/10.1038/nbt1209-1135>
- Orimaye, O. J. (2017). Density and Abundance of the Red-Capped Mangabey (*Cercocebus Torquatus*) In Omo Biosphere Reserve and Idanre Forest Reserve, South Western Nigeria. *MOJ Proteomics & Bioinformatics*, 5(2). <https://doi.org/10.15406/mojpb.2017.05.00156>
- Özyürek, A., & Woll, B. (2019). Language in the Visual Modality: Co- speech Gesture and Sign Language. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 67–83). MIT Press, Cambridge, MA.
- Palagi, E. (2008). Sharing the motivation to play: the use of signals in adult bonobos. *Animal Behaviour*, 75(3), 887–896. <https://doi.org/10.1016/j.anbehav.2007.07.016>
- Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion*, 7(1), 172–181. <https://doi.org/10.1037/1528-3542.7.1.172>
- Partan, S., & Marler, P. (1999). Communication Goes Multimodal. *Science*, 283(5406), 1272–1273. <https://doi.org/10.1126/science.283.5406.1272>
- Partan, S. R., & Marler, P. (2005). Issues in the Classification of Multimodal Communication Signals. *The American Naturalist*, 166(2), 231–245. <https://doi.org/10.1086/431246>
- Partan, S. R. (2002). Single and multichannel signal composition: facial expressions and vocalizations of rhesus macaques (*Macaca mulatta*). *Behaviour*, 139(8), 993–1027.
- Partan, S. R., Fulmer, A. G., Gounard, M. A. M., & Redmond, J. E. (2010). Multimodal alarm behavior in urban and rural gray squirrels studied by means of observation and a mechanical robot. *Current Zoology*, 56(3), 313–326. <https://doi.org/10.1093/czoolo/56.3.313>
- Partan, S. R., & Marler, P. (2005). Issues in the Classification of Multimodal Communication Signals. *The American Naturalist*, 166(2), 231–245. <https://doi.org/10.1086/431246>
- Pika, S. (2008). Gestures of apes and pre-linguistic human children: Similar or different? *First Language*, 28(2), 116–140. <https://doi.org/10.1177/0142723707080966>
- Pika, S., Liebal, K., Call, J., & Tomasello, M. (2005). The Gestural Communication of Apes. *Gesture*, v.5, 41–56 (2005), 5. <https://doi.org/10.1075/gest.5.1.05pik>
- Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, learning, and use. *American Journal of Primatology*, 60(3), 95–111. <https://doi.org/10.1002/ajp.10097>
- Pika, S., Liebal, K., & Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): Repertoire and use. *American Journal of Primatology*, 65(1), 39–61. <https://doi.org/10.1002/ajp.20096>
- Pollick, A. S., & De Waal, F. B. (2007). Ape gestures and language evolution. *Proceedings of the National*

- Academy of Sciences*, 104(19), 8184–8189. <https://doi.org/10.1073/pnas.0702624104>
- Pollick, A. S., Jeneson, A., & de Waal, F. B. M. (2008). Gestures and Multimodal Signaling in Bonobos. In T. Furuichi & J. Thompson (Eds.), *The Bonobos* (pp. 75–94). Springer New York, NY. https://doi.org/10.1007/978-0-387-74787-3_5
- Prieur, J., Barbu, S., Blois-Heulin, C., & Lemasson, A. (2019). The origins of gestures and language: history, current advances and proposed theories. *Biological Reviews*, n/a(n/a). <https://doi.org/10.1111/brev.12576>
- Prieur, J., Lemasson, A., Barbu, S., & Blois-Heulin, C. (2018). Challenges Facing the Study of the Evolutionary Origins of Human Right-Handedness and Language. *International Journal of Primatology*. <https://doi.org/10.1007/s10764-018-0038-6>
- R Core Team. (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. <http://www.R-project.org/>
- Reichert, K. E., Heistermann, M., Hodges, J. K., Boesch, C., & Hohmann, G. (2002). What Females Tell Males About Their Reproductive Status: Are Morphological and Behavioural Cues Reliable Signals of Ovulation in Bonobos (*Pan paniscus*)? *Ethology*, 108(7), 583–600. <https://doi.org/10.1046/j.1439-0310.2002.00798.x>
- Roberts, A. I., Roberts, S. G. B., & Vick, S.-J. (2014). The repertoire and intentionality of gestural communication in wild chimpanzees. *Animal Cognition*, 17(2), 317–336. <https://doi.org/10.1007/s10071-013-0664-5>
- Roberts, A. I., Vick, S.-J., Roberts, S. G. B., Buchanan-Smith, H. M., & Zuberbühler, K. (2012). A structure-based repertoire of manual gestures in wild chimpanzees: Statistical analyses of a graded communication system. *Evolution and Human Behavior*, 33(5), 578–589. <https://doi.org/10.1016/j.evolhumbehav.2012.05.006>
- Rose, P. E., & Croft, D. P. (2020). Evaluating the social networks of four flocks of captive flamingos over a five-year period: Temporal, environmental, group and health influences on assortment. *Behavioural Processes*, 175, 104118. <https://doi.org/10.1016/j.beproc.2020.104118>
- Sasahara, K., Cody, M. L., Cohen, D., & Taylor, C. E. (2012). Structural Design Principles of Complex Bird Songs: A Network-Based Approach. *PLOS ONE*, 7(9), e44436. <https://doi.org/10.1371/journal.pone.0044436>
- Scott-Phillips, T. C. (2015). Nonhuman Primate Communication, Pragmatics, and the Origins of Language. *Current Anthropology*, 56(1), 56–80. <https://doi.org/10.1086/679674>
- Schel, A. M., Bono, A., Aychet, J., Pika, S., Lemasson, A. (under review) Intentional gesturing in red-capped mangabeys (*Cercocebus torquatus*).
- Sievers, C., & Gruber, T. (2020). Can nonhuman primate signals be arbitrarily meaningful like human words? An affective approach. *Animal Behavior and Cognition*, 7(2), 140–150. <https://doi.org/10.26451/abc.07.02.08.2020>
- Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: the need for multimodality in primate communication research. *Animal Behaviour*, 81(5), 919–924. <https://doi.org/10.1016/j.anbehav.2011.02.002>
- Sosa, S., Puga-Gonzalez, I., Hu, F., Pansanel, J., Xie, X., & Sueur, C. (2020). A multilevel statistical toolkit to study animal social networks: the Animal Network Toolkit Software (ANTs) R package. *Scientific Reports*, 10(1), 1–8. <https://doi.org/10.1038/s41598-020-69265-8>
- Sosa, S., Sueur, C., & Puga-Gonzalez, I. (2020). Network measures in animal social network analysis: Their strengths, limits, interpretations and uses. *Methods in Ecology and Evolution*, n/a(n/a). <https://doi.org/10.1111/2041-210X.13366>
- Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2), 481–511. <https://doi.org/10.1111/rssa.12125>
- Suzuki, T. N., & Zuberbühler, K. (2019). Animal syntax. *Current Biology*, 29(14), R669–R671. <https://doi.org/10.1016/j.cub.2019.05.045>
- Tagliatela, J. P., Russell, J. L., Pope, S. M., Morton, T., Bogart, S., Reamer, L. A., Schapiro, S. J., & Hopkins, W. D. (2015). Multimodal communication in

