



Turn-taking in grooming interactions of sooty mangabeys (*Cercocebus atys*) in the wild

Pharisemène Tibesar¹ · Catherine Crockford^{2,3,4} · Auriane Le Floch^{2,3,5,6} · Simone Pika¹

Received: 24 June 2025 / Revised: 22 December 2025 / Accepted: 24 December 2025
© The Author(s) 2026

Abstract

Language is widely regarded as a uniquely human trait, rooted in mechanisms of social interaction such as conversational turn-taking. It has been proposed that turn-taking may not be exclusive to humans but may represent an ancestral trait shared across primates and beyond. Early studies on great apes indicated a specialization in gestural forms of turn-taking, whereas monkeys appeared to rely more heavily on vocal exchanges. However, this ape–monkey difference may simply be due to a lack of research. To bridge this gap, the present study investigated turn-taking behaviors during grooming in an Afro-Eurasian monkey species, sooty mangabeys (*Cercocebus atys*). We addressed the following three key questions: (1) What turn-taking patterns, if any, characterize grooming interactions in sooty mangabeys? (2) Do sooty mangabeys preferentially use vocal turn-taking? (3) Which elements characterizing turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking? We collected 105.3 h of video footage from 22 adult female sooty mangabeys in Taï National Park (between May 2022 and August 2023), yielding 428 female–female grooming videos, of which 71 were coded. An additional 595 grooming interactions were described directly in the field. Results showed that exchanges involving both actions and signals were the most frequent (e.g., a present gesture followed by grooming). The most common type of signal exchange was gestural turn-taking, a form previously reported only in great apes. The study identified key elements of conversational turn-taking: rapid responses, adjacency pair-like sequences, and participation frameworks. These findings suggest that the fundamental structural principles of human conversation likely predate our species, offering valuable insight into the evolutionary history of language.

Keywords Social interaction · Participation framework · Temporal relationships · Adjacency pairs · Evolution of language · Gestures

✉ Pharisemène Tibesar
ftibesar@uni-osnabrueck.de

¹ Comparative BioCognition, Institute of Cognitive Science, Osnabrück University, Artilleriestrasse 34, 49076 Osnabrück, Germany

² The Ape Social Mind Lab, Institut Des Sciences Cognitives Marc Jeannerod, CNRS, 67 Boulevard PinelBron, 69675 Lyon, France

³ Taï Chimpanzee Project, Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, Abidjan, Ivory Coast

⁴ Department of Human Behaviour, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

⁵ Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

⁶ Taï Monkey Project, Centre Suisse de Recherches Scientifique en Côte d'Ivoire, Abidjan, Ivory Coast

Introduction

Although it has intrigued scientists for centuries, the evolution of language still remains a mystery (Fitch 2002; Hewes et al. 1973; Knight et al. 2000). While the specific processes underlying language evolution are not fully understood (Bickerton 2016; Chomsky 2010), it is, however, widely accepted that language emerged relatively recently in human history (Diller & Cann 2009; Miyagawa et al. 2025; Nóbrega & Miyagawa 2015). This temporal proximity supports the hypothesis that the entire cognitive apparatus required for language could have developed based on already existing mechanisms and abilities (Hewes et al. 1973; Jackendoff 1999; Levinson & Holler 2014). Currently, the two predominant hypotheses suggest evolutionary precursors in either vocal (Hockett 1960; Lemasson et al. 2011; Locke & Hauser 1999) or gestural means (Arbib et

al. 2008; Call & Tomasello 2020; Corballis 1999; Hewes et al. 1973). Recently, however, Levinson (2016a, 2019) proposed that the evolution of language is deeply intertwined with a range of interactional abilities, collectively referred to as the interaction engine. This hypothesis suggests that human communication is evolutionarily stratified, with language emerging by using pre-existing abilities, such as joint attention, mutual gaze, and cooperative turn-taking. The latter stands out as representing one of the most ancient layers (Levinson & Holler 2014; Pika et al. 2018), with scholars suggesting a possible phylogenetic continuity between human social action during conversations and communicative interactions of other primate species and beyond (Abreu & Pika 2022; Levinson 2016b; Pika et al. 2018).

Turn-taking refers to the structured organization of participation in various activities, in which individuals take turns (Sacks et al. 1974). Turn-taking plays a crucial role across diverse contexts such as allocating political office, games, traffic management, and debates (Helbing et al. 2005; Natalia et al. 2019; Smith 2020). Turn-taking may involve either actions (e.g., children building a toy tower by alternating placing the blocks) or signals (e.g., speech) or a combination. Particular research attention has been dedicated to conversational turn-taking, which is part of the ‘speech exchange systems’ in humans (Sacks et al. 1974). Conversational turn-taking refers to the cooperative, rapid, and reciprocal exchange of signals, with turns being allocated between at least two participants (Sacks et al. 1974). While significant overlaps and gaps between turns are rare, conversational turn-taking is highly flexible, and interactants use specific turn-allocation techniques (e.g., vocal cues such as “go ahead” or pausing to signal the end of a turn), as well as repair mechanisms (e.g., clarification requests or self-corrections to address turn-taking errors or violations) (Sacks et al. 1974).

Turn-taking has been proposed to have preceded language in phylogeny due to several reasons (Levinson 2016b). First, although the over 6,700 languages so far described show a high degree of diversity (Comrie 2003; Pereltsvaig 2020), the involved turn-taking structures seem to exhibit strong universality (Stivers et al. 2009). Second, a study across 10 different languages revealed a consistent pattern in response offsets, with a unimodal peak occurring within 200 ms after a question was posed (Stivers et al. 2009). This is extremely fast, considering that it takes between 600 and 1000 ms to plan and produce a single-word response in a conversation (Bates et al. 2003; Indefrey and Levelt 2004). Hence, the processes of speech comprehension and production must overlap significantly to achieve such rapid response timing (Levinson 2016b; Levinson & Holler 2014). Third, turn-taking emerges very early in infancy, with babies at the age of three months already starting to coordinate smiles, gazes,

and facial expressions with their caregivers (e.g., Bates et al. 1975; Bateson 1975; Gratier et al. 2015). Fourth, evidence suggested that some form of turn-taking may be present in several nonhuman primate species (hereafter primates) (Fröhlich et al. 2016a, b; Lemasson et al. 2011; Levinson 2016b; Takahashi et al. 2013). This includes great apes such as bonobos (*Pan paniscus*; Fröhlich et al. 2016a; Levréro et al. 2019; Rossano 2013) and chimpanzees (*Pan troglodytes*; Fröhlich et al. 2016a); smaller apes such as lar gibbons (*Hylobates lar*; Terleph et al. 2022) and Mueller’s gibbons (*Hylobates muelleri*; Inoue et al. 2013); Afro-Eurasian monkeys such as Campbell’s monkeys (*Cercopithecus campbelli*; Lemasson et al. 2011) and Diana monkeys (*Cercopithecus diana*; Candiotti et al. 2012); and Neotropical monkeys such as common marmosets (*Callithrix jacchus*; Takahashi et al. 2013) and spider monkeys (*Ateles geoffroyi*; Briseño-Jaramillo et al. 2018). If this hypothesis is correct, investigating turn-taking can offer critical insights into the underlying structures and mechanisms that existed prior to the emergence of human language and contributed to its evolution (Levinson 2016b). One useful approach to this investigation is the comparative approach, which examines the behavior of other animals to then infer the abilities of our common ancestor and gain an understanding of the adaptations that enabled the evolution of behaviors, such as language (Hodos & Campbell 1969; Pika 2008; Zuberbühler 2005). Systematic and informed cross-species comparisons further refine this approach by enabling the reconstruction of evolutionary trajectories across lineages (Fitch 2002; MacLean et al. 2012; Pika et al. 2018). Comparisons with distantly related species provide complementary insights, as they allow the identification of patterns due to analogies, in which similar traits arise independently in taxa exposed to comparable selective pressures (Anderson et al. 2024; Lankester 1870). The reliance on comparative methodologies underscores the necessity of a comparative framework, as it allows for meaningful comparisons between studies and species by addressing differences in terminology and methods (Pika et al. 2018).

Evidence from nonhuman species suggested that some features characterizing human conversational turn-taking may be widespread across the animal kingdom. Some features have been documented in non-primate mammals, including African elephants (*Loxodonta africana*; Soltis et al. 2005), beluga whales (*Delphinapterus leucas*) (Morisaka et al. 2013), and roe deer (*Capreolus capreolus*) (Reby et al. 1999); in birds, such as robins (*Erithacus rubecula*) (Dabelsteen et al. 1997), starlings (*Sturnus vulgaris*) (Henry et al. 2015), and large-billed crows (*Corvus macrorhynchos*) (Kondo et al. 2010); in anurans, including green frogs (*Rana clamitans*) (Wells 1977) and Sri Lankan tree frogs (*Philautus leucorhinus*) (Arak 1983); and in insects, such

as North American katydids (*Amblycorypha parvipennis*), stoneflies (*Eucoptura xanthenes*), and tarbush grasshoppers (*Schistocerca obscura*; Bailey 2003) (Greenfield & Minckley 1993). These species all exhibit alternating vocal or auditory exchanges, the evolutionary basis of which remains unclear and may reflect shared evolutionary origins or convergent evolution (Pika et al. 2018). Although turn-taking is observed in many species, its structure is not uniform and may be species-specific, varying across multiple dimensions: form (e.g., flexibility of turn size), sequence (e.g., adjacency pair-like sequences), timing (e.g., inter-turn gaps), and distribution (e.g., turn allocation) (Pika et al. 2018; Sacks et al. 1974; Vlaeyen et al., submitted). In recent years, research interest in turn-taking has increased substantially, with researchers exploring vocal, gestural, and general turn-taking abilities in great apes (bonobos: Fröhlich et al. 2016a; Levréro et al. 2019; Rossano 2013; chimpanzees: Kolff & Pika, 2025; Pournault et al. 2021; van Boekholt & Pika, 2025; western lowland gorillas, *Gorilla gorilla gorilla*: Pournault et al. 2022). Concurrently, Afro-Eurasian monkeys have also been the subject of increased study. For instance, a study on adult Japanese macaques showed that, similarly to humans, individuals flexibly adjusted their response interval to the response latency of their interaction partners (Katsu et al. 2019). Likewise, studies on juvenile individuals of Japanese macaques and Campbell's monkeys (*Cercopithecus campbelli*) suggested that specific rules, such as call matching, needed to be learned to master turn-taking exchanges (Lemasson et al. 2011, 2013). Furthermore, red-capped mangabeys (*Cercocebus torquatus*) engaged in structured contact call exchanges, where overlapping calls were treated as violations, reflected by reduced listener attention (Meunier et al. 2024). These interaction patterns were further shaped by dominance rank and the level of social integration (Meunier et al. 2023). Based on existing evidence on turn-taking in great apes and monkeys, Levinson (2016b) proposed that vocal turn-taking may be ancestral within the primate order, with great apes having specialized in gestural rather than vocal turn-taking, a proposal hereafter referred to as the 'gestural specialization' hypothesis. Amid continuing debate about which modality most plausibly represents the evolutionary precursor of human language, identifying the modality underlying the emergence of turn-taking would therefore offer a significant contribution to this discussion (Arbib et al. 2008; Fröhlich et al. 2019; Pollick & De Waal 2007; Rodrigues et al. 2021). However, current knowledge of turn-taking in primates is limited by a research bias on several levels, including foci on experimental approaches, captive settings, vocalizations, specific features of conversational turn-taking, and particular species, especially common marmosets (see for an overview Pika et al. 2018). In monkeys specifically, gestural and

facial modalities have received virtually no research attention. Consequently, there is not yet enough systematic quantitative data to verify the gestural specialization hypothesis, and clarify the evolutionary trajectory of turn-taking abilities and specializations in primates (Pika et al. 2018; Schegluff 2007; but see van Boekholt & Pika 2025).

In studies of primate turn-taking, grooming emerged as a key interactional context for examining interactional roles and how these roles are alternated and negotiated (Henzi & Barrett 1999; Pika 2014). Grooming serves a central function in maintaining social bonds and group cohesion, yet its performance is constrained by its high time and energy demands (Dunbar 1991). The "grooming-at-a-distance" hypothesis proposed that vocal communication evolved to maintain social bonds when direct grooming is impractical (Arlet et al. 2015; Dunbar 1991). This hypothesis is supported by observations that primates with stronger social bonds engaged in more frequent vocal exchanges, suggesting that vocalizations complement grooming to sustain social relationships across space (Levréro et al. 2019). This situates grooming within the evolutionary trajectory of vocal interactions, making it an ideal context to study the emergence of language-related abilities, such as turn-taking (Dunbar 1991; Kolff & Pika 2025).