- chimpanzees. *American Journal of Primatology*, 77(11), 1143–1148. <https://doi.org/10.1002/ajp.22449>
- Tempelmann, S., & Liebal, K. (2012). Spontaneous use of gesture sequences in orangutans: a case for strategy? In *Developments in Primate Gesture Research* (pp. 73–92). John Benjamins Pub. Company.
- Thierry, B. (2007). Unity in diversity: lessons from macaque societies. *Evolutionary Anthropology*, 16, 224–238.
- Thierry, B., Iwaniuk, A. N., & Pellis, S. M. (2000). The Influence of Phylogeny on the Social Behaviour of Macaques (Primates: Cercopithecidae, genus *Macaca*). *Ethology*, 106(8), 713–728. <https://doi.org/10.1046/j.1439-0310.2000.00583.x>
- Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology*, 42(1), 144–168. [https://doi.org/10.1016/0022-0965\(86\)90020-2](https://doi.org/10.1016/0022-0965(86)90020-2)
- Tomasello, M., & Call, J. (2018). Thirty years of great ape gestures. *Animal Cognition*. <https://doi.org/10.1007/s10071-018-1167-1>
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpentier, M. (1994). The learning and use of gestural signals by young chimpanzees: a trans-generational study. *Primates*, 35(2), 137–154. <https://doi.org/10.1007/BF02382050>
- Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Farrar, & Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, 14(2), 175–186. [https://doi.org/10.1016/S0047-2484\(85\)80005-1](https://doi.org/10.1016/S0047-2484(85)80005-1)
- Tomasello, M., Gust, D., & Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates*, 30(1), 35–50. <https://doi.org/10.1007/BF02381209>
- Tronick, E., Brazelton, T. B., & Als, H. (1978). The structure of face-to-face interaction and its developmental functions. *Sign Language Studies*, 18, 1–16. JSTOR.
- Vauclair, J. (2013). Communication des primates humains et non humains : à la recherche des origines du langage – Conclusions. *Revue de primatologie*, 5. <https://doi.org/10.4000/primatologie.1740>
- Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0292>
- Waller, B. M., & Cherry, L. (2012). Facilitating Play Through Communication: Significance of Teeth Exposure in the Gorilla Play Face: Facilitating Play Through Communication. *American Journal of Primatology*, 74(2), 157–164. <https://doi.org/10.1002/ajp.21018>
- Waller, B. M., & Dunbar, R. I. M. (2005). Differential Behavioural Effects of Silent Bared Teeth Display and Relaxed Open Mouth Display in Chimpanzees (*Pan troglodytes*). *Ethology*, 111(2), 129–142. <https://doi.org/10.1111/j.1439-0310.2004.01045.x>
- Waller, B. M., Lembeck, M., Kuchenbuch, P., Burrows, A. M., & Liebal, K. (2012). GibbonFACS: A Muscle-Based Facial Movement Coding System for Hylobatids. *International Journal of Primatology*, 33(4), 809–821. <https://doi.org/10.1007/s10764-012-9611-6>
- Waller, B. M., Liebal, K., Burrows, A. M., & Slocombe, K. E. (2013). How can a multimodal approach to primate communication help us understand the evolution of communication? *Evolutionary Psychology*, 11(3), 147470491301100320.
- Waller, B. M., Whitehouse, J., & Micheletta, J. (2016). Macaques can predict social outcomes from facial expressions. *Animal Cognition*, 19(5), 1031–1036. <https://doi.org/10.1007/s10071-016-0992-3>
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Weiss, M., Hultsch, H., Adam, I., Scharff, C., & Kipper, S. (2014). The use of network analysis to study complex animal communication systems: A study on nightingale song. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785), 20140460. <https://doi.org/10.1098/rspb.2014.0460>

- Whitehead, H. (1999). Testing association patterns of social animals. *Animal Behaviour*, 57, 26–29. 305–316. <https://doi.org/10.1016/j.anbehav.2016.10.024>
- Whitehead, H. (2008). *Analyzing Animal Societies* (The University of Chicago Press). <https://www.press.uchicago.edu/ucp/books/book/chicago/A/bo5607202.html>
- Whitehead, H. (2009). SOCPROG programs: analysing animal social structures. *Behavioral Ecology and Sociobiology*, 63(5), 765–778. <https://doi.org/10.1007/s00265-008-0697-y>
- Wiley, R. H. (1975). Multidimensional Variation in an Avian Display: Implications for Social Communication. *Science*, 190(4213), 482–483. <https://doi.org/10.1126/science.190.4213.482>
- Wilke, C., Kavanagh, E., Donnellan, E., Waller, B. M., Machanda, Z. P., & Slocombe, K. E. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, *Pan troglodytes schweinfurthii*. *Animal Behaviour*, 123, 305–316. <https://doi.org/10.1016/j.anbehav.2016.10.024>
- Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (*Taeniopygia guttata*). *Journal of Experimental Biology*, 204(20), 3497–3506. <https://doi.org/10.1242/jeb.204.20.3497>
- Wu, Y. C., & Coulson, S. (2007). How iconic gestures enhance communication: An ERP study. *Brain and Language*, 101(3), 234–245. <https://doi.org/10.1016/j.bandl.2006.12.003>
- Zuberbühler, K. (2005). The Phylogenetic Roots of Language: Evidence From Primate Communication and Cognition. *Current Directions in Psychological Science*, 14(3), 126–130. <https://doi.org/10.1111/j.0963-7214.2005.00357.x>
- Zuberbühler, K. (2020). Syntax and compositionality in animal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), 20190062. <https://doi.org/10.1098/rstb.2019.0062>

FIGURES

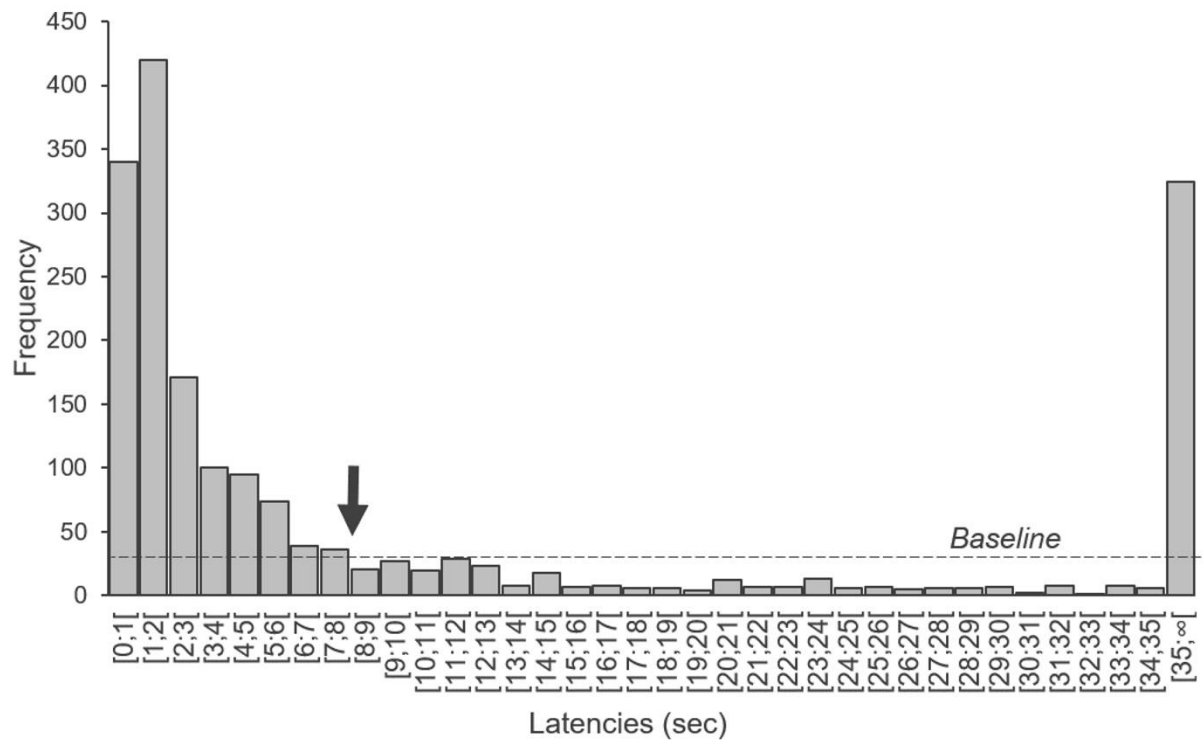


Figure 1. Distribution of the latencies between two consecutive signal productions from a same signaller toward a same receiver (n = 1878 latencies). Black arrow: threshold indicating the maximum latency between two signals from the same communicative sequence. Grey dotted line: baseline representing apparent “background noise” in frequency values.

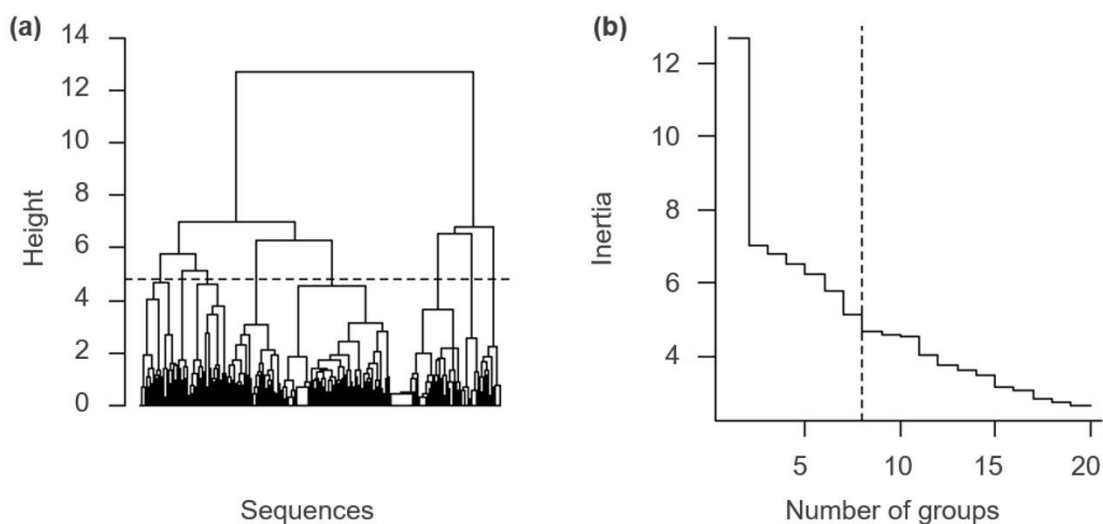


Figure 2. (a) Dendrogram of the hierarchical clustering (Ward method) made on sequence dissimilarities (N = 424). (b) Representation of the dendrogram inertia. The 8-group solution (indicated by a dotted line in both graphs) presented the best between-cluster distance.

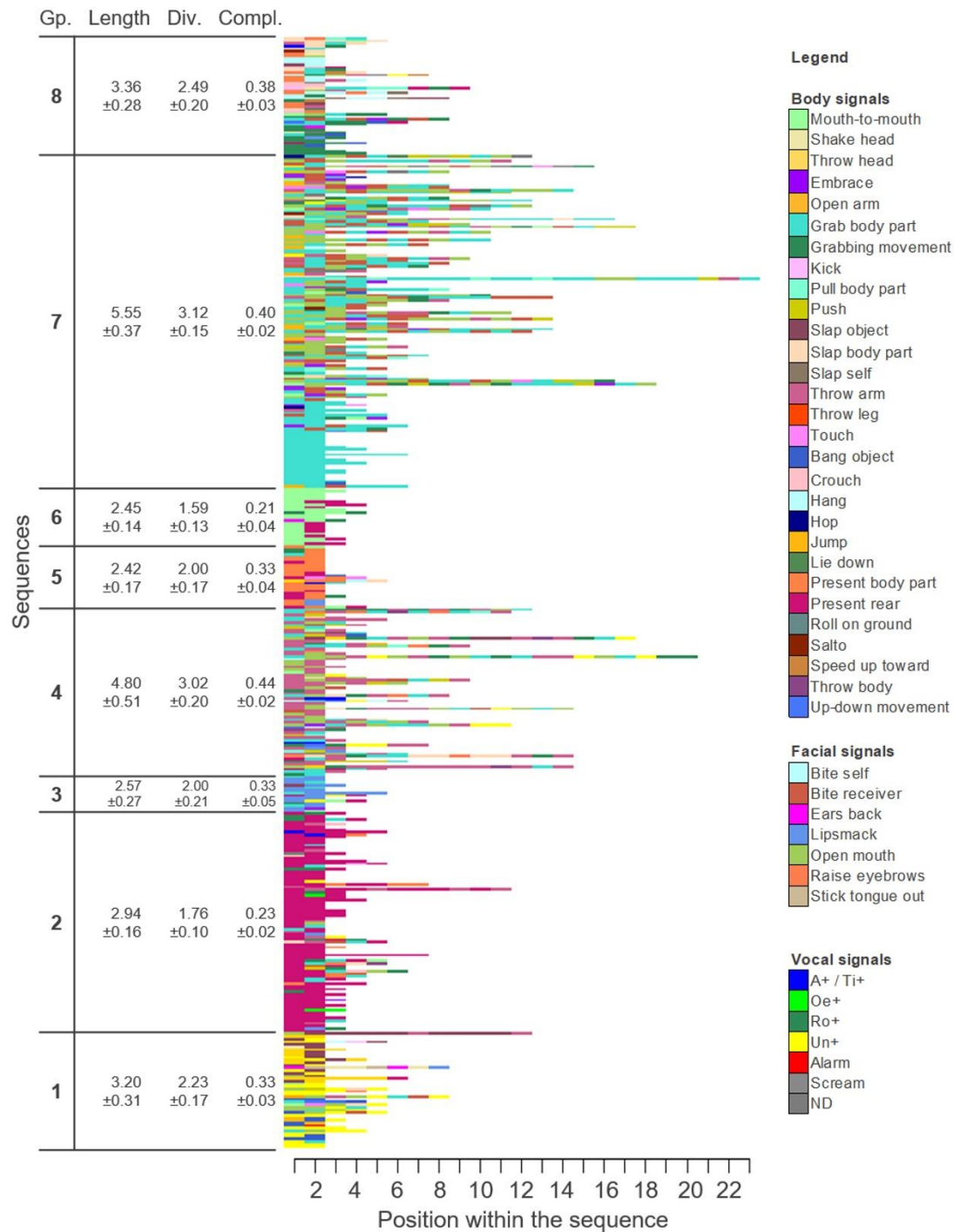


Figure 3. Sequence index plot and sequence group characteristics. Each sequence is represented by a line and each signal by a colour. Sequences are sorted by groups, and characteristics are given on the left as mean values $\pm S.E.$ **Gp.:** group number; **Length:** number of signals within the sequence; **Div.:** diversity (number of different signals within the sequence); **Compl.:** complexity index (from 0 to 1, based on longitudinal entropy and number of transitions within the sequence. Gabadinho et al., 2010, 2011). Different capital letters indicate significant differences between groups (Mann-Whitney tests with FDR correction: $P < 0.05$).