In this study, we aimed to gain insight into turn-taking skills in an Afro-Eurasian monkey species, the sooty mangabey (*Cercocebus atys*), living in their natural environment. Sooty mangabeys have been suggested to exchange calls, making them a useful candidate for studying turn-taking (Range & Fischer 2004). In addition, a closely related species, red-capped mangabeys, have been shown to communicate via gestures and employ them in intentional ways, with some similarities to great apes and other monkey species (Gupta & Sinha 2019; Molesti et al. 2020; Schel et al. 2022). Sooty mangabeys are a terrestrial member of the Cercopithecidae family and the Papionini tribe, which can be found along the west coast of Africa (Groves 1978, 2001). They live in multi-male and multi-female groups averaging 70 to 120 individuals, with females typically remaining in their natal group while males emigrate (Range & Noë, 2002). We specifically focused on social interactions of adult individuals since several studies have suggested that some elements of turn-taking systems may not be consistently exhibited in younger individuals (e.g., Abreu & Pika 2022; Lemasson et al. 2011; Takahashi et al. 2016). As in many primate species, grooming plays an important role in sooty mangabeys' social life, functioning not only to maintain hygiene but also to reinforce social bonds within the group (Dunbar 2010). For instance, a study conducted at the Taï National Park, Côte d'Ivoire, showed that individuals dedicated approximately 15% of their activity budget to grooming (Fruteau et al. 2011). Grooming interactions lasted an average of five

minutes and exhibited a high degree of within-interaction grooming reciprocity, with 97.8% of interactions involving bidirectional grooming (Fruteau et al. 2011). In addition, females typically focused their grooming efforts on two to four partners, who were often close in rank (Fruteau et al. 2011).

The present study addressed the following three key questions: (1) What turn-taking patterns, if any, characterize grooming interactions in sooty mangabeys? To address this question, we adopted the definition of Pika and colleagues (2018) (see also Vlaeyen et al., submitted), who view turn-taking as the structured and coordinated exchange of communicative signals or actions between two or more individuals, characterized by principles that regulate how turns are transferred and by temporal regularities. Signals and actions differ in that signals (e.g., facial expressions, gestures, or vocalizations) require a response from the recipient to achieve their intended outcome, whereas actions (e.g., grooming) can achieve their intended goal without eliciting a response from the recipient (Fröhlich et al. 2020; Genty et al. 2009; Hobaiter & Byrne 2011; Kolff 2025; van Boekholt 2024). We measured the presence of turn-taking by identifying the occurrence of turn transitions, which were defined as instances where one individual produced a unit (e.g., a vocalization, a gesture, a facial expression, or an action), and the next unit was initiated by the other individual (Sacks et al. 1974). Subsequently, to characterize the usage of turn-taking, we examined associations between the constituent units of turn transitions and quantified grooming reciprocity using a reciprocity index (Kolff 2025; Mitani 2009; Nishida 1988; van Boekholt et al. 2025). (2) Do sooty mangabeys preferentially use vocal turn-taking? In this study, we offered a first step to test the gestural specialization hypothesis (Levinson 2016b). To explore this question, we focused on vocalizations, facial expressions, and the production of gestures and tested whether sooty mangabeys engaged predominantly in facial turn-taking, gestural turn-taking, vocal turn-taking, or multimodal turn-taking (the first and second units of the turn transition are different modalities). (3) Which elements characterizing signal–signal turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking? Given that different types of turn transitions may follow distinct coordination rules, we focused on the type that may be the most comparable to humans: signal-to-signal exchanges. For example, during turn-taking in pool skateboard sessions, the current rider did not select the next, the content of a turn bore no relation to preceding or subsequent turns, the end of a turn was not predictable, and riders avoided overlapping due to safety considerations (Ivarsson & Greiffenhagen 2015). These properties contrast with conversational turn-taking, where next speakers are allocated through a set of practices (including selection by

the current speaker), turns are meaningfully connected to prior and subsequent turns, turn endings are often predictable at potential transition points, and overlap is handled through a more flexible set of coordination practices (Sacks et al. 1974). We examined three key features of human conversational turn-taking, following the framework of Pika et al. (2018):

- Element B, ‘who is taking the next turn,’ concerns the establishment of a participation framework (the initiator indicates who the intended recipient is and who will participate in the communicative exchange before producing a signal), and was assessed by investigating the parameters of gaze and body direction of the signaler during the production of signals used to initiate turn transitions and interactions.
- Element C, ‘when do response turns occur,’ refers to the temporal relationship between two turns. It was operationalized by calculating the intervals from the start of the initial unit to the start of the response, and from the end of the initial unit to the start of the response.
- Element D, ‘what should the next turn do,’ focuses on adjacency pairs and refers to two adjacent turns by participants, typically organized into first and second pair parts (e.g., question and answer, or a greeting and its response; Sacks et al. 1974; Schegloff 2007).

Methods

Study site and species

This study was conducted in the Taï National Park, Côte d’Ivoire, which spans 3,500 km² and is the largest remaining area of primary forest in West Africa (Boesch & Boesch 1983). This dense evergreen forest is home to a diverse range of large mammals, including African forest elephants (*Loxodonta cyclotis*), African forest buffaloes (*Syncerus caffer nanus*), and leopards (*Panthera pardus*), as well as 10 different primate species (Boesch & Boesch 1983). Research has been ongoing in the park for decades, involving a long-term study on western chimpanzees (*Pan troglodytes verus*) since 1979 (Boesch & Boesch 1983).

For this study, we investigated the behavior of individuals of a group of sooty mangabeys, which consisted of around 70 individuals at the start of the first field season. The group has been habituated to human presence since 2013 (Mielke et al. 2017), allowing for the reliable identification of the majority of group members and the collection of high-quality behavioral and video data. In addition, for the majority of individuals in the group, extensive long-term behavioral and genealogical data were available.

Data collection

Video data were collected by PT for a total of 12 months between May 2022 and August 2023 from 06:30 to 18:00. Continuous one-hour focal sampling was used on a total of 28 adult individuals (six males and 22 females) (Altmann 1974). Due to the small number of adult males and their frequent months-long absences during the study period, their data were not included in the subsequent analyses. Each focal individual was observed in a randomized order to prevent bias in the time of data collection. Following the completion of observations for all individuals, the sequence was reinitiated in a different randomized order. Particular attention was paid to ensuring systematic data collection across all subjects throughout each study month. One focal hour was counted if the focal individual had not been out of sight for more than 30 min in total, or 10 min consecutively. Using a digital camera (Sony Handycam FDR-AX100E) with an externally attached microphone (Sony ECM-CG60), recording of the focal animal took place at a distance of three to five meters when there was at least one other individual within five meters of the focal animal. Moreover, two trained research assistants collected direct observational behavioral data through half-day focal follows conducted from 06:30 to 18:00 (Altmann 1974). One observation day was counted when the group was followed for at least five hours.

Data coding

Video recordings were extracted to a computer (Dell Latitude 5420) and coded using the program ELAN (version 6.4) (Hellwig and Sloetjes 2021). For nine non-related females, we coded 10 to 11 dyadic grooming interactions involving another female partner (Table 1). A maximum of one interaction per dyad was coded to capture a diverse

Table 1 Demographic and observational data of the individuals as a function of name (as a three-letter code), the approximate age of each subject (in years), total observation hours per individual (in hours), and the total number of coded grooming interactions per individual

Focals	Approximate age (years)	Observation hours	Coded interactions
SON	15	11	10
BZA	10	15	10
GIS	5	14	11
MER	7	13	10
AMB	17	13	10
TAS	7	16	10
LUS	9	14	10
RAM	9	16	10
MAL	6	14	10

^aFrom here on, GESTURES are depicted in small capitals

range of social partners. Dyads were selected by prioritizing videos in which image quality, filming angle, and vegetation density provided optimal visibility of the behaviors performed during the interaction. Out of 22 adult females initially considered for data coding and analysis, the data of one female were excluded from the subsequent analysis due to her early disappearance during the data collection. This resulted in a total of 71 coded interactions, corresponding to approximately five hours of video footage. Of these, 27 interactions were complete, meaning that they captured both the beginning and the end of the grooming interaction. Behavioral definitions were based on behaviors described in ethograms of sooty mangabeys (Grampp et al. 2023; Range & Noë, 2002) and chimpanzees (Goodall 1986; Grund et al. 2024; Hobaiter & Byrne 2011; Pika 2014).

For each interaction, the following information was coded:

- Identity of involved individuals.
- Interaction number: Unique number for the interaction.
- Actions performed by both individuals: Socially directed embodied behavior that leads to the perceived goal through direct manipulation of another's body/body parts or the movement of one's own body (Kolff & Pika 2025; van Boekholt & Pika 2025).
- Potential gestures and facial expressions: Socially produced, mechanically ineffective movement of the body, face, or limbs that could potentially result in a response from the recipient (Gupta & Sinha 2019; Pika 2008).
- Vocalizations: Vocalizations produced by both individuals during the interaction, based on the vocal repertoire of Range and Fischer (2004).
- Gaze and body direction of involved individuals (see Figure SM 1) (Abreu & Pika 2022; Pika et al. 2003; van Boekholt & Pika 2025; Vlaeyen et al., submitted):
 - o Directed: angle from the midline (direction of the nose or nipples) and the body of the other individual ≤ 45 degrees
 - o Not directed: angle from the midline (direction of the nose or nipples) and the body of the other individual > 45 degrees from the midline

Eighteen percent of the videos were coded by a second observer blind to the study's hypothesis and research questions. A "substantial" to "almost perfect" agreement was obtained using ELAN's Cohen's Kappa coefficient from the "EasyDIAG" package using a 60% overlap (gestures: $K=0.95$, vocalizations: $K=0.77$, actions: $K=0.95$, gaze direction: $K=0.96$, body direction: $K=0.96$) (Holle & Rein 2015).

Behavioral operationalization & statistical analysis

Definition of grooming interactions and bouts

Following van Boekholt and Pika (2025), a grooming interaction was defined as an encounter between individuals that involved grooming. This began when one individual produced a signal or performed an action directed towards a recipient. The recipient could respond with a signal or action. The interaction ended when one individual moved away or when both individuals remained inactive for over 30 s (Aureli & Yates 2010; Goodall 1986; Kaburu & Newton-Fisher 2016; Newton-Fisher & Lee 2011). Grooming involved using one hand to move the partner's hair aside, while the other hand inspected and cleaned the exposed skin, sometimes using the mouth to remove scabs, parasites, or foreign material (Goodall 1986). A grooming bout was defined as a distinct period of grooming engagement in which two individuals participated in continuous grooming, with one consistently acting as the groomer and the other as the recipient without switching roles (Foster et al. 2009). Therefore, a grooming interaction could consist of multiple grooming bouts if the roles of groomer and recipient changed during the interaction.

What turn-taking patterns, if any, characterize grooming interactions in sooty mangabeys?

To investigate the presence of turn-taking, we analyzed turn transitions. Turn transitions were classified into four categories based on the nature of the first and second units (signal or action): signal–signal, signal–action, action–signal, and action–action. Signal–action turn transitions present more basic forms of interaction, commonly observed across many species (Belding's ground squirrel (*Urocitellus beldingi*), alarm whistle signal/flee-to-burrow response (Sherman 1977); jackdaw (*Corvus monedula*), allopreening solicitation posture/preening response (Katzir 1983)). By contrast, signal–signal turn transitions are used during conversational turn-taking and are seen in association with various conversation rules (e.g., overlap avoidance) (Sacks et al. 1974). Action–action turn transitions are also observed in humans, where they can share features of conversational turn-taking (Hofstetter 2021; Ivarsson & Greiffenhagen 2015; Sacks et al. 1974). For example, in board games, players manage turn transitions by monitoring when another's move is nearing completion, recognizing natural endpoints, and using subtle cues (such as stepping back) to signal that the next player may begin (Hofstetter, 2021a). Interactions can include one type of turn transition or a combination of several with variable numbers and proportions.

To evaluate whether turn transitions occurred in a pattern consistent with random behaviors, we conducted a Chi-Square Goodness-of-Fit test. We first used the overall proportion of actions and signals across the dataset, which were approximately balanced at 52% and 48% respectively. Assuming that each turn occurred independently of the previous one, we calculated the expected frequencies for each transition type. These expected proportions were then converted into expected counts. We compared these expected counts to the observed transition counts using the Chi-Square test implemented using the `chisq.test()` function, with the `rescale.p=TRUE` option to match the scale of the observed data. We then conducted post hoc pairwise comparisons using pairwise proportion tests with Bonferroni correction using the function `pairwise.prop.test()`, to determine which transition types occurred significantly more or less often than others.