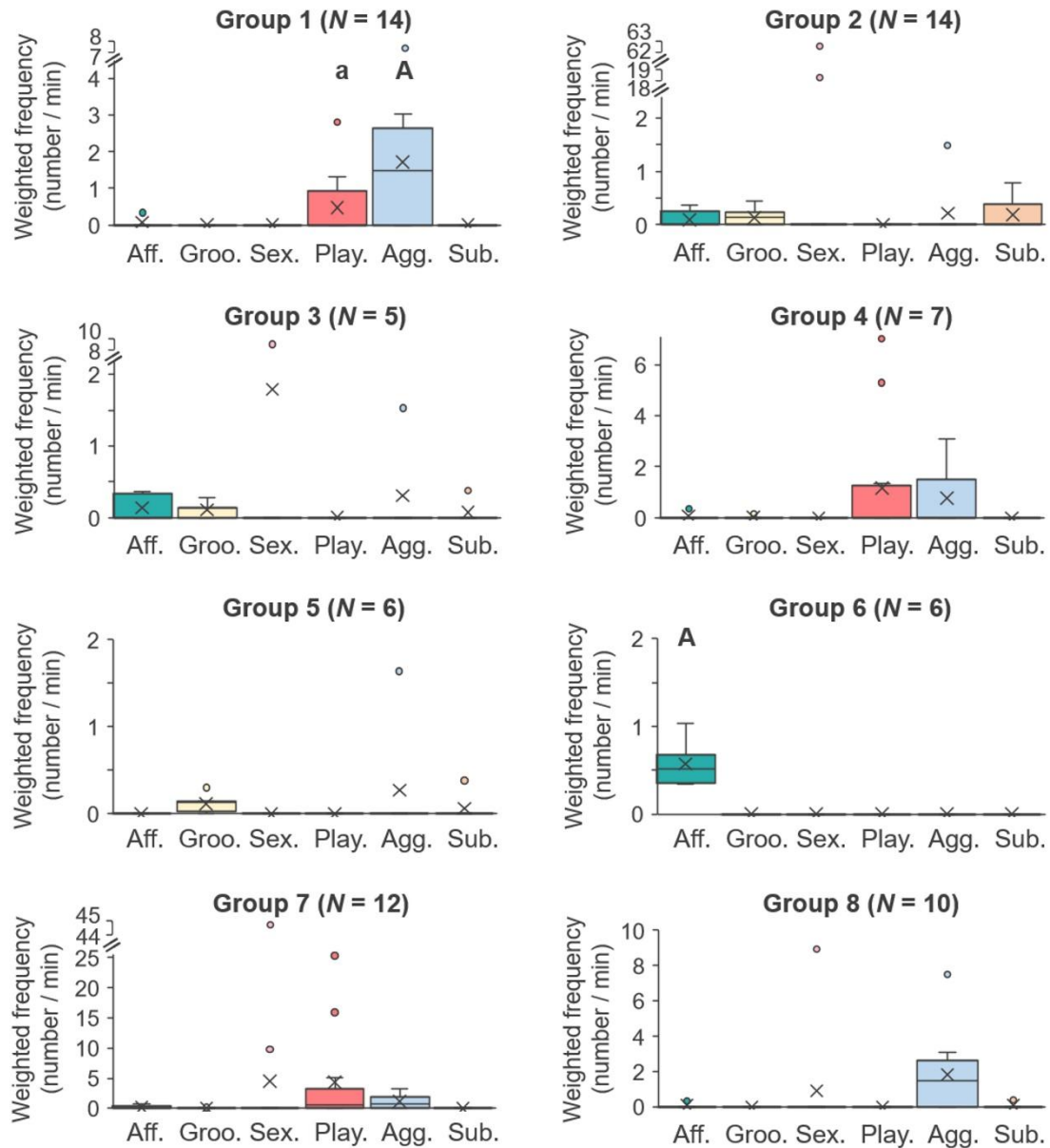


Figure 4. Social contexts of sequence production of different groups. N: number of signallers. Different capital letters indicate significant differences between groups (Wilcoxon signed rank test with FDR correction: $P < 0.05$), and minor letters are used when these are tendencies: $P < 0.01$. **Aff.:** affiliation; **Groo.:** grooming; **Sex.:** sexual context; **Agg.:** aggression; **Sub.:** submission.

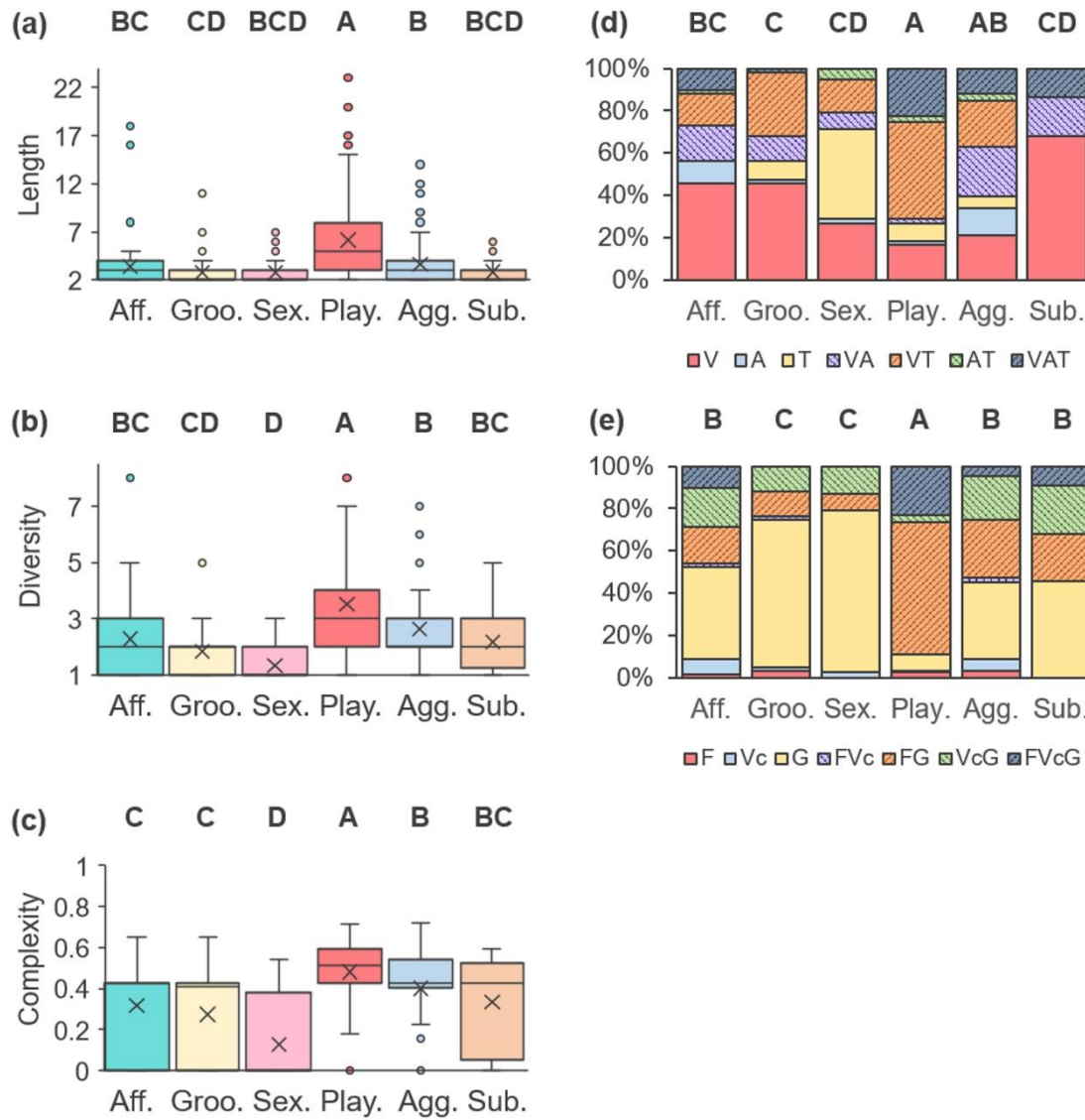


Figure 5. Sequence characteristics depending on social contexts. (a) **Sequence length** (number of signals within the sequence), (b) **diversity** (number of different signals within the sequence), (c) **complexity** (indices from 0 to 1, based on longitudinal entropy and number of transitions within the sequence. Gabadinho et al., 2010, 2011), (d) **multimodality** (association of different signal modalities) and (e) **multicomponentiality** (association of different signal types) of communicative sequences were compared between contexts using respectively a GLMM of Gamma family, of Poisson family, a LMM, and GLMMs of the Binomial family. Different capital letters indicate significant differences between contexts, i.e. **Aff.:** affiliative; **Groo.:** grooming; **Sex.:** sexual; **Play.:** playful; **Agg.:** aggressive and **Sub.:** submissive contexts. **V:** “visual only” signals; **A:** audible signals; **T:** tactile signals; **F:** facial signals; **Vc:** vocal signals; **G:** gestural and body signals. Plain filling indicates unimodal or unicomponent sequences, hatching indicates multimodal or multicomponent sequences.