We used Multiple Distinctive Collocation Analysis (MDCA) with the program `Coll.analysis 3.2a` (Gries 2007), to measure the association between two units in a given turn transition. MDCA is a technique developed in linguistics to explore how often two specific words appear together (for review, see Gries 2013) and has recently also been applied to studies of animal communication (Bosshard et al. 2022; Kolff 2025). MDCA evaluates all possible unit combinations to see if they occur more or less frequently than chance would predict (Gries & Stefanowitsch 2004). Because it uses the binomial probability mass function, MDCA is particularly effective for analyzing skewed, non-random, and small datasets, which are often found in studies of animal communication (Bosshard et al. 2022; Gries & Stefanowitsch 2004; Leroux et al. 2021). The log-transformed results of MDCA (p_{bin}) indicate the degree of exclusivity between combinations, with positive values showing a higher-than-expected occurrence and negative values indicating a lower-than-expected occurrence (Bosshard et al. 2022; Gries and Stefanowitsch 2004).

Within-interaction grooming reciprocity, that is, the alternation of grooming roles during a single interaction, has been documented in several primate species and represents a potential source of turn-taking (Chancellor & Isbell 2009; Fruteau et al. 2011; Newton-Fisher & Lee 2011). To assess whether it played a role in shaping turn-taking in sooty mangabeys, grooming reciprocity within interactions was quantified with a Grooming Reciprocity Index (GRI) (Mitani 2009; Nishida 1988). In this index, g_{AB} represents the duration of grooming provided by individual A to individual B, and g_{BA} refers to the grooming duration from individual B to individual A. The sum of $g_{AB} + g_{BA}$ refers to the total grooming time between the two individuals. The index ranges from zero (indicating no reciprocity) to one (indicating complete reciprocity).

$$\text{GRI} = 1 - \left| \frac{g_{AB}}{g_{AB} + g_{BA}} - \frac{g_{BA}}{g_{AB} + g_{BA}} \right|$$

Do sooty mangabeys preferentially use vocal turn-taking?

Since gestures and facial expressions have not yet been studied in the context of grooming in sooty mangabeys, we first created a repertoire of the gestures and facial expressions produced by the focal animals in the observation period (but see Grampp et al. 2023). We defined gestures as movements of the extremities or head and body postures that can be directed to a social partner, are mechanically ineffective (produced to request or reach a goal, requiring active participation of the interactant), and potentially elicit a response by the recipient of the gesture (Pika 2008). Only the gestures and facial expressions that had been produced at least twice by two individuals or three times by one individual were included in the repertoire (Hobaiter & Byrne 2011; Pika et al. 2003). The repertoire was created by using the “verb first” principle (Nishida et al. 1999), as well as using, when possible, terms from previous gesture work on mangabeys and great apes (Aychet et al. 2021; Grampp et al. 2023; Pika et al. 2003, 2005; Schel et al. 2022). To ensure a comprehensive measurement of the entire gestural repertoire, we constructed a graph plotting repertoire size against the total number of observed gestures, aiming to reach an asymptote (Fröhlich et al. 2016b; Genty et al. 2009; Hobaiter & Byrne 2011). We assessed signal–signal turn transitions, which are cases where gestures, facial expressions, or vocalizations produced by one individual are directly followed by gestures, facial expressions, or vocalizations produced by another individual. They were categorized as facial, gestural, vocal, or multimodal turn transitions. To determine whether a dominant modality emerged in the distribution of turn transitions, we conducted a Chi-Square Goodness-of-Fit test comparing the observed counts of each category against an expected uniform distribution, where each category was assumed to occur with equal frequency (25%). We then performed pairwise comparisons using a **pairwise proportion test** with Bonferroni correction to identify which specific categories differed significantly from each other.

Which elements characterizing signal–signal turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking?

We first investigated Element C by measuring the response time for each signal–signal turn transition. The onset-to-onset timing was measured as the time interval between the start of one unit by one individual and the start of the next unit produced by another individual (Kolff & Pika 2025; van Boekholt & Pika 2025; Vlaeyen et al., submitted). Although

less commonly used (but see Fischer et al. 2000; Kolff & Pika 2025; van Boekholt & Pika 2025), it offers the advantage of circumventing methodological discrepancies related to gesture duration assessment (Grund et al. 2024). Offset-to-onset timing was measured as the interval between when the first unit produced by one individual ended and when the next unit produced by another individual began (Lemasson et al. 2013; Stivers et al. 2009; Takahashi et al. 2013; Vlaeyen et al., submitted). Concerning gestures, the offset was set when the gesture reached its final form, before its retraction, thus excluding the optional holding phase, which likely adds no semantic value (De Vos et al. 2015; Grund et al. 2024). When comparing signal–signal turn transitions to other types of turn transitions, the offset of actions was similarly defined as the moment when the movement required to achieve the intended goal ceased, excluding any subsequent retraction or holding movements (Vlaeyen et al., submitted). To determine whether a majority of turn transitions exhibited overlap avoidance, we performed a one-sided binomial test using the `binom.test()` function. We tested whether the observed number of turn transitions showing overlap avoidance was significantly greater than what would be expected by chance. The null hypothesis assumed that overlap avoidance occurred in half of the transitions, while the alternative hypothesis predicted a greater proportion, indicating a dominant tendency to avoid overlap. Additionally, we quantified the typical response window by identifying the time period during which the number of observed turn transitions exceeded the number of observed turn transitions for unlikely responses, defined as instances in which the second unit was potentially unrelated to the preceding unit (Lemasson et al. 2018; Levréro et al. 2019). This window thus reflects the timeframe in which responses are most likely to be genuine. Subsequently, we compared timings across turn transition types and communication modalities, enabling a deeper understanding of the factors shaping turn-taking timing. Differences were assessed using Kruskal–Wallis tests. Additionally, the overlap rate for each turn transition type and modality of signal–signal turn-taking was calculated. Differences across turn transition types were evaluated using Chi-squared tests, while differences across modalities of signal–signal turn-taking were assessed using Fisher’s exact test to account for the small sample size and ensure accurate inference.

Element B was assessed by measuring gaze and body direction parameters of signals used during signal–signal turn transitions. Within the 10 s preceding a signal being produced, we measured whether the gaze and body direction of the emitter were directed or not directed towards the recipient (Abreu & Pika 2022; Pika et al. 2003). We used one-tailed exact binomial tests to determine whether communicative behaviors were more likely directed toward

the recipient (via gaze or body orientation) (50%). Three separate tests were conducted: one for gestures initiating signal–signal turn transitions, one for gestures initiating interactions, and one for vocalizations. No test was conducted for facial expressions, as none were observed in the species' repertoire for the grooming context. Vocalizations initiating interactions were likewise not subjected to statistical analysis due to an insufficient number of observations, with only four instances recorded. To evaluate whether the proportion of signals accompanied by directed gaze and body orientation differed across signal types, we conducted proportion tests for overall comparisons. Post hoc pairwise comparisons using Fisher's exact test were performed with Bonferroni correction to evaluate differences between the three signal categories: gestures initiating signal turn transitions, gestures initiating interactions, and vocalizations.

To evaluate Element D, we identified adjacency pair-like sequences. We used Multiple Distinctive Collocation Analysis (MDCA) to identify non-randomly associated signal pairs. To compare the timing of non-randomly associated signal pairs with all other events, we excluded the non-random pairs from the full dataset. We then performed a one-sided Wilcoxon rank-sum test to assess whether the offset-to-onset timings of non-random pairs were shorter (i.e., faster) than those of the remaining events.

All statistical analyses were performed using R version 4.3.2. Statistical significance was set at an alpha level of 0.05.

Results

Gestural repertoire

To allow further analyses, we compiled the repertoire of facial expressions and gestures for the focal individuals. We coded a total of 444 potential gestures and facial expressions, with 426 of them being subsequently classified into eight different gesture types (see Table 2). All gestures were either visual (generating a mainly visual component with no physical contact) or tactile (involving contact with the recipient) (Pika et al. 2003); no audible gestures (producing sound) or facial expressions met the criteria to be included in the repertoire. Based on the cumulative frequency of gesture types employed in the grooming context, it appeared that the group's repertoire had reached an asymptote (Figure SM 2). This suggested that the full repertoire had been observed (Hobaiter & Byrne 2011; Pika 2007).

Table 2 Gestural repertoire of our study individuals in the context of grooming during the observation period, including the definition of each gesture, the number of occurrences, and the number of individuals observed performing the gestures ^a(Hobaiter & Byrne 2011; Pika et al. 2003, 2005; Schel et al. 2022)

Gesture type	Definition ^a	Sensory modality	Number of occurrences	Number of individuals
EMBRACE ^a	The signaler hugs the recipient by wrapping one or both arms or legs around him/her	Tactile	17	11
EXTEND LIMB	The signaler stretches out a limb (arm or leg) horizontally toward or within the attentional field of the receiver	Visual	32	12
MAINTAIN CONTACT	The signaler stops grooming, but one or both hands remain in touch with the body of the recipient for at least one second	Tactile	66	16
PRESENT	The signaler moves his/her body or body part to expose an area (excluding fingers, hand, toes, or foot) to the receiver's attention and maintains that position	Visual	220	21
PULL	The signaler's hand or foot is firmly clasped over a part of the recipient's body or a handful of hair and softly pulled toward the signaler	Tactile	18	10
PUSH	The signaler makes contact with the recipient's body, typically with a hand or a foot, exerting force	Tactile	13	10
RAISE	The signaler lifts a body part, such as a hand, arm, or leg, in a generally vertical movement, often pausing briefly near the top. A lifted hand involves the elbow not being above the shoulder; otherwise, it is considered lifting the arm	Visual	51	18
TOUCH	The signaler lightly comes in contact with the recipient's body, typically with fingers, knuckles, hand, or foot	Tactile	16	10

Descriptive data

Direct behavioral observations allowed us to document a total of 595 grooming interactions among the 22 females. This dataset was used to assess the daily grooming rate.

For the video dataset, we conducted a total of 319.55 h of observation across 22 females (mean \pm SD = 14.52 \pm 1.55 h per female), yielding 105.3 h of video footage (mean \pm SD = 4.79 \pm 0.75 h per female). This resulted in 428 female–female grooming videos (mean \pm SD = 19.45 \pm 8.04 per female), of which 71 interactions were selected for coding. All other analyses were performed using this dataset.

Focal sooty mangabeys engaged in grooming an average of 2.21 times per day. Grooming interactions involved 1.34 \pm 0.65 bouts on average (mean \pm SD). The maximum of three bouts was observed in seven interactions. The 27 complete interactions lasted 95 \pm 98 s on average (mean \pm SD). In 90.77% of interactions with a recorded ending (59/65), the interaction ended with one individual moving away from the recipient by at least three meters. In the remaining cases, one individual was chased by a third individual, or both individuals stopped interacting for more than 30 s while remaining in close proximity.

In the coded videos, a total of 1,075 units were recorded, comprising 557 actions (mean \pm SD = 57.15 \pm 16.54% per dyad), 426 gestures (mean \pm SD = 36.65 \pm 13.93% per dyad), and 92 vocalizations (mean \pm SD = 6.20 \pm 12.09% per dyad). No facial expressions were observed (Tables SM 1 and SM 2, Figure SM 3).

What turn-taking patterns, if any, characterize grooming interactions in sooty mangabeys?

A total of 92.96% of all observed interactions involved at least one turn transition, with an average of 7.86 \pm 6.47 turn transitions per interaction (mean \pm SD). In the remaining interactions, all behaviors were performed by a single individual, with the recipient passively receiving grooming and showing no further response. In total, 519 turn transitions were identified and analyzed. Of these, 55 transitions (10.40%) were classified as action–action, 165 (31.79%) as action–signal, 215 (41.61%) as signal–action, and 84 (16.19%) as signal–signal. Based on the overall behavior rates, the expected proportions were 27.04% for action–action, 24.96% for both action–signal and signal–action, and 23.04% for signal–signal. The observed distribution of turn transitions differed significantly from what would be expected under random usage ($\chi^2 = 128.56$, $df = 3$, $p = 2.2 \times 10^{-16}$). This deviation indicated a structured, non-random pattern in how actions and signals are sequenced during turn transitions. Both mixed transitions, signal–action and action–signal, occurred significantly more often

than action–action transitions ($p = 9 \times 10^{-3}$ and $p < 2 \times 10^{-16}$, respectively). Signal–signal transitions were significantly less frequent than signal–action ($p < 2 \times 10^{-16}$) and action–signal transitions ($p = 3.6 \times 10^{-8}$), but did not differ significantly from action–action transitions ($p = 0.06$). The average percentages of turn transitions per dyad were 13.12 \pm 21.61% (mean \pm SD) for action–action, 27.56 \pm 18.34% (mean \pm SD) for action–signal, 38.75 \pm 22.49% (mean \pm SD) for signal–action, and 13.53 \pm 17.96% (mean \pm SD) for signal–signal (Figure SM 4).

We found that 25 turn transitions were associated at a frequency higher than expected by chance. Among these, 15 were observed at least three times, thereby representing the most relevant pairs (Fig. 1 and Table SM 3) (Mielke & Carvalho 2022). For the remaining ten combinations, their relevance in the study individuals' grooming interactions could not be fully established due to a relatively low production in the current dataset. The highest association was observed between the action grooming and the gesture PRESENT. Interestingly, grooming was strongly negatively associated with itself, similarly to PRESENT.