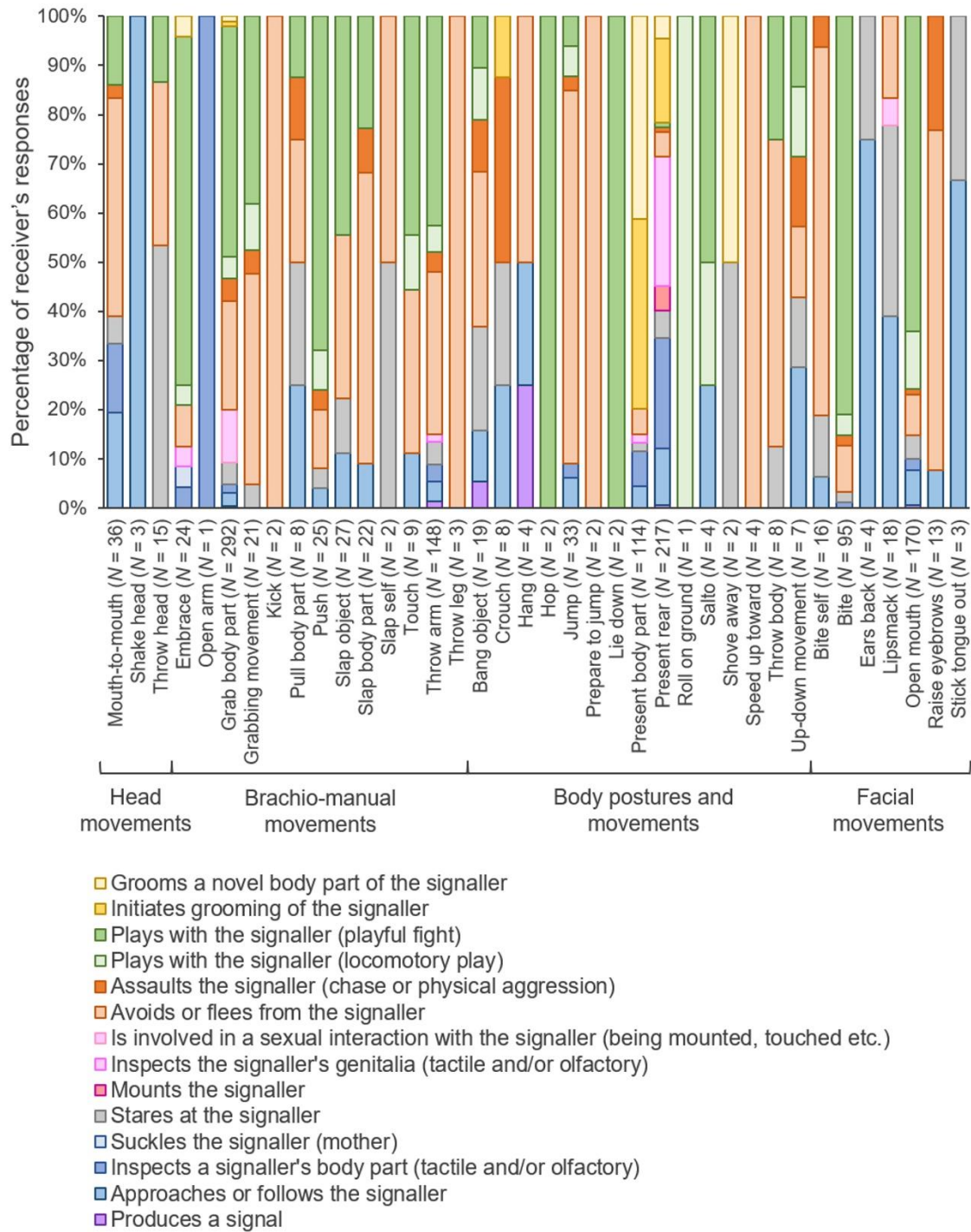


Figure A1. Responses of receivers to body and facial signals. N = 1384 occurrences of body and facial signals induced a response from the receiver, i.e. a change of behaviour within the 5 seconds following the signal (e.g. Aychet et al., 2021; Liebal et al., 2004, 2006; Roberts et al., 2014; Schel et al., under review).

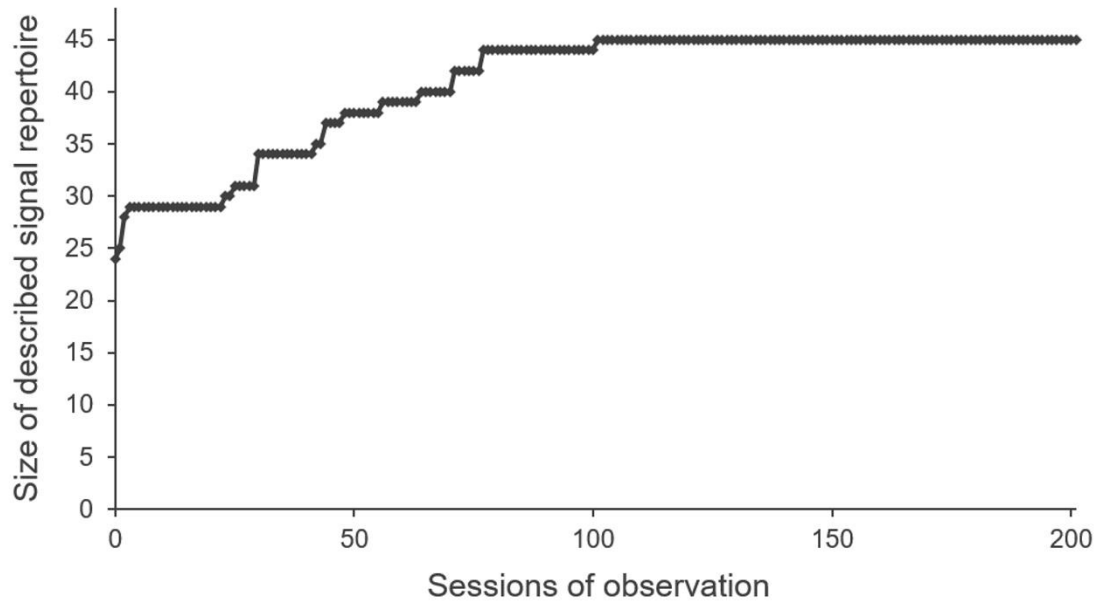


Figure A2. Cumulative number of described signals over observation sessions. At the moment we began our study, 24 signals have already been described in captive red-capped mangabeys (i.e. vocal signals described through 77 hours of recordings: by Bouchet et al., 2010; body signals described through 34 hours of observations: Schel et al., under review). A maximum number of 45 signals was reached at our 101st session of observation in the studied captive population.

TABLES

Table 1. Characteristics of the red-capped mangabeys

Social group	Subject	Sex	Date of birth	Age Category
I.	Triskelle	Female	21 April 2015	Juvenile
	Chipse	Female	03 January 2006	Middle-aged
	Julie	Female	08 May 2004	Middle-aged
	Maillette	Female	29 December 2009	Middle-aged
	Many	Female	14 August 2008	Middle-aged
	Bell	Female	31 March 2002	Old adult
	Chipie	Female	28 June 1992	Old adult
	Gofrette	Female	08 November 1996	Old adult
	Joly	Female	22 October 2000	Old adult
	Zunie	Female	03 July 1987	Old adult
	Litchi	Male	20 April 2015	Juvenile
	Pouët	Male	14 March 2015	Juvenile
	Kargi	Male	19 May 2005	Middle-aged
	Kamel	Male	07 September 2010	Middle-aged
II.	Roby	Male	18 November 2010	Middle-aged
	Coët	Male	31 August 2011	Middle-aged
III.	Tips	Male	10 July 2011	Middle-aged
	Bandit	Male	15 June 1991	Old adult
	Carillon	Male	02 April 2007	Middle-aged
IV.	Elky	Male	06 November 2009	Middle-aged
	George	Male	05 June 2006	Middle-aged
	Pirate	Male	17 October 1992	Old adult
	Isba	Male	20 April 2004	Middle-aged
V.	Lenni	Male	07 October 2006	Middle-aged
	Marti	Male	16 October 1998	Old adult

Table 2. Social contexts of signal production

Social context	Associated behaviours in signaller and receiver
Affiliative	Physical proximity between subjects (at less than one arm length), calm approach of recipient or gentle physical contact
Grooming	Grooming, tactile examination
Sexual	Touching, smelling, licking genital parts or mounting
Playful	Play-fight, rough or gentle (involving manual fighting, biting, gentle or rough touching and grabbing), or locomotor-rotational play (play with few physical contacts, but involving pursuits, jumps, somersaults)
Aggressive	Physical aggression of recipient by signaller (biting, beating, rough manipulation), flight or avoidance of signaller by recipient, or intergroup conflict in which signaller and recipient were in different social groups
Submissive	Flight or avoidance of recipient by signaller
Unclear	Behaviours accompanying the signal were not sufficient to classify the context, or the interaction could be classified in several categories.

Table 3. Repertoire of communicative signals observed in captive red-capped mangabeys. Modality: main modality on which the signal could be perceived; **Nb:** total number of occurrences observed, with number of occurrences recorded within communicative sequences (as opposed to single occurrences) indicated within brackets; **N:** number of signallers.

Type	Subtype	Signal	Description	Modality	Nb	N
Body signals	Head movements	Mouth-to-mouth	Signaller approaches his/her nose close to the receiver's face (few centimetres or less).	Visual	106 (59)	21
		Shake head ¹	Signaller moves his/her head with quick repeated movements, horizontally or vertically.	Visual	6 (6)	3
		Throw head ¹	Signaller produces a brief head movement toward the receiver.	Visual	26 (19)	9
	Brachio-manual movements	Embrace	Signaller puts one or two arms around the receiver's body.	Tactile	35 (27)	9
		Open arm	Signaller opens one or two arms in the direction of the receiver, like before embracing, without contact.	Visual	4 (1)	4
		Grab body part ¹	Signaller closes one or two hands or feet on a receiver's body part.	Tactile	412 (364)	21
		Grabbing movement ¹	Signaller throws one or two arms in the receiver's direction, with hands closing at the end of the movement.	Visual	24 (22)	9
		Kick	Signaller hits the receiver with one or two feet.	Tactile	4 (2)	3
		Pull body part	Signaller holds and pulls a receiver's body part.	Tactile	16 (16)	7
		Push	Signaller pushes the receiver away, with hands or feet.	Tactile	29 (27)	10
		Slap object ¹	Signaller slaps cage element or ground with an open hand.	Audible	47 (33)	12
		Slap body part ¹	Signaller hits a receiver's body part with an open hand.	Tactile	41 (37)	11
		Slap self	Signaller hits him/herself with an open hand, in a unique or repeated movement.	Visual	3 (3)	1
		Throw arm ¹	Signaller throws one or two arms in the receiver's direction.	Visual	204 (169)	21
		Throw leg	Signaller throws one leg in the receiver's direction.	Visual	3 (2)	2
		Touch ¹	Signaller gently puts one or two open hands on receiver's body.	Tactile	19 (14)	6
	Body postures and movements	Bang object ¹	Signaller jumps on cage wall or perches after having taken a run-up, or grabs and shakes cage wall or perches (with hands only or with hands and feet), producing tremors and noise.	Audible	44 (19)	16
		Crouch ¹	Signaller bents his/her four limbs, belly touching the ground.	Visual	11 (8)	6