The average within-interaction grooming reciprocity index was 0.12 \pm 0.25 (mean \pm SD), indicating a low rate of reciprocity, and 23.94% of interactions included role reversals.

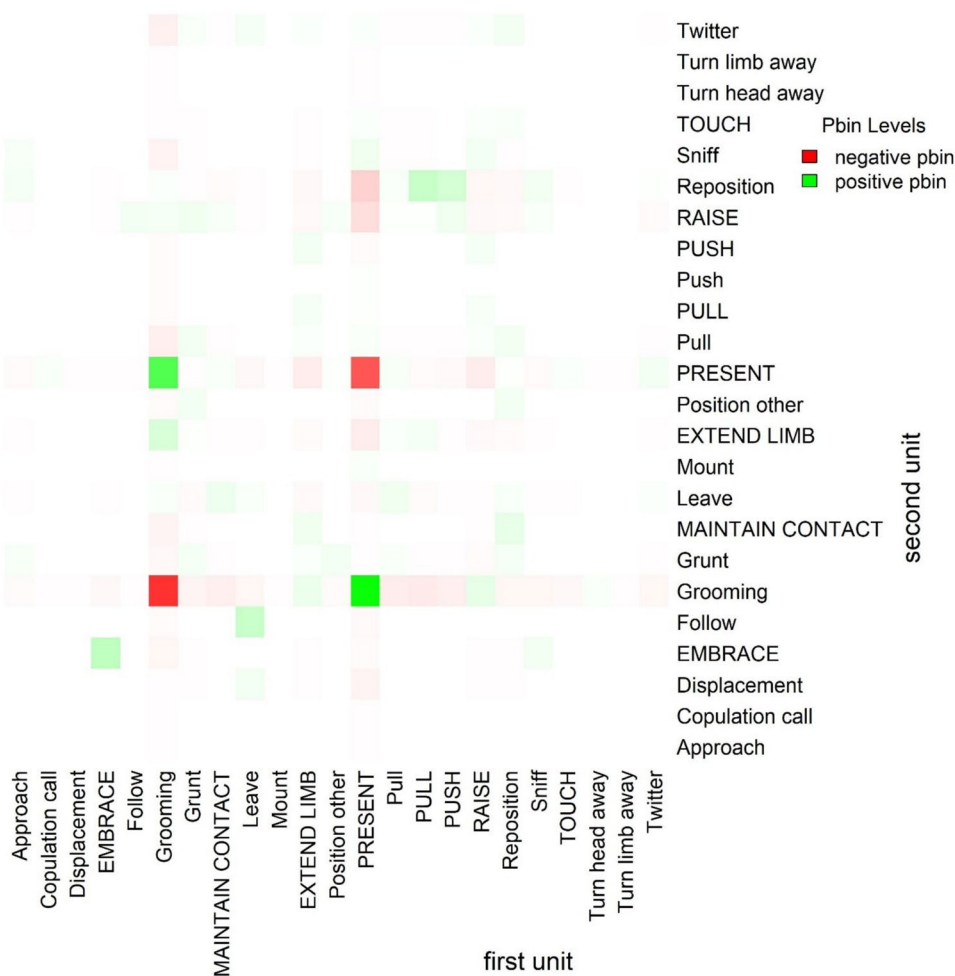
Do sooty mangabeys preferentially use vocal turn-taking?

Signal–signal turn transitions were observed in 52.11% of dyads. Among these, 60.71% were gestural, 3.57% were vocal, 0% were facial, and 35.71% were multimodal (gesture and vocalization in either order) (Figure SM 5). We observed a highly significant difference from the expected equal distribution ($\chi^2 = 83.143$, $df = 3$, $p < 2 \times 10^{-16}$). Gestural turn transitions were used significantly more than multimodal ($p = 0.01$), vocal turn transitions ($p = 4.9 \times 10^{-14}$), and facial turn transitions ($p < 2.2 \times 10^{-16}$). Multimodal turn transitions were used more frequently than vocal turn transitions ($p = 2.7 \times 10^{-6}$) and facial turn transitions ($p = 3.1 \times 10^{-8}$). No significant differences were found between vocal and facial turn transitions ($p = 1.00$).

Which elements characterizing signal–signal turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking?

Examining element C, 'when do response turns occur,' we found that onset-to-onset gaps of signal–signal turn transitions were highly variable, ranging from 0.009 s to 109.85 s (mean \pm SD = 9.16 \pm 18.86 s, median = 1.52 s) (Fig. 2). Similarly, offset-to-onset gaps spanned from -2.97 s to 109.61 s

Fig. 1 Matrix with columns and rows representing the first and second units within turn transitions, respectively. Positive pbin values in green represent higher-than-expected occurrences, negative values in red indicate lower-than-expected occurrences, and zero values appear in white. Color intensity reflects the magnitude of the pbin values, with more opaque colors indicating higher magnitudes (range: -22.43 to 27.03). Gestures are indicated in all capital letters



(mean ± SD = 7.72 ± 19.09 s, median = 0.09 s) (Fig. 2). Negative offset-to-onset gaps, representing overlapping turn transitions, occurred in 41.67% of cases, which was not significantly different from 50% ($p = 0.156$), suggesting that overlap avoidance did not dominate in this dataset. To identify the typical response times, we analyzed histograms of turn transitions. A horizontal threshold line was drawn that represents how often unlikely response turn transitions or uncommon response intervals occur (Lemasson et al. 2018; Levréro et al. 2019). Intervals above this line fell inside the typical range. Based on this criterion, responses generally occurred within 1 s of the offset of another signal or within 4 s of the onset of a signal. The typical median onset-to-onset was 0.68 s (mean ± SD = 1.08 ± 0.99 s) and the typical median offset-to-onset was -0.09 s (mean ± SD = -0.32 ± 0.75 s). Within the range of typical responses, 61.40% of offset-to-onset turn transitions overlapped, with a mean overlap duration of 0.71 s.

Signal-signal turn-taking exhibited shorter onset-to-onset durations compared to other transition types (all transitions: median = 2.89 s; action-action: median = 8.32 s; action-signal: median = 14.51 s; signal-action: median = 1.59 s) (Figure

SM 6). The difference between the timing of these turn transitions was significant ($\chi^2 = 131.93$, $df = 3$, $p < 2 \times 10^{-16}$). Signal-signal turn-taking exhibited shorter offset-to-onset durations compared to other transition types (all transitions: median = 0.35 s; action-action: median = 1.18 s; action-signal: median = 0.55 s; signal-action: median = 0.13 s) (Figure SM 7). However, these differences were not statistically significant ($\chi^2 = 6.277$, $df = 3$, $p = 0.099$). The rate of overlap was higher for signal-signal transitions (41.67%) compared to action-action (29.63%) and signal-action (37.5%), and similar to action-signal (41.21%) transitions, but these differences were likewise not statistically significant ($\chi^2 = 1.250$, $df = 3$, $p = 0.741$).

Gestural offset-to-onset turn transitions had a mean duration of 5.82 s and a median of 0.09 s, with an overlap rate of 45.1%. Multimodal turn transitions were longer, with a mean of 11.69 s and a median of 0 s, and an overlap rate of 33.3%. Vocal turn transitions were shorter, with a mean of 0.47 s and a median of 0 s, but exhibited the highest overlap rate at 66.7%. However, the differences in timing across modalities were not significant ($\chi^2 = 3.040$, $df = 2$, $p = 0.219$), nor were the differences in overlap rates ($p = 0.405$).

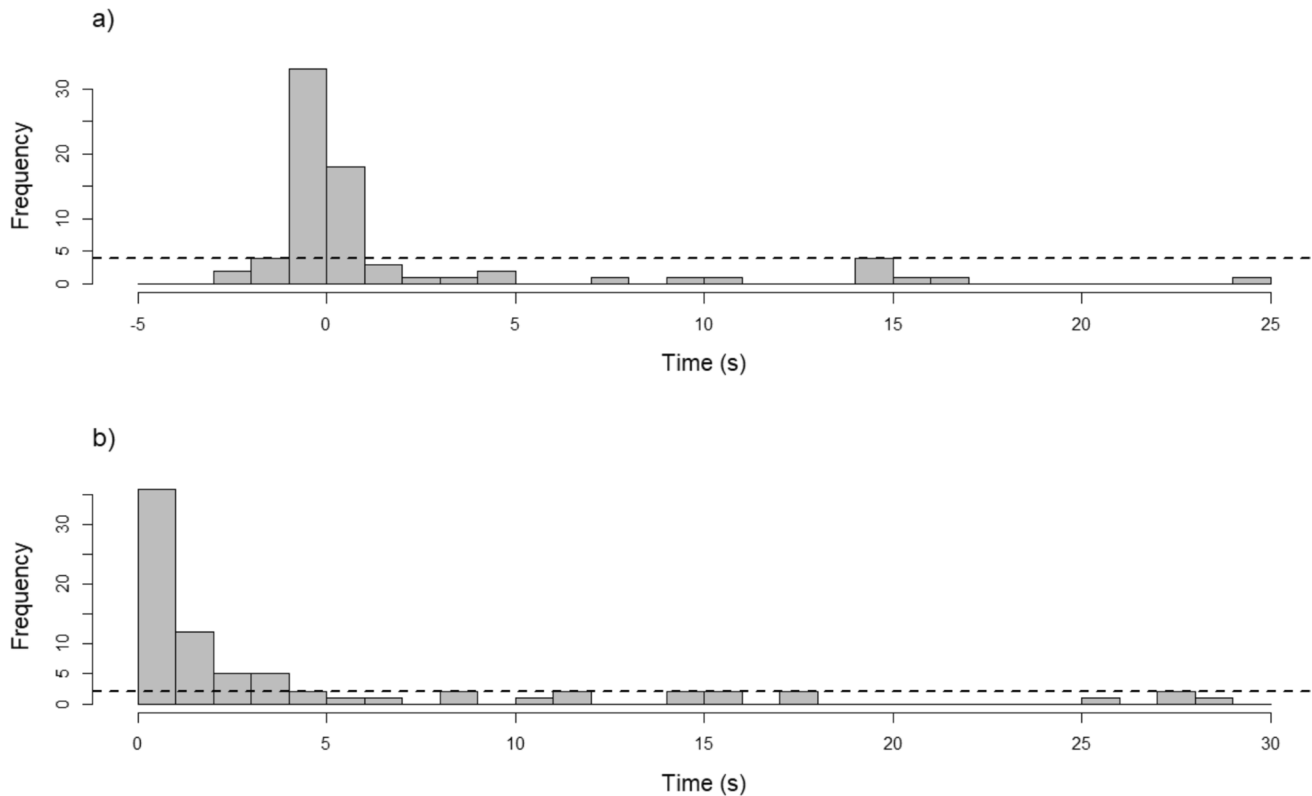


Fig. 2 Overview of the temporal relationships of turns for signal-signal turn transitions with (a) showing the distribution of offset-to-onset gaps, and (b) depicting the distribution of onset-to-onset gaps. To

improve readability, the plot was limited to the range from -5 s to 25 s for (a) and 0 s to 30 s for (b)

For element B, ‘who is taking the next turn,’ directed gaze accompanied gestures that initiated signal-signal turn transitions in 56.45% of cases ($35/62$, $p=0.1871$), which did not differ significantly from chance. Directed body orientation occurred in 67.74% of gestures that initiated signal-signal turn transitions ($42/62$, $p=3.57 \times 10^{-3}$), a proportion significantly above chance. Focusing specifically on gestures that initiated interactions, directed gaze was observed in 79.32% of cases ($23/29$, $p=1.16 \times 10^{-3}$), and directed body orientation in 75.86% ($22/29$, $p=4.07 \times 10^{-3}$), both significantly above chance. For vocalizations that initiated signal-signal turn transitions, directed gaze accompanied the vocalization in 88.24% of cases ($15/17$, $p=1.18 \times 10^{-3}$), and directed body orientation occurred in 76.47% ($13/17$, $p=2.45 \times 10^{-2}$), both exceeding chance levels. Comparing directed gaze across these signal types revealed a significant overall difference ($p=1.33 \times 10^{-2}$). Post hoc pairwise comparisons did not reach significance (signal-signal gestures vs. initiating interaction gestures, $p=0.12$; signal-signal gestures vs. vocalizations, $p=6.54 \times 10^{-2}$; initiating interaction gestures vs. vocalizations, $p=1.00$), likely due to reduced statistical power from splitting the data and applying Bonferroni correction. For directed body orientation, differences across

these signal types were not significant ($p=0.64$), indicating similar orientation rates across categories.

For element D, ‘what should the next turn do,’ we found that three signal pairs were not randomly associated (PUSH—RAISE, grunt—RAISE, EMBRACE—EMBRACE). The 11 non-randomly associated turn transitions had a median offset-to-onset timing of -0.24 s, compared with 0.09 s for all other events excluding these non-random pairs, a difference that was statistically significant, indicating that non-random pairs occurred faster ($p=0.035$). Notably, these pairs exhibited a high overlap rate of 81.82%.

Discussion

This study aimed to contribute to testing the hypothesis that turn-taking constitutes an ancient layer of human language, conserved across all major primate clades, reflecting phylogenetic continuity (Levinson 2016b; Pika et al. 2018). To address this objective, we investigated turn-taking interactions in female sooty mangabeys living in their natural environment with a special focus on the cooperative context of grooming. We examined three key questions: (1) What turn-taking patterns, if any, characterize grooming interactions in

sooty mangabeys? (2) Do sooty mangabeys preferentially use vocal turn-taking? and (3) Which elements characterizing turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking?

What turn-taking patterns, if any, characterize grooming interactions in sooty mangabeys?

We found that nearly all grooming interactions in sooty mangabeys involved turn-taking. These findings mirror recent results of a study on turn-taking in male Eastern chimpanzees (*Pan troglodytes schweinfurthii*) showing that more than eighty percent of all grooming interactions included at least one turn transition (Kolff 2025). These findings suggest that turn-taking could be a widespread feature in grooming interactions among these two species. This behavior helps to mediate grooming interactions either to enable even distribution of benefits and balance the roles of grooming and being groomed, to negotiate change in grooming location on the body, or both.

Like chimpanzees, the interactions of the sooty mangabeys studied here also combined all types of turn transitions (Kolff 2025; van Boekholt & Pika 2025). This interactional pattern could represent a transitional stage toward signal-based conversational turn-taking, a feature considered foundational in the evolution of language (Levinson 2016b). Nevertheless, differences between chimpanzees and sooty mangabeys were evident. We found that action–signal and signal–action transitions dominated sooty mangabeys’ grooming interactions, whereas in adult male chimpanzees, action–action and signal–action transitions were most frequent, with signal–signal occurring at comparable rates in both species (Kolff 2025). Turn-taking patterns in mother–infant chimpanzees differed further, with all types of turn transitions used at balanced rates during food-sharing, grooming, joint travel, and nursing (van Boekholt & Pika 2025). Thus, sooty mangabeys predominantly relied on signal–action turn transitions, whereas adult chimpanzees more frequently employed action–action turn transitions, which in humans share some, but not all, organizational properties with conversational turn-taking (Hofstetter 2021; Ivarsson & Greiffenhagen 2015). One possible explanation for these differences is that, unlike in chimpanzees, grooming by individual A in mangabeys was rarely immediately followed by grooming from individual B (action–action) (Kolff 2025). Role reversals were relatively rare overall, and when they occurred, they were often facilitated by communicative signals (action–signal followed by signal–action). These patterns may reflect species-specific social dynamics and communication styles: chimpanzees more frequently engage in mutual and reciprocal grooming, which requires little signaling, whereas sooty mangabeys rely more heavily

on communicative signals to coordinate the less common grooming role reversals (Fröhlich et al. 2016a). Another contributing factor may be that sooty mangabeys have a larger number of grooming partners relative to chimpanzees (Newton-Fisher & Lee 2011; Watts 2000). This may lead to lower familiarity with each partner and necessitate more explicit communicative signals to coordinate grooming, such as indicating when to change grooming location. This interpretation is supported by the frequent use of gestures aimed at presenting a specific body part for grooming, including PRESENT, EXTEND LIMB, and RAISE. However, this hypothesis remains tentative, as evidence from male Eastern chimpanzees suggests that the social bond strength does not significantly influence the type of turn transition employed during grooming (Kolff & Pika 2025). Overall, this cross-species comparison suggests that action–action turn transitions in grooming interactions may be favored in groups where reciprocal grooming is common and expectations are predictable. By contrast, lower partner familiarity, and infrequent role reversals appear to increase reliance on communicative signals, thereby enhancing the prevalence of action–signal exchanges. However, since the present study is based on a single sooty mangabey group, further work is needed to determine how broadly these findings apply across the species. Additionally, direct cross-species comparisons using standardized methods are necessary to clarify how social and ecological factors shape patterns of turn transition use during grooming coordination and across other interactional contexts. This will allow for the identification of the mechanisms that facilitate signal–signal turn transitions and conversation-like interactions, and clarify their role in the evolutionary origins of conversational turn-taking and language.

This study revealed that two action–signal pairs, grooming–PRESENT and PRESENT–grooming were the predominant pairs observed. This pattern indicates that communicative exchanges during grooming interactions in our study individuals were predominantly focused on the location of grooming (“groom here”) rather than on the negotiation of grooming roles (“you groom me”) and grooming tasks (“your turn now”). A total of 15 combinations were frequently associated at a rate higher than predicted by chance. Some were variations of PRESENT–grooming, such as RAISE–grooming and EXTEND LIMB–grooming, while others involved changes in body posture, including PULL–reposition and PUSH–RAISE.

Similarly, in the study on adult Eastern chimpanzees, nine non-randomly associated signal–action pairs were observed during grooming, four of which were identical to those found in sooty mangabey grooming interactions: PRESENT–grooming, RAISE–grooming, PULL–reposition, and PUSH–reposition (Kolff 2025). These similarities suggest

that in both species, in grooming interactions, turn-taking is used to facilitate coordinated exchanges, specifically, requests for grooming and changes in the recipient's body position. Further data and comparative studies are needed to clarify the role of exchanges produced during turn-taking in grooming interactions and beyond, providing insight into the contexts in which turn-taking occurs, the types of interactions in which it first emerged, and how it may have evolved.

Do sooty mangabeys preferentially use vocal turn-taking?

Concerning question two, we found that our study individuals used two different modalities during signal–signal turn-taking: gestural, and vocal. They combined them in multimodal turn transitions. Although the female mangabeys showed some instances of vocal turn-taking using grunt and twitter calls, only gestural adjacency pair-like sequences were observed. Whereas vocal turn-taking has been suggested to exist in sooty mangabeys (Range & Fischer 2004) and other Afro-Eurasian monkey species (Katsu et al. 2019; Lemasson et al. 2011; Richman 1976), investigations into gestural signaling were, until now, very limited (but see, Gupta & Sinha 2019; Laidre 2011; Molesti et al. 2020). Our study reveals that, in this specific context of grooming, gestural and multimodal signaling were the most crucial means, emphasizing the need to gain a more detailed understanding of non-vocal communicative modalities beyond the great ape species. Our results also challenge the hypothesis that great apes are the taxa specialized in gestural rather than vocal turn-taking, suggesting instead that gestural turn-taking may represent an early-emerging feature of turn-taking across the primate lineage (Levinson 2016b). Amid persistent debate over the modality in which language originated, this study provides support for the view that gestural communication played a crucial role in the emergence of language-linked cognitive and communicative capacities (Arbib et al. 2008; Fröhlich et al. 2019; Pollick & De Waal 2007; Rodrigues et al. 2021). In chimpanzees, grooming and play contexts preferentially involved gestural signals, whereas feeding interactions were dominated by vocalizations (Fröhlich et al. 2016c, 2020; Hobaiter et al. 2017). Modality choice was shaped by factors such as individual matrices, mechanical constraints, social circumstance, the potential cost of misinterpretation, and the degree of privacy afforded by different signal types (Fröhlich et al. 2016c, 2017; Hobaiter et al. 2017). These findings highlight the importance of examining multiple interactional contexts to characterize broader patterns of modality use for turn-taking within a species. Accordingly, systematic comparative studies examining the frequency and complexity of gestural and vocal turn-taking

across apes and other primate species in various contexts are needed to fully evaluate this hypothesis, and such data are currently lacking (but see Southern et al., submitted).

Which elements characterizing signal–signal turn-taking patterns of sooty mangabeys are shared with human conversational turn-taking?

Regarding question three, we examined three hallmark features of human conversational turn-taking with the aim of comparing them across primate taxa. This comparison aimed to identify shared characteristics and to test the interaction engine hypothesis, which proposes that these features emerged through homology within the primate lineage.

With regard to element C, 'when do response turns occur,' we found that despite the wide range of response times, the studied sooty mangabeys interacted within a relatively narrow window of a few seconds. Such specific timing likely plays a critical role in enabling both the signaler and the recipient to recognize that they are engaged in a coordinated interaction. This temporal framework may thus aid in clearly delineating the boundaries, beginnings, and endings of turns, thereby facilitating smoother coordination during grooming.

Compared to humans, the typical median signal–signal offset-to-onset found in this study (-0.09 s) was faster than the timing reported in a study investigating conversational turn-taking across several different languages (0.10 s) (Stivers et al. 2009). Turn-taking timing in sooty mangabeys was more variable, evidenced by a comparatively higher standard deviation in typical offset-to-onset timing (Stivers et al. 2009). Turn-taking during grooming interactions did not involve overlap avoidance, contrary to conversational turn-taking (Sacks et al. 1974). The overlap rate of typical answers was one and a half to four times higher than that observed in human question–answer interactions across spoken languages and sign languages (Beukeleers & Lepout 2023; De Vos et al. 2015; Heldner 2011). However, it is important to note that turn-taking timing in humans can be influenced by factors such as age and social relationships (Beaumont & Cheyne 1998; Dewi & Munir 2018; Sprecher & Treger 2015). Investigating these variables in future research may enable more meaningful comparisons and deepen our understanding of the evolution of turn-taking dynamics.

To date, no turn-taking studies have been conducted on sooty mangabeys. In red-capped mangabeys, typical response times for the exchange of contact calls fell within 2 s, although overlapping exchanges were rare, in contrast to what we observed (Meunier et al. 2023). The reported average response time of 0.5 s is slower than our observed typical average of -0.32 s, although this difference

may result from the study's exclusion of overlapping calls (Meunier et al. 2023). Compared to other studies on Afro-Eurasian monkeys, the typical signal–signal offset-to-onset timing observed in this study was faster than that reported for vocal turn-taking in Japanese macaques and Campbell's monkeys. The authors reported that overlapping calls were extremely rare (Katsu et al. 2019; Lemasson et al. 2010). This difference may be due to inherent distinctions between communication modalities. It has been proposed that vocal communication involves less overlap than visual gestural communication because overlapping vocalizations can lead to perceptual masking and production difficulties, while overlapping visual gestures are typically easier to perceive simultaneously without such interferences (Coates & Sutton-Spence 2001; Emmorey et al. 2009). Some gestures observed in this study, such as PULL or TOUCH, included a tactile component. Research has shown that the somatosensory system has a limited capacity to process multiple tactile inputs simultaneously (Kusnir et al. 2023). Even so, issues with overlap between turns would likely arise only when individuals are both receiving and producing tactile gestures without visual contact, an uncommon scenario, as it would be difficult to accurately touch someone without seeing them. Our findings did not provide evidence that gestural turn transitions permit greater overlap than vocal turn transitions. However, the low frequency of vocal turn transitions constrained the interpretability of these comparisons. Further studies in Afro-Eurasian monkeys are needed to determine whether communicative modality influences turn-taking timing. Additionally, the present study did not exclude the preparation phase of gestures from the analysis. It has been argued that these phases should be omitted (McCleary & Leite 2024) because, in vocal communication, pre-vocal preparations such as in-breaths and pre-speech tongue movements are usually not included in response-time measures (Palo et al. 2014). Excluding gestural preparation phases might therefore have produced response times more comparable to those reported for vocal exchanges.

When contrasted with results from studies on great apes, the general median offset-to-onset timing in sooty mangabeys (0.35 s) was faster than that observed in joint travel between mother-infant chimpanzees and bonobos, which was close to 1 s (Fröhlich et al. 2016a). However, it was closer to that reported for adult male Eastern chimpanzees during grooming interactions, who exhibited no delay between turns (Kolff & Pika 2025). These differences suggest that developmental factors may influence the timing of turn-taking in chimpanzees (Fröhlich et al. 2016a; Lemasson et al. 2013). Similarly to chimpanzees, overlap avoidance, a key feature of human conversational turn-taking, was not observed during grooming interactions in sooty mangabeys (Kolff 2025). This suggests that sooty

mangabeys anticipated the goal of the signaler and initiated their response before the signal had ended (Badihi et al. 2024; Kolff 2025). Despite slight differences, the offset-to-onset timings of turn-taking observed in our study were rapid, similar to those observed in other Afro-Eurasian monkeys, great apes, and humans, further contributing to the growing body of evidence supporting ancestral quick turn-taking abilities in the primate lineage (Abreu & Pika 2022; Levinson 2016b; Pika et al. 2018).

Onset-to-onset timing is typically not presented in studies on humans or Afro-Eurasian monkeys. Relative to great apes, the observed general median onset-to-onset timing (2.89 s) was longer than that reported for mother-infant and adult chimpanzees, whose median timings were closer to two seconds (Kolff 2025; van Boekholt & Pika 2025). This difference may be attributed to the nature of the grooming context, where units are structured to be relatively long, while the study on mother-infant chimpanzees also focuses on the contexts of food-sharing, joint travel, and nursing (van Boekholt & Pika 2025). Another explanation could be that due to their bigger neocortex and involved neurons, chimpanzees may require less pre-processing time than sooty mangabeys, enabling them to initiate responses quicker (Herculano-Houzel 2017). This interpretation supports the broader hypothesis that chimpanzees possess higher cognitive processing abilities than Afro-Eurasian monkeys (Byrne 1995; Matsuzawa 2010).