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

		Hang	Signaller hangs with arms or legs to the cage or perches, in front of the receiver.	Visual	6 (5)	3
		Hop ¹	Signaller jumps repetitively, in one position or moving around.	Visual	3 (3)	3
		Jump	Signaller jumps and land in front of the receiver or on him/her.	Visual	37 (17)	10
		Prepare to jump	Signaller is oriented toward the receiver, limbs bent, in tension, staring at the receiver.	Visual	2 (0)	2
		Lie down ¹	Signaller lays down on his/her back, in front of the receiver.	Visual	2 (2)	1
		Present body part ¹	Signaller exposes one particular body part (head, neck, limb, back, chest, belly, side or crotch) to the receiver, by orienting his/her body appropriately in the direction and/or close to him/her.	Visual	161 (41)	22
		Present rear ¹	Signaller puts his/her rear in the direction and/or close to the receiver.	Visual	351 (210)	24
		Roll on ground ¹	Signaller rolls on ground in front of the receiver.	Visual	2 (1)	1
		Salto ¹	Signaller jumps and turns in the air in front of the receiver.	Visual	4 (3)	2
		Shove away	Signaller produces a brief movement with a body part that the receiver is touching, to push him away.	Tactile	3 (0)	3
		Speed up toward	Signaller runs toward the receiver, in a fast movement, and stops close to him/her.	Visual	4 (3)	2
		Throw body ¹	Signaller produces a brief movement toward the receiver with his/her upper body.	Visual	12 (5)	11
		Up-down movement	Signallers produces unique or repeated up-and-down movements with his/her body, in bipedal or quadrupedal posture.	Visual	8 (7)	5
		Bite self	Signaller bites herself/himself or an object, staring at the receiver.	Visual	25 (20)	3
		Bite	Signaller closes his/her mouth on a body part of the receiver.	Tactile	114 (114)	9
Facial signals	Facial displays	Ears back ²	Signaller puts his/her ears backwards, stretching the upper part of his/her face.	Visual	5 (4)	3
		Lipsmack ²	Signaller sticks out his/her tongue and moves it between the lips with rapid repeated movements.	Visual	62 (30)	19

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

Vocal signals ³	Open mouth ²	Signaller opens his/her mouth completely (uncovering upper and lower teeth), half or slightly (forming an oval shape with the lips, teeth covered or slightly visible).	Visual	223 (209)	14
	Raise eyebrows ²	Signaller raises his/her eyebrows, enlarging the white part of eyelids.	Visual	22 (21)	14
	Stick tongue out ²	Signaller sticks his/her tongue out and retracts it briefly (not repeated).	Visual	3 (2)	2
	A+ / Ti+	Signaller produces series of "A" or "Ti" vocal units ("food call", "contact call").	Audible	18 (8)	5
	Oe+	Signaller produces series of "Oe" vocal units ("sexual call").	Audible	12 (3)	4
	Ro+	Signaller grunts, producing series of "Ro" vocal units ("contact call", "grunt").	Audible	123 (87)	19
	Un+	Signaller produces series of "Un" vocal units ("threat call").	Audible	87 (61)	20
	Alarm	Signaller produces WaHoo calls ("alarm calls") or Whoop-Gobble ("loud call").	Audible	1 (1)	1
	Scream	Signaller screams, producing "Wi" vocal unit ("distress call").	Audible	2 (2)	1
	ND	Signaller produces unidentified vocalisations.	Audible	12 (12)	5

¹Schel et al., under review; ²Aychet et al., 2021; ³Bouchet et al., 2010

Table 4. Sequence group description and summary of production context. **Gp.:** group number; **Nb (%)**: number of sequences within the group, and corresponding percentage of the total sample; **N**: number of different signallers; **Nb.sq**: number of representative sequences. Representative sequences are ordered by respective coverage importance. Each representative sequence is described by separating consecutive signals by a semicolon, and if signals are consecutively repeated, “x k” indicates the number k of consecutive instances.

Gp.	Nb (%)	N	Representative sequences		Sequence production		
			Nb.sq	Coverage	Description	Main social contexts	Age and sex of signallers
1	44 (10.4%)	18	3	27.3%	- Un+ x2 - Slap object; Un+ - Throw head; Slap object	Aggressive	No specific age and sex category
2	84 (19.8%)	17	4	51.2%	- Present rear x2 - Present rear x3 - Present rear; Grab body part - Present rear; Throw arm	-	Females
3	14 (3.3%)	9	3	50.0%	- Lipsmack x3 - Grab body part; Lipsmack - Lipsmack; Grab body part	-	No specific age and sex category
4	64 (15.1%)	15	5	25.0%	- Throw arm x2 - Open mouth; Throw arm - Grab body part x2; Throw arm - Throw arm; Open mouth - Throw arm; Open mouth; Throw arm	Aggressive and playful	Juveniles
5	24 (5.7%)	10	3	50.0%	- Present body part x2 - Present rear; Present body part - Present body part; Lipsmack	-	No specific age and sex category
6	22 (5.2%)	9	2	63.6%	- Mouth-to-mouth x2 - Mouth-to-mouth; Present rear	Affiliative	Juveniles
7	127 (30.0%)	13	8	25.2%	- Grab body part x2 - Grab body part x3	Playful and aggressive	Juveniles

8	45 (10.6%)	14	8	26.7%	- Grab body part x4	Aggressive	No specific age and sex category
					- Grab body part; Bite		
					- Embrace; Bite; Embrace		
					- Jump; Throw arm		
					- Grab body part; Open mouth		
					- Ro+ x2		
					- Bite self x2		
					- Ro+ x4		
					- Ro+ x3		
					- Ro+; Bang object		
					- Present rear; Ro+ x2; Bang object		
					- Ro+; Lipsmack; Ro+		
					- Grab body part; Ro+; Bang object		

Table 5. Signal network statistics. Values are given as median \pm IQR for each category. *N*: number of different signals in each category. Different letters in superscript indicate significant differences (Mann-Whitney tests with FDR correction: $P < 0.05$)

Signal categories		Degree	Strength	Affinity	Clustering coefficient
Type	Body signals ($N = 29$)	11.00 \pm 17.00	1.47 \pm 4.22	3.52 \pm 1.99	0.27 \pm 0.06
	Facial signals ($N = 7$)	15.00 \pm 8.00	1.96 \pm 1.80	3.35 \pm 1.26	0.25 \pm 0.13
	Vocal signals ($N = 7$)	7.00 \pm 23.00	1.33 \pm 3.53	3.54 \pm 1.63	0.28 \pm 0.11
Modality	Visual only ($N = 26$)	10.00 \pm 10.25	1.29 \pm 1.18 ^b	3.40 \pm 1.41	0.25 \pm 0.12
	Audible ($N = 9$)	12.00 \pm 15.50	1.47 \pm 2.56 ^{ab}	3.54 \pm 1.33	0.28 \pm 0.11
	Tactile ($N = 8$)	16.00 \pm 7.75	2.92 \pm 2.61 ^a	3.83 \pm 0.93	0.31 \pm 0.17

Table A1. Mangabey communicative sequences' characteristics depending on social contexts: detailed results of the mixed models ($N = 424$ sequences). Social contexts are included as fixed effects and the identities of the signallers ($N_{\text{signallers}} = 24$) and receivers ($N_{\text{receivers}} = 26$). **GLMM**: Generalized Linear Mixed Model; **LMM**: Linear Mixed Model; **Aff.**: affiliation; **Groo.**: grooming; **Sex.**: sexual context; **Agg.**: aggression; **Sub.**: submission; **Un.**: Unclear social context. **S.E.**: Standard error; **df**: degrees of freedom; **Var.**: Variance; **S.D.**: Standard deviation; **Min.**: Minimum; **Q1**: First quartile; **Med.**: Median; **Q3**: Third quartile; **Max.**: Maximum. Significance of statistical tests: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.