Regarding element B, 'who is taking the next turn,' we found that gaze and body direction were used to initiate interactions and during signal–signal turn transitions, suggesting that these cues play a role in coordinating interactions. Specifically, we observed that directed body orientation was significantly present in signal–signal initiating gestures, interaction-initiating gestures, and vocalizations, whereas directed gaze was significantly present only for interaction-initiating gestures and vocalizations. The role of directed eye gaze during vocal turn-taking remains largely unexplored. Our study provided evidence that vocalizations used during social interactions were frequently accompanied by a directed gaze toward the recipient. Research on chimpanzees has shown that in the alarm context, they similarly use directed gaze specifically during certain vocalizations, such as Alarm Hoots and Waa Barks, although this behavior has not been documented during turn-taking interactions (Schel et al. 2013).

When comparing with human communication, we observed that in both species, directional cues play a role in managing turn-taking. Human language is hypothesized to have evolved primarily for face-to-face interactions, with gaze playing a central role (Kendon 1967; Levinson 2016b). Eye gaze is not only used to signal the end of a turn but also to manage repair and interruptions (Brône et al. 2017; Ho et

al. 2015; Kendon 1967; Rossano, 2009, 2013; for review, see Degutye & Astell 2021). However, the use of directed gaze varies depending on the type of interaction. For instance, questions are more likely to be accompanied by a directed gaze toward the recipient, whereas free-flowing conversation or storytelling often involves less eye contact (Eberhard & Nicholson 2010; Kendon 1967; Rossano et al. 2009). Cultural differences further influence these dynamics, with studies reporting that in some cultures, questions are less likely to be asked while looking at the recipient than in others (Rossano et al. 2009). Research on gaze direction duration during free-flowing conversations remains limited; nevertheless, one study found that Japanese speakers maintained eye contact with their conversational partner only around a quarter of the time while speaking (Ijuin et al. 2018). Further studies could investigate whether, as in humans, contextual and social factors influence gaze behavior during turn-taking in sooty mangabeys. Such research may enable holistic comparisons and contribute to our understanding of the evolution of participation framework establishment in turn-taking.

Concerning findings on other Afro-Eurasian monkeys, our results showed that just over half of the gestures used to initiate signal–signal turn transitions involved directed gaze. This contrasts with previous studies reporting that about ninety percent of gestures in captive olive baboons and nearly all gestures observed in captive red-capped mangabeys were produced with directed gaze toward the recipient (Molesti et al. 2020; Schel et al. 2022). Differences between the studies may be due to the use of different methodologies, with the two above-mentioned studies focusing on other contexts, such as affiliation, aggression, sex, and social play, rather than only the grooming context. In a context where both interactants are in body contact and where body positions are constrained, such as grooming, checking the attentional state of the recipient before signaling could be less common than in other contexts. This finding should be considered in future research examining gaze and body direction in the grooming context (for example, for first-order intentionality). Furthermore, we argue that gaze may play a different role with regard to the function and intended meaning of the accompanying signal. For instance, if an individual “opens” the interaction and “conversation” (Pillet-Shore 2018), then directed signals may be crucial to emphasize this request and the transferred meaning. In contrast, in an ongoing and already established interaction, directed signals might not be needed to transfer the information and reach the desired goal. This hypothesis is supported by our data, which showed that direct eye gaze was employed primarily at the initiation of interactions rather than continuously throughout.

In great ape studies, some researchers classify a behavior as a gesture only if it includes markers of intentional signaling such as goal persistence and directed gaze (Cartmill & Byrne 2010; Genty et al. 2009; Hobaiter & Byrne 2011; Pollick & De Waal 2007). This approach excludes gestures that occur within an already established interaction and no longer require directed gaze (e.g., tactile gestures used to request a change in the recipient’s position), or that are immediately successful, such that response waiting or goal persistence is no longer necessary. However, some studies differentiated between intentionally produced gestures showing more than one key characteristic of intentional communication and gestures showing only one key characteristic of intentionality (Fröhlich et al. 2016a), enabling a more comprehensive understanding of gestural production and usage. They also showed that in mother-infant dyads of chimpanzees and bonobos, signalers used directed eye gaze and body orientation to initiate joint travel in fewer than sixty percent of cases, a higher rate than that observed in sooty mangabeys (Fröhlich et al. 2016a). Notably, species- and age-related differences were apparent: bonobos relied more on directed eye gaze than chimpanzees, while older infants of both species tended to use body orientation more frequently than younger ones (Fröhlich et al. 2016a). These findings suggested that gaze direction during turn-taking may be subject to developmental and inter-species variations, which may help explain the discrepancies between our findings and those reported in this previous study. Despite varying rates, great apes and the sooty mangabeys of the present study used directional cues during turn-taking, suggesting that this element of the conversational turn-taking system may have ancient evolutionary roots across the primate lineage.

The analysis of element D, ‘what should the next turn do,’ identified three signal sequences akin to adjacency pair-like sequences, which were both more frequently associated than expected by chance and faster than other types of turn transitions.

When compared to humans, our study revealed that grooming interactions in sooty mangabeys were similarly organized into signal adjacency pair-like sequences, which occurred faster than other turn transitions (Roberts et al. 2015; Sacks et al. 1974). This indicated that signal exchanges were not random or incidental, but instead formed a characteristic feature of mangabeys’ grooming interactions. We also found that adjacency pair-like sequences showed a high rate of overlap, which stands in contrast to the general avoidance of overlap documented for human conversation (Sacks et al. 1974). This pattern may reflect the nature of the specific pairs involved. The PUSH–RAISE pair (e.g., individual A gently pushing individual B’s arm, followed by individual B lifting that arm) resembles agreement sequences, which are known to display increased overlap

(Pomerantz 1984; Vatanen 2018). The EMBRACE–EMBRACE pair (i.e., mutual hugging) is comparable to reciprocal greeting sequences, where overlap is also reported (Pillet-Shore 2012). Because EMBRACE involves tactile contact, additional mechanical constraints may further contribute to overlap. Further work will be useful to determine whether adjacency pair-like sequences in sooty mangabeys and other primate species reliably exhibit elevated overlap and to identify the mechanisms that produce this pattern. In this study, we did not assess the intentionality or communicative meaning of the signals involved, which would strengthen the claim that the second unit functioned as a contingent response to the first (Schegloff 2007). Addressing this limitation should be a key objective for future research seeking to identify and analyze structured patterns of communicative exchange.

Considering Afro-Eurasian monkeys, studies demonstrating adjacency pair-like sequences remain rare. However, Lemasson and colleagues (2011) highlighted call matching in Japanese macaques, suggesting a conversational rule in which vocalizations were answered with matching calls. Although this indicated that vocal interactions in Japanese macaques were governed by structured rules, similar evidence for other modalities and species remains lacking. Regarding great apes, our results were consistent with observations of adult Eastern chimpanzees and mother-infant chimpanzees, where adjacency pair-like sequences have also been documented (Fröhlich et al. 2016a, b; Kolff 2025; van Boekholt et al. 2025). However, since some studies focused exclusively on gesture–response pairs, direct comparisons of the modalities and units used are not possible (Fröhlich et al. 2016a, b; Kolff 2025). Our results contrast with findings on interactions of mother-infant chimpanzees, where signal–signal adjacency pair-like sequences have not been observed; instead, only either signal–action (e.g., OPEN ARM–nursing) or action–action pairs (e.g., nursing–grooming) have been described (van Boekholt et al. 2025). Further research, particularly involving adult individuals, is necessary to determine whether such patterns emerge at later developmental stages, as age has been shown to influence the production and use of gestures, and gesture–response pairs (Fröhlich et al. 2016a; Pika 2007).

To date, mother-infant ape interactions have served as the primary model for investigating gestural turn-taking (Fröhlich et al. 2016a; Rossano 2013). However, our findings suggested that systematic studies into the behaviour of Afro-Eurasian monkeys living in their natural environments may offer valuable insights and could serve as promising subjects for advancing research in this domain.

Taken together, the present findings add another piece of evidence to support the hypothesis that turn-taking is an ancient mechanism already present across the whole primate lineage (Levinson 2016b, 2019; Levinson & Holler 2014;

Pika et al. 2018). More specifically, these abilities may have already been present in the last common ancestor of humans and catarrhine monkeys, which lived approximately 20 to 30 million years ago (Raum et al. 2005; Steiper & Young 2006). This is consistent with the idea that turn-taking constitutes one of the earliest layers of human communication.

Conclusion

This study explored turn-taking patterns in grooming interactions among adult female sooty mangabeys living in a single wild population. The findings strengthened the view that turn-taking is a fundamental feature of primates' social and cooperative exchanges. Importantly, this research provided evidence of gestural turn-taking in Afro-Eurasian monkeys, underscoring the importance of expanding comparative studies to the wide diversity of monkey species. Future studies will be important to test the gestural specialization hypothesis (Levinson 2016b). Comparisons with temporal features and the organization of interactions into adjacency pair-like sequences revealed parallels with humans and great apes. Similarly, the use of directional cues aligned with communication patterns in both humans and non-human primates. These similarities suggested that the roots of conversational turn-taking could extend deep into primate evolution. Continued comparative research will be essential to unravel how the building blocks of language have evolved in the primate lineage.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10071-025-02040-2>.

Acknowledgements We are grateful to the Ministère de l'Enseignement supérieur et de la Recherche scientifique, the Ministry of Water and Forests of Côte d'Ivoire, and the Office Ivoirien des Parcs et Réserves for granting permission to conduct this research. We sincerely thank the Centre Suisse de Recherches Scientifiques en Côte d'Ivoire and the team at the Taï Chimpanzee Project for their generous logistical and institutional support. PT and SP extend their appreciation to R. M. Wittig and C. Crockford for granting them access to the field site, for supporting field administration, and data collection. Special thanks go to Simon Kannieu and Constant Kaye, the dedicated field assistants involved in the long-term data monitoring. Finally, we are thankful to all members of the Comparative BioCognition Research Group for their insightful discussions, and constructive feedback, and administrative support throughout this project.

Author contributions S.P.: Idea, Conceptualization, funding acquisition, methodology, supervision, project administration, writing and shaping of manuscript, writing—review & editing P.T.: Conceptualization, funding acquisition, data curation, data analysis, investigation, methodology, Writing—original draft, Writing—review & editing C.C.: Data curation, Writing—review & editing A.L.F.: Data curation, Writing—review & editing.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded by an EU-Consolidator grant (772000, TurnTaking) to SP of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program. Additional funding was provided to PT by the Leakey Foundation (Grant 64083009).

Data availability The data generated by this study are available at <http://doi.org/10.6084/m9.figshare.29304983>.

Declarations

Competing interests The authors declare no competing interests.

Ethics approval This study was strictly non-invasive and observational, adhering fully to ethical standards and guidelines set by the involved research institutions. It adhered to the guidelines set forth by the United Kingdom's Animals (Scientific Procedures) Act of 1986, as well as the ethical standards for the treatment of non-human primates outlined by the American Society of Primatologists. This study was categorized as a non-animal experiment under the German Animal Welfare Act of May 25, 1998 (Section V, Article 7), and therefore did not require approval from an ethics committee. To minimize the risk of disease transmission, observers followed strict hygiene protocols, including a five-day quarantine period and wearing face masks during data collection. All applicable international, national, and institutional regulations regarding animal care and use were strictly followed.