Dependant variable	Model family	Fixed effects: Social contexts							Random effects			Residuals				
Length	GLMM Gamma	Reference	Levels	Estimate	S.E.	t	P	Significance	Variables	Var.	S.D.	Min.	Q1	Med.	Q3	Max.
		Aff.	Agg.	-0.020	0.026	-0.797	0.426		Signaller	0.001	0.031	-1.255	-0.599	-0.326	0.306	6.200
			Groo.	0.058	0.031	1.904	0.057		Receiver	0.001	0.036					
			Play.	-0.116	0.024	-4.801	< 0.001	***	Residual	0.367	0.606					
			Sex.	0.043	0.041	1.054	0.292									
			Sub.	0.050	0.044	1.121	0.262									
			Un.	0.094	0.040	2.353	0.019	*								
		Agg.	Groo.	0.079	0.029	2.745	0.006	**								
			Play.	-0.096	0.022	-4.366	< 0.001	***								
			Sex.	0.063	0.039	1.627	0.104									
			Sub.	0.070	0.043	1.631	0.103									
			Un.	0.115	0.039	2.967	0.003	**								
		Groo.	Play.	-0.175	0.029	-5.974	< 0.001	***								
			Sex.	-0.016	0.043	-0.364	0.716									
			Sub.	-0.009	0.046	-0.191	0.849									
			Un.	0.036	0.042	0.841	0.401									
		Play.	Sex.	0.159	0.040	3.983	< 0.001	***								
			Sub.	0.166	0.044	3.807	< 0.001	***								
			Un.	0.211	0.039	5.424	< 0.001	***								
		Sex.	Sub.	0.007	0.053	0.126	0.900									

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

			Un.	0.051	0.049	1.038	0.299										
		Sub.	Un.	0.045	0.052	0.848	0.396										
Diversity	GLMM Poisson	Refer ence	Levels	Estimate	S.E.	Z	P	Signific ance	Variable s	Var.	S.D.	Min.	Q1	Med.	Q3	Max.	
		Aff.	Agg.	0.146	0.108	1.360	0.174		Signaller	0.000	0.000	-1.341	-0.398	-0.232	0.466	3.776	
			Groo.	-0.214	0.129	-1.661	0.097		Receiver	0.000	0.000						
			Play.	0.429	0.099	4.338	< 0.001	***									
			Sex.	-0.534	0.164	-3.246	0.001	**									
			Sub.	-0.048	0.168	-0.283	0.777										
			Un.	-0.462	0.167	-2.772	0.006	**									
		Agg.	Groo.	-0.360	0.115	-3.120	0.002	**									
			Play.	0.282	0.081	3.502	< 0.001	***									
			Sex.	-0.680	0.154	-4.410	< 0.001	***									
			Sub.	-0.194	0.158	-1.226	0.220										
			Un.	-0.608	0.157	-3.883	< 0.001	***									
		Groo.	Play.	0.642	0.107	5.984	< 0.001	***									
			Sex.	-0.320	0.170	-1.884	0.060										
			Sub.	0.166	0.173	0.960	0.337										
			Un.	-0.248	0.172	-1.444	0.149										
		Play.	Sex.	-0.962	0.148	-6.492	< 0.001	***									
			Sub.	-0.476	0.152	-3.127	0.002	**									
			Un.	-0.891	0.151	-5.905	< 0.001	***									
		Sex.	Sub.	0.486	0.201	2.416	0.016	*									
			Un.	0.071	0.200	0.356	0.722										
		Sub.	Un.	-0.415	0.203	-2.042	0.041	*									
Complexit y	LMM	Refer ence	Levels	Estimate	S.E.	df	t	P	Sign ifica nce	Variable s	Var.	S.D.	Min.	Q1	Med.	Q3	Max.

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

		Aff.	Agg.	0.090	0.033	358.962	2.724	0.007	***	Signaller	0.002	0.043	-2.666	-0.662	0.171	0.693	2.169
			Groo.	-0.042	0.035	413.715	-1.192	0.234		Receiver	0.000	0.008					
			Play.	0.163	0.033	162.871	5.006	< 0.001	***	Residual	0.036	0.190					
			Sex.	-0.170	0.044	182.387	-3.861	< 0.001	***								
			Sub.	0.023	0.048	405.949	0.479	0.632									
			Un.	-0.173	0.041	408.414	-4.175	< 0.001	***								
	Agg.		Groo.	-0.132	0.033	379.599	-4.022	< 0.001	***								
			Play.	0.073	0.030	114.965	2.409	0.018	*								
			Sex.	-0.260	0.041	177.529	-6.390	< 0.001	***								
			Sub.	-0.067	0.047	374.052	-1.438	0.151									
			Un.	-0.263	0.039	383.494	-6.719	< 0.001	***								
	Groo.		Play.	0.205	0.034	119.470	6.104	< 0.001	***								
			Sex.	-0.127	0.043	202.935	-2.932	0.004	**								
			Sub.	0.065	0.048	391.401	1.353	0.177									
			Un.	-0.131	0.041	371.860	-3.151	0.002	**								
	Play.		Sex.	-0.333	0.042	101.137	-7.912	< 0.001	***								
			Sub.	-0.140	0.047	242.855	-2.963	0.003	**								
			Un.	-0.336	0.040	206.384	-8.442	< 0.001	***								
	Sex.		Sub.	0.193	0.055	271.635	3.533	< 0.001	***								
			Un.	-0.003	0.048	268.921	-0.068	0.946									
	Sub.		Un.	-0.196	0.053	410.533	-3.713	< 0.001	***								
Multimodality	GLMM Binomial	Reference	Levels	Estimate	S.E.	Z	P	Significance		Variables	Var.	S.D.	Min.	Q1	Med.	Q3	Max.
		Aff.	Agg.	-0.641	0.346	-1.855	0.064			Signaller	0.000	0.000	-2.926	-0.797	-0.528	0.868	1.892
			Groo.	0.036	0.377	0.096	0.924			Receiver	0.062	0.249					
			Play.	-1.250	0.348	-3.590	< 0.001	***									
			Sex.	0.737	0.459	1.607	0.108										
			Sub.	0.552	0.537	1.029	0.303										
			Un.	1.787	0.598	2.987	0.003	**									

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

			Agg.	Groo.	0.677	0.346	1.959	0.050	*								
				Play.	-0.609	0.315	-1.936	0.053									
				Sex.	1.378	0.431	3.195	0.001	**								
				Sub.	1.193	0.516	2.315	0.021	*								
				Un.	2.428	0.580	4.187	< 0.001	***								
			Groo.	Play.	-1.286	0.353	-3.643	< 0.001	***								
				Sex.	0.701	0.455	1.542	0.123									
				Sub.	0.516	0.538	0.959	0.337									
				Un.	1.751	0.600	2.916	0.004	**								
			Play.	Sex.	1.988	0.437	4.549	< 0.001	***								
				Sub.	1.803	0.520	3.467	< 0.001	***								
				Un.	3.037	0.583	5.209	< 0.001	***								
			Sex.	Sub.	-0.185	0.595	-0.311	0.756									
				Un.	1.049	0.652	1.610	0.107									
			Sub.	Un.	1.235	0.709	1.740	0.082									
Multicom ponentiali ty	GLMM Binomial	Refer ence	Levels		Estimate	S.E.	Z	P	Signific ance	Variable s	Var.	S.D.	Min.	Q1	Med.	Q3	Max.
				Aff.	Agg.	-0.262	0.356	-0.737	0.461								
					Groo.	1.029	0.414	2.485	0.013								
					Play.	-2.197	0.428	-5.131	< 0.001								
					Sex.	1.331	0.530	2.513	0.012								
					Sub.	-0.284	0.524	-0.542	0.588								
					Un.	1.468	0.535	2.741	0.006								
				Agg.	Groo.	1.291	0.383	3.371	< 0.001								
					Play.	-1.935	0.404	-4.788	< 0.001								
					Sex.	1.593	0.494	3.227	0.001								
					Sub.	-0.022	0.505	-0.044	0.965								
					Un.	1.730	0.519	3.336	0.001								
				Groo.	Play.	-3.226	0.465	-6.944	< 0.001								

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

	Sex.	0.302	0.533	0.566	0.571	
	Sub.	-1.314	0.549	-2.392	0.017	*
	Un.	0.438	0.565	0.775	0.438	
Play.	Sex.	3.528	0.563	6.263	< 0.001	***
	Sub.	1.913	0.569	3.364	< 0.001	***
	Un.	3.665	0.577	6.353	< 0.001	***
Sex.	Sub.	-1.616	0.635	-2.545	0.011	*
	Un.	0.137	0.652	0.209	0.834	
Sub.	Un.	1.752	0.645	2.715	0.007	**

Table A2. Frequency of sequence production depending on social context: detailed results of Wilcoxon signed-rank tests with FDR corrections.