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

References

- Abreu F, Pika S (2022) Turn-taking skills in mammals: A systematic review into development and acquisition. *Front Ecol Evol* 10:1168. <https://doi.org/10.3389/fevo.2022.987253>
- Altmann J (1974) Observational study of behavior: Sampling methods. *Behaviour* 49(4):227–267. <https://doi.org/10.1163/156853974X00534>
- Anderson NK, Preininger D, Fuxjager MJ (2024) Physiological basis of convergent evolution in animal communication systems. *Integr Comp Biol* 64(5):1422–1436. <https://doi.org/10.1093/icb/icae091>
- Arak A (1983) Vocal interactions, call matching and territoriality in a Sri Lankan treefrog, *Philautus leucorhinus* (Rhacophoridae). *Anim Behav* 31:292–302. [https://doi.org/10.1016/S0003-3472\(83\)80199-7](https://doi.org/10.1016/S0003-3472(83)80199-7)
- Arbib MA, Liebal K, Pika S (2008) Primate vocalization, gesture, and the evolution of human language. *Curr Anthropol* 49(6):1053–1076. <https://doi.org/10.1086/593015>
- Arlet M, Jubin R, Masataka N, Lemasson A (2015) Grooming-at-a-distance by exchanging calls in non-human primates. *Biol Lett* 11(10):20150711. <https://doi.org/10.1098/rsbl.2015.0711>
- Aureli F, Yates K (2010) Distress prevention by grooming others in crested black macaques. *Biol Lett* 6:27–29. <https://doi.org/10.1098/rsbl.2009.0513>
- Aychet J, Blois-Heulin C, Palagi E, Lemasson A (2021) Facial displays in red-capped mangabeys (*Cercocebus torquatus*): Repertoire, social context, and potential intentionality. *J Comp Psychol* 135:98–113. <https://doi.org/10.1037/com0000252>
- Badihi G, Graham KE, Grund C, Safryghin A, Soldati A, Donnellan E, Hobaiter C (2024) Chimpanzee gestural exchanges share temporal structure with human language. *Curr Biol* 34(14):R673–R674. <https://doi.org/10.1016/j.cub.2024.06.009>
- Bailey WJ (2003) Insect duets: underlying mechanisms and their evolution. *Physiol Entomol* 28(3):157–174. <https://doi.org/10.1046/j.1365-3032.2003.00337.x>
- Bates E, Camaioni L, Volterra V (1975) The acquisition of performatives prior to speech. *Merrill-Palmer quarterly of behavior and development should be Merrill-Palmer Quarterly of Behavior and Development*, 21(3):205–226. <http://www.jstor.org/stable/23084619>
- Bates E, D'Amico S, Jacobsen T, Székely A, Andonova E, Devescovi A et al (2003) Timed picture naming in seven languages. *Psychon Bull Rev* 10(2):344–380. <https://doi.org/10.3758/bf03196494>
- Bateson MC (1975) Mother-infant exchanges: The epigenesis of conversational interaction. *Ann N Y Acad Sci* 263:101–113. <https://doi.org/10.1111/j.1749-6632.1975.tb41575.x>
- Beaumont SL, Cheyne JA (1998) Interruptions in adolescent girls' conversations: Comparing mothers and friends. *J Adolesc Res* 13(3):272–292. <https://doi.org/10.1177/0743554898133003>
- Beukeleers I, Lepeut A (2023) Who's got the upper hand?: A cross-linguistic study on overlap in VGT and LSF. *Belg J Linguist* 36:212–246. <https://doi.org/10.1075/bjl.00076.beu>
- Bickerton D (2016) Language and human behavior. University of Washington Press. <https://doi.org/10.1515/9780295801049>
- Boesch C, Boesch H (1983) Optimisation of nut-cracking with natural hammers by wild chimpanzees. *Behaviour* 83(4):265–286. <https://doi.org/10.1163/156853983X00192>
- Bosshard AB, Leroux M, Lester NA, Bickel B, Stoll S, Townsend SW (2022) From collocations to call-ocations: Using linguistic methods to quantify animal call combinations. *Behav Ecol Sociobiol* 76(9):1–8. <https://doi.org/10.1007/s00265-022-03224-3>
- Briseño-Jaramillo M, Ramos-Fernandez G, Palacios-Romo TM, Sosa-López JR, Lemasson A (2018) Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behav Ecol Sociobiol* 72(12):192. <https://doi.org/10.1007/s00265-018-2615-2>
- Brône G, Oben B, Jehoul A, Vranjes J, Feyaerts K (2017) Eye gaze and viewpoint in multimodal interaction management. *Cognitive Linguistics* 28(3):449–483. <https://doi.org/10.1515/cog-2016-0119>
- Byrne RW (1995) Primate cognition: Comparing problems and skills. *Am J Primatol* 37(2):127–141. <https://doi.org/10.1002/ajp.1350370206>
- Call J, Tomasello M (eds) (2020) *The gestural communication of apes and monkeys*. Psychology Press. <https://doi.org/10.4324/9781003064541>
- Candiotti A, Zuberbühler K, Lemasson A (2012) Convergence and divergence in Diana monkey vocalizations. *Biol Lett* 8(3):382–385. <https://doi.org/10.1098/rsbl.2011.1182>
- Cartmill EA, Byrne RW (2010) Semantics of primate gestures: intentional meanings of orangutan gestures. *Anim Cogn* 13(6):793–804. <https://doi.org/10.1007/s10071-010-0328-7>
- Chancellor RL, Isbell LA (2009) Female grooming markets in a population of gray-cheeked mangabeys (*Lophocebus albigena*). *Behav Ecol* 20:79–86. <https://doi.org/10.1093/beheco/arn117>

- Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In: *The evolution of human language: Bio-linguistic perspectives*, pp. 45–62. Cambridge University Press. <https://doi.org/10.1017/cbo9780511817755.003>
- Coates J, Sutton-Spence R (2001) Turn-taking patterns in deaf conversation. *J Socioling* 5(4):507–529. <https://doi.org/10.1111/1467-9481.00162>
- Comrie B (ed) (2003) *The Major Languages of South Asia, the Middle East and Africa*. Taylor & Francis.
- Corballis MC (1999) The gestural origins of language. *Am Sci* 87(2):138–145. <https://doi.org/10.1511/1999.20.810>
- Dabelsteen T, McGregor PK, Holland JO, Tobias JA, Pedersen SB (1997) The signal function of overlapping singing in male robins. *Anim Behav* 53(2):249–256. <https://doi.org/10.1006/anbe.1996.0369>
- Degutye Z, Astell A (2021) The role of eye gaze in regulating turn taking in conversations: a systematized review of methods and findings. *Front Psychol* 12:616471. <https://doi.org/10.3389/fpsyg.2021.616471>
- De Vos C, Torreira F, Levinson SC (2015) Turn-timing in signed conversations: Coordinating stroke-to-stroke turn boundaries. *Front Psychol* 6:268. <https://doi.org/10.3389/fpsyg.2015.00268>
- Dewi RF, Munir A (2018) Turn-taking strategies and its relation to EFL learners' personality and power in the interaction of English conversation class. *Eng Teach Learn Res J* 4(2):288–305. <https://doi.org/10.24252/Eternal.V42.2018.A12>
- Diller KC, Cann RL (2009) Evidence against a genetic-based revolution in language 50,000 years ago. In: *The cradle of language*, Vol. 12, pp. 135–149. <https://doi.org/10.1093/oso/9780199545858.003.0007>
- Dunbar RI (1991) Functional significance of social grooming in primates. *Folia Primatol* 57(3):121–131. <https://doi.org/10.1159/000156574>
- Dunbar RI (2010) The social role of touch in humans and primates: Behavioural function and neurobiological mechanisms. *Neurosci Biobehav Rev* 34(2):260–268. <https://doi.org/10.1016/j.neubiorev.2008.07.001>
- Eberhard K, Nicholson H (2010). Coordination of understanding in face-to-face narrative dialogue. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* Vol. 32, No. 32. <https://escholarship.org/uc/item/6bk8h630>
- Emmorey K, Bosworth R, Kraljic T (2009) Visual feedback and self-monitoring of sign language. *J Mem Lang* 61(3):398–411. <https://doi.org/10.1016/j.jml.2009.06.001>
- Fischer J, Cheney DL, Seyfarth, RM (2000) Development of infant baboons' responses to graded bark variants. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1459):2317–2321. <https://doi.org/10.1098/rspb.2000.1285>
- Fitch WT (2002) The evolution of language comes of age. *Trends Cogn Sci* 6(7):278–279. [https://doi.org/10.1016/s1364-6613\(02\)01925-3](https://doi.org/10.1016/s1364-6613(02)01925-3)
- Foster MW, Gilby IC, Murray CM, Johnson A, Wroblewski EE, Pusey AE (2009) Alpha male chimpanzee grooming patterns: Implications for dominance “style.” *Am J Primatol* 71(2):136–144. <https://doi.org/10.1002/ajp.20632>
- Fröhlich M, Kuchenbuch P, Müller G, Fruth B, Furuichi T, Wittig RM, Pika S (2016a) Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci Rep* 6(1):1–14. <https://doi.org/10.1038/srep25887>
- Fröhlich M, Müller G, Zeiträg C, Wittig RM, Pika S (2017) Gestural development of chimpanzees in the wild: The impact of interactional experience. *Animal Behaviour*, 134(Special Issue: Communicative Complexity). 271–282. <https://doi.org/10.1016/j.anbehav.2016.12.018>
- Fröhlich M, Müller G, Zeiträg C, Wittig RM, Pika S (2020) Begging and social tolerance: food solicitation tactics in young chimpanzees (*Pan troglodytes*) in the wild. *Evol Hum Behav* 41(2):126–135. <https://doi.org/10.1016/j.evolhumbehav.2019.11.002>
- Fröhlich M, Sievers C, Townsend SW, Gruber T, van Schaik CP (2019) Multimodal communication and language origins: integrating gestures and vocalizations. *Biol Rev* 94(5):1809–1829. <https://doi.org/10.1111/brv.12535>
- Fröhlich M, Wittig RM, Pika S (2016b) Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. *R Soc Open Sci*. <https://doi.org/10.1098/rsos.160278>
- Fröhlich M, Wittig RM, Pika S (2016c) Play-solicitation gestures in chimpanzees in the wild: Flexible adjustment to social circumstances and individual matrices. *Royal Soc Open Sci* 3(8):160278. <https://doi.org/10.1098/rsos.160278>
- Fruteau C, Lemoine S, Hellard E, van Damme E, Noë R (2011) When females trade grooming for grooming: testing partner control and partner choice models of cooperation in two primate species. *Anim Behav* 81(6):1223–1230. <https://doi.org/10.1016/j.anbeha.2011.03.008>
- Genty E, Breuer T, Hobaiter C, Byrne RW (2009) Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim Cogn* 12(3):527–546. <https://doi.org/10.1007/s10071-009-0213-4>
- Goodall J (1986) *The chimpanzees of Gombe: Patterns of behavior*. Harvard University Press. <https://ci.nii.ac.jp/naid/10010164813>
- Gramp M, Samuni L, Girard-Buttoz C, León J, Zuberbühler K, Tkaczynski P, Crockford C (2023) Social uncertainty promotes signal complexity during approaches in wild chimpanzees (*Pan troglodytes verus*) and mangabeys (*Cercocebus atys atys*). *R Soc Open Sci* 10(11):231073. <https://doi.org/10.1098/rsos.231073>
- Gratier M, Devouche E, Guellai B, Infanti R, Yilmaz E, Parlato-Oliveira E (2015) Early development of turn-taking in vocal interaction between mothers and infants. *Front Psychol* 6:1374. <https://doi.org/10.3389/fpsyg.2015.01167>
- Greenfield MD, Minckley RL (1993) Acoustic dueling in tarbush grasshoppers: settlement of territorial contests via alternation of reliable signals. *Ethology* 95(4):309–326. <https://doi.org/10.1111/j.1439-0310.1993.tb00480.x>
- Gries ST (2007) Coll. analysis 3.2 a. A program for R for Windows 2
- Gries STh (2013) 50-something years of work on collocations: What is or should be next.... *Int J Corpus Linguist* 18:137–166. <https://doi.org/10.1075/ijcl.18.1.09gri>
- Gries STh, Stefanowitsch A (2004) Extending collocation analysis: A corpus-based perspective on ‘alternations.’ *Int J Corpus Linguist* 9:97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>
- Groves CP (1978) Phylogenetic and population systematics of the mangabeys (primates: Cercopithecoidea). *Primates* 19:1–34. <https://doi.org/10.1007/BF02373224>
- Groves CP (2001) *Primate taxonomy*. Smithsonian Institution, pp 1–350
- Grund C, Badihi G, Graham KE, Safryghin A, Hobaiter C (2024) GesturalOrigins: a bottom-up framework for establishing systematic gesture data across ape species. *Behav Res Methods* 56(2):986–1001. <https://doi.org/10.3758/s13428-023-02082-9>
- Gupta S, Sinha A (2019) Gestural communication of wild bonnet macaques in the Bandipur National Park Southern India. *Behav Process* 168:103956. <https://doi.org/10.1016/j.beproc.2019.103956>
- Helbing D, Schönhof M, Stark HU, Holyst JA (2005) How individuals learn to take turns: emergence of alternating cooperation in a congestion game and the prisoner's dilemma. *Adv Complex Syst* 8:87–116. <https://doi.org/10.1142/S0219525905000361>
- Heldner M (2011) Detection thresholds for gaps, overlaps, and no-gap-no-overlaps. *J Acoust Soc Am* 130:508–513. <https://doi.org/10.1121/1.3598457>