Significance: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.

Compared social contexts	Sequence groups							
	Group 1 (N = 14)	Group 2 (N = 14)	Group 3 (N = 5)	Group 4 (N = 14)	Group 5 (N = 6)	Group 6 (N = 6)	Group 7 (N = 12)	Group 8 (N = 10)
Affiliative vs Aggressive	V = 6, P = 0.032 *	V = 10, P = 1.000	V = 3, P = 1.000	V = 3, P = 0.170	V = 0, P = 1.000	V = 21, P = 0.036 *	V = 3, P = 0.110	V = 3, P = 0.170
Affiliative vs Grooming	V = 6, P = 0.186	V = 18, P = 0.890	V = 6, P = 1.000	V = 12, P = 0.400	V = 0, P = 0.590	V = 21, P = 0.036 *	V = 25, P = 0.140	V = 3, P = 0.620
Affiliative vs Playful	V = 6, P = 0.204	V = 10, P = 0.410	V = 3, P = 1.000	V = 6, P = 0.210	-	V = 21, P = 0.036 *	V = 3, P = 0.110	V = 3, P = 0.620
Affiliative vs Sexual	V = 6, P = 0.186	V = 10, P = 0.830	V = 3, P = 1.000	V = 6, P = 0.290	-	V = 21, P = 0.036 *	V = 10, P = 1.000	V = 3, P = 1.000
Affiliative vs Submissive	V = 6, P = 0.186	V = 10, P = 0.410	V = 3, P = 1.000	V = 6, P = 0.290	V = 0, P = 1.000	V = 21, P = 0.036 *	V = 15, P = 0.120	V = 3, P = 0.770
Affiliative vs Unclear	V = 6, P = 0.186	V = 24, P = 1.000	V = 3, P = 1.000	V = 9, P = 0.300	V = 0, P = 1.000	V = 21, P = 0.036 *	V = 15, P = 0.120	V = 7, P = 0.770
Aggressive vs Grooming	V = 45, P = 0.032 *	V = 17, P = 0.830	V = 3, P = 1.000	V = 25, P = 0.170	V = 4, P = 1.000	-	V = 27, P = 0.110	V = 21, P = 0.170
Aggressive vs Playful	V = 57, P = 0.091	V = 3, P = 0.810	V = 1, P = 1.000	V = 23, P = 1.000	V = 1, P = 1.000	-	V = 19, P = 0.520	V = 21, P = 0.170
Aggressive vs Sexual	V = 45, P = 0.032 *	V = 3, P = 0.750	V = 1, P = 1.000	V = 15, P = 0.170	V = 1, P = 1.000	-	V = 10, P = 1.000	V = 16, P = 0.620
Aggressive vs Submissive	V = 45, P = 0.032 *	V = 15, P = 0.950	V = 2, P = 1.000	V = 15, P = 0.170	V = 2, P = 1.000	-	V = 21, P = 0.110	V = 27, P = 0.170
Aggressive vs Unclear	V = 45, P = 0.032 *	V = 13, P = 1.000	V = 1, P = 1.000	V = 20, P = 0.170	V = 2, P = 1.000	-	V = 21, P = 0.110	V = 33, P = 0.170
Grooming vs Playful	V = 0, P = 0.151	V = 36, P = 0.250	V = 6, P = 1.000	V = 3, P = 0.170	V = 10, P = 0.590	-	V = 3, P = 0.110	-

SEQUENTIAL AND NETWORK ANALYSES OF MULTIPLE SIGNAL USE IN CAPTIVE MANGABEYS

Grooming vs Sexual		$V = 28,$ $P = 1.000$	$V = 3,$ $P = 1.000$	$V = 3,$ $P = 0.460$	$V = 10,$ $P = 0.590$		$V = 1,$ $P = 0.520$	$V = 0,$ $P = 1.000$
-						-		
Grooming vs Submissive		$V = 19,$ $P = 0.830$	$V = 6,$ $P = 1.000$	$V = 3,$ $P = 0.460$	$V = 10,$ $P = 1.000$		$V = 3,$ $P = 0.520$	$V = 0,$ $P = 0.620$
-						-		
Grooming vs Unclear		$V = 30,$ $P = 1.000$	$V = 6,$ $P = 1.000$	$V = 3,$ $P = 1.000$	$V = 11,$ $P = 1.000$		$V = 4,$ $P = 0.920$	$V = 0,$ $P = 0.620$
-						-		
Playful vs Sexual	$V = 10,$ $P = 0.151$	$V = 0,$ $P = 0.640$	$V = 0,$ $P = 1.000$	$V = 15,$ $P = 0.170$			$V = 21,$ $P = 0.480$	$V = 0,$ $P = 1.000$
					-	-		
Playful vs Submissive	$V = 10,$ $P = 0.151$	$V = 0,$ $P = 0.250$	$V = 0,$ $P = 1.000$	$V = 15,$ $P = 0.170$	$V = 0,$ $P = 1.000$		$V = 21,$ $P = 0.110$	$V = 0,$ $P = 0.620$
						-		
Playful vs Unclear	$V = 10,$ $P = 0.151$	$V = 0,$ $P = 0.250$		$V = 20,$ $P = 0.170$	$V = 0,$ $P = 1.000$		$V = 21,$ $P = 0.110$	$V = 0,$ $P = 0.620$
			-			-		
Sexual vs Submissive		$V = 18,$ $P = 0.830$	$V = 2,$ $P = 1.000$		$V = 0,$ $P = 1.000$		$V = 3,$ $P = 0.520$	$V = 3,$ $P = 1.000$
-				-		-		
Sexual vs Unclear		$V = 18,$ $P = 0.830$	$V = 1,$ $P = 1.000$	$V = 0,$ $P = 1.000$	$V = 0,$ $P = 1.000$		$V = 3,$ $P = 0.520$	$V = 3,$ $P = 1.000$
-						-		
Submissive vs Unclear		$V = 32,$ $P = 0.750$	$V = 1,$ $P = 1.000$	$V = 0,$ $P = 1.000$	$V = 2,$ $P = 1.000$		$V = 0,$ $P = 1.000$	$V = 5,$ $P = 0.650$
-						-		

Table A3. Effect of signaller's age and sex on sequence production: detailed results of Mann-Whitney tests with FDR corrections. $N_{females} = 10$; $N_{males} = 14$; $N_{juveniles} = 3$; $N_{middle-aged} = 14$; $N_{old\ adults} = 7$. Significance: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.

Compared signaller's age and sex categories	Sequence groups							
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Females vs Males	$W = 47,$ $P = 0.171$	$W = 126.5,$ $P < 0.001$ ***	$W = 69,$ $P = 0.967$	$W = 89.5,$ $P = 0.248$	$W = 66,$ $P = 0.788$	$W = 86,$ $P = 0.232$	$W = 72.5,$ $P = 0.900$	$W = 68.5,$ $P = 0.948$
Juveniles vs Middle-age adults	$W = 24,$ $P = 0.750$	$W = 33,$ $P = 0.332$	$W = 18,$ $P = 0.574$	$W = 40,$ $P = 0.041$ *	$W = 24,$ $P = 0.696$	$W = 40,$ $P = 0.005$ **	$W = 40,$ $P = 0.042$ *	$W = 24,$ $P = 0.733$
Juveniles vs Old adults	$W = 15.5,$ $P = 0.372$	$W = 12,$ $P = 0.820$	$W = 6,$ $P = 0.393$	$W = 19,$ $P = 0.101$	$W = 13,$ $P = 0.696$	$W = 19,$ $P = 0.078$	$W = 20,$ $P = 0.051$	$W = 14.5,$ $P = 0.733$
Middle-aged vs Old adults	$W = 72.5,$ $P = 0.221$	$W = 33,$ $P = 0.332$	$W = 33,$ $P = 0.365$	$W = 28,$ $P = 0.105$	$W = 54,$ $P = 0.696$	$W = 38.5,$ $P = 0.221$	$W = 48.5,$ $P = 1.000$	$W = 56,$ $P = 0.733$