- Hellwig B, Sloetjes H (2021) ELAN - linguistic annotator. <https://www.mpi.nl/corpus/html/elan/index.html>
- Henry L, Craig AJFK, Lemasson A, Hausberger M (2015) Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Front Psychol* 6:1416 <https://doi.org/10.3389/fpsyg.2015.01416>
- Henzi SP, Barrett L (1999) The value of grooming to female primates. *Primates* 40:47–59. <https://doi.org/10.1007/BF02557701>
- Herculano-Houzel S (2017) Numbers of neurons as biological correlates of cognitive capability. *Curr Opin Behav Sci* 16:1–7. <https://doi.org/10.1016/j.cobeha.2017.02.004>
- Hewes GW, Andrew RJ, Carini L, Choe H, Gardner RA, Kortlandt A et al (1973) Primate communication and the gestural origin of language [and comments and reply]. *Curr Anthropol* 14(1/2):5–24. <https://doi.org/10.1086/201401>
- Hobaiter C, Byrne RW (2011) The gestural repertoire of the wild chimpanzee. *Anim Cogn* 14(5):745–767. <https://doi.org/10.1007/s10071-011-0409-2>
- Hobaiter C, Byrne RW, Zuberbühler K (2017) Wild chimpanzees' use of single and combined vocal and gestural signals. *Behav Ecol Sociobiol* 71(6):96. <https://doi.org/10.1007/s00265-017-2325-1>
- Hockett CF (1960) The origin of speech. *Sci Am* 203(3):88–97. <https://doi.org/10.2307/24940617>
- Hodos W, Campbell CBG (1969) *Sacra naturae: Why there is no theory in comparative psychology*. *Psychol Rev* 76(4):337–350. <http://s://doi.org/10.1037/h0027523>
- Hofstetter E (2021) Achieving preallocation: Turn transition practices in board games. *Discourse Process* 58(2):113–133. <https://doi.org/10.1080/0163853X.2020.1816401>
- Holle H, Rein R (2015) EasyDIAg: A tool for easy determination of interrater agreement. *Behav Res Methods* 47(3):837–847. <https://doi.org/10.3758/s13428-014-0506-7>
- Ho S, Foulsham T, Kingstone A (2015) Speaking and listening with the eyes: Gaze signaling during dyadic interactions. *PLoS ONE* 10(8):e0136905. <https://doi.org/10.1371/journal.pone.0136905>
- Ijuin K, Umata I, Kato T, Yamamoto S, Nakayama J (2018) Difference in eye gaze for floor apportionment in native- and second-language conversations. *J Nonverbal Behav*. <https://doi.org/10.1007/s10919-017-0262-3>
- Indefrey P, Levelt WJM (2004) The spatial and temporal signatures of word production components. *Cognition* 92(1–2):101–144. <https://doi.org/10.1016/j.cognition.2002.06.001>
- Inoue Y, Sinun W, Yosida S, Okanoya K (2013) Intergroup and intragroup antiphonal songs in wild male Mueller's gibbons (*Hyllobates muelleri*). *Interact Stud* 14:24–43. <https://doi.org/10.1075/is.14.1.03ino>
- Ivarsson J, Greiffenhagen C (2015) The organization of turn-taking in pool skate sessions. *Res Lang Soc Interact* 48(4):406–429. <https://doi.org/10.1080/08351813.2015.1090114>
- Jackendoff R (1999) Possible stages in the evolution of the language capacity. *Trends Cogn Sci* 3(7):272–279. [https://doi.org/10.1016/S1364-6613\(99\)01333-9](https://doi.org/10.1016/S1364-6613(99)01333-9)
- Kaburu SS, Newton-Fisher NE (2016) Bystanders, parcelling, and an absence of trust in the grooming interactions of wild male chimpanzees. *Sci Rep* 6:20634. <https://doi.org/10.1038/srep20634>
- Katsu N, Yamada K, Okanoya K, Nakamichi M (2019) Temporal adjustment of short calls according to a partner during vocal turn-taking in Japanese macaques. *Curr Zool* 65:99–105. <https://doi.org/10.1093/cz/zoy077>
- Katzir G (1983) Bowing and allopreening of captive jackdaws *Corvus monedula*. *Ibis* 125(4):516–523. <https://doi.org/10.1111/j.1474-919X.1983.tb03145.x>
- Kendon A (1967) Some functions of gaze-direction in social interaction. *Acta Psychol* 26(C):22–63. [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4)
- Knight C, Studdert-Kennedy M, Hurford JR (2000) *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511606441>
- Kolff K (2025) *Communicative underpinnings in grooming interactions of wild-living chimpanzees* (Doctoral dissertation), Universität Osnabrück. <https://doi.org/10.48693/708>
- Kolff K, Pika S (2025) Turn-taking in grooming interactions of chimpanzees (Pan troglodytes schweinfurthii) in the wild: The role of demographic and social factors. *Anim Cogn* 28:1–14. <https://doi.org/10.1007/s10071-025-01940-7>
- Kondo N, Watanabe S, Izawa EI (2010) A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithol Sci* 9:83–91. <https://doi.org/10.2326/osj.9.83>
- Kusnir F, Pesin S, Landau AN (2023) Hello from the other side: Robust contralateral interference in tactile detection. *Atten Percept Psychophys* 86:295. <https://doi.org/10.3758/s13414-023-02801-6>
- Laidre ME (2011) Meaningful gesture in monkeys? Investigating whether mandrills create social culture. *PLoS ONE* 6(2):e14610. <https://doi.org/10.1371/journal.pone.0014610>
- Lanckester ER (1870) On the use of the term homology in modern zoology, and the distinction between homogenetic and homoplastic agreements. *Ann Mag Nat Hist* 6(31):34–43. <https://doi.org/10.1080/00222937008696201>
- Lemasson A, Gandon E, Hausberger M (2010) Attention to elders' voice in non-human primates. *Biol Lett* 6(3):325–328. <https://doi.org/10.1098/rsbl.2009.0875>
- Lemasson A, Glas L, Barbu S, Lacroix A, Guilloux M, Remeuf K, Koda H (2011) Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Sci Rep* 1:1–4. <https://doi.org/10.1038/srep00022>
- Lemasson A, Guilloux M, Barbu S, Lacroix A, Koda H (2013) Age- and sex-dependent contact call usage in Japanese macaques. *Primates* 54(3):283–291. <https://doi.org/10.1007/s10329-013-0347-5>
- Lemasson A, Pereira H, Levréro F (2018) Social basis of vocal interactions in western lowland gorillas (*Gorilla g. gorilla*). *J Comp Psychol* 132(2):141. <https://doi.org/10.1037/com0000105>
- Leroux M, Bosshard AB, Chandia B, Manser A, Zuberbühler K, Townsend SW (2021) Chimpanzees combine pant hoots with food calls into larger structures. *Anim Behav* 179:41–50. <https://doi.org/10.1016/j.anbehav.2021.06.026>
- Levinson SC (2016a) Interactional foundations of language: The interaction engine hypothesis. In *Language in interaction summer school on human language: From genes and brains to behavior*, pp. 189–200. <https://hdl.handle.net/11858/00-001M-0000-002B-A7E9-2>
- Levinson SC (2016b) Turn-taking in human communication - origins and implications for language processing. *Trends Cogn Sci* 20:6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levinson SC, Holler J (2014) The origin of human multi-modal communication. *Philos Trans R Soc B: Biol Sci*. <https://doi.org/10.1098/rstb.2013.0302>
- Levréro F, Touitou S, Frédet J, Nairaud B, Guéry JP, Lemasson A (2019) Social bonding drives vocal exchanges in bonobos. *Sci Rep*. <https://doi.org/10.1038/s41598-018-36024-9>
- Locke JL, Hauser MD (1999) Sex and status effects on primate volubility: Clues to the origin of vocal languages? *Evol Hum Behav* 20(3):151–158. [https://doi.org/10.1016/S1090-5138\(99\)00003-3](https://doi.org/10.1016/S1090-5138(99)00003-3)
- MacLean EL, Matthews LJ, Hare BA, Nunn CL, Anderson RC, Aureli F, Wobber V (2012) How does cognition evolve? Phylogenetic comparative psychology. *Anim Cogn* 15(2):223–238. <https://doi.org/10.1007/s10071-011-0448-8>
- Matsuzawa T (2010) Cognitive development in chimpanzees: a trade-off between memory and abstraction? In *The making of human*

- concepts, p. 227. <https://doi.org/10.1093/acprof:oso/9780199549221.003.11>
- McCleary L, de Leite T, A. (2013) Turn-taking in Brazilian Sign Language. *J Interact Res Commun Disord* 4:123–154. <https://doi.org/10.1558/jirecd.v4i1.123>
- Meunier B, Durier V, Giacalone A, Coye C, Lemasson A (2023) Social factors drive vocal exchanges in red-capped mangabeys (*Cercocebus torquatus*). *Sci Nat* 110:2. <https://doi.org/10.1007/s00114-023-01835-y>. (PMID: 36651987)
- Meunier B, Durier V, Rossard A, Lemasson A (2024) A quick response, but not too much: experienced mangabeys (*Cercocebus torquatus*) lose interest in contact call exchanges that do not respect this “conversational rule.” *PLoS ONE* 19(11):e0310857. <https://doi.org/10.1371/journal.pone.0310857>
- Mielke A, Carvalho S (2022) Chimpanzee play sequences are structured hierarchically as games. *PeerJ* 10:e14294. <https://doi.org/10.7717/peerj.14294>
- Mielke A, Samuni L, Preis A, Gogarten JF, Crockford C, Wittig RM (2017) Bystanders intervene to impede grooming in Western chimpanzees and sooty mangabeys. *R Soc Open Sci*. <https://doi.org/10.1098/rsos.171296>
- Miller CT, Beck K, Meade B, Wang X (2009) Antiphonal call timing in marmosets is behaviorally significant: Interactive playback experiments. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 195(8):783. <https://doi.org/10.1007/s00359-009-0456-1>
- Mitani JC (2009) Male chimpanzees form enduring and equitable social bonds. *Anim Behav* 77(3):633–640. <https://doi.org/10.1016/j.anbehav.2008.11.021>
- Miyagawa S, DeSalle R, Nóbrega VA, Nitschke R, Okumura M, Tattersall I (2025) Linguistic capacity was present in the *Homo sapiens* population 135 thousand years ago. *Front Psychol* 16:1503900. <https://doi.org/10.3389/fpsyg.2025.1503900>
- Molesti S, Meguerditchian A, Bourjade M (2020) Gestural communication in olive baboons (*Papio anubis*): Repertoire and intentionality. *Anim Cogn* 23:19–40. <https://doi.org/10.1007/s10071-019-01312-y>
- Morisaka T, Yoshida Y, Akune Y, Mishima H, Nishimoto S (2013) Exchange of “signature” calls in captive belugas (*Delphinapterus leucas*). *J Ethol* 31(2):141–149. <https://doi.org/10.1007/s10164-013-0358-0>
- Natalia DD, Subekti F, Mirahayuni NK (2019) Turn taking strategies in political debates. *Anaphora J Lang Lit Cult Stud* 2:56–63. <https://doi.org/10.30996/anaphora.v2i2.3365>
- Nóbrega VA, Miyagawa S (2015) The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Front Psychol* 6:271. <https://doi.org/10.3389/FPSYG.2015.00271/BIBTEX>
- Newton-Fisher NE, Lee PC (2011) Grooming reciprocity in wild male chimpanzees. *Anim Behav* 81(2):439–446. <https://doi.org/10.1016/J.ANBEHAV.2010.11.015>
- Nishida T (1988) Development of social grooming between mother and offspring in wild chimpanzees. *Folia Primatol* 50(1–2):109–123. <https://doi.org/10.1159/000156335>
- Nishida T, Kano T, Goodall J, McGrew WC, Nakamura M (1999) Ethogram and ethnography of Mahale chimpanzees. *Anthropol Sci* 107(2):141–188. <https://doi.org/10.1537/ASE.107.141>
- Palo P, Schaeffler S, Scobbie JM (2014, May). Pre-speech tongue movements recorded with ultrasound. In: Proceedings of the 10th International Seminar in Speech Production (ISSP 10).
- Pereltsvaig A (2020) *Languages of the World*. Cambridge University Press. <https://doi.org/10.1017/9781108783071>
- Pika S. (2007). Gestures in subadult gorillas (*Gorilla gorilla*). In *The Gestural Communication of Apes and Monkeys*, pp. 99–130. Taylor and Francis Inc. <https://doi.org/10.4324/9781003064541-5>
- Pika S (2008) Gestures of apes and pre-linguistic human children: Similar or different? *First Lang* 28(2):116–140. <https://doi.org/10.1177/0142723707080966>
- Pika S (2014) Chimpanzee grooming gestures and sounds: What might they tell us about how language evolved? In *The Social Origins of Language*, pp. 128–140. Oxford University Press. <https://doi.org/10.1093/ACPROF:OSO/9780199665327.003.0010>
- Pika S, Liebal K, Call J, Tomasello M (2005) The gestural communication of apes. *Gesture* 5(1/2):41–56. <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. *Am J Primatol* 60(3):95–111. <https://doi.org/10.1002/AJP.10097>
- Pika S, Wilkinson R, Kendrick KH, Vernes SC (2018) Taking turns: Bridging the gap between human and animal communication. *Proc R Soc B*. <https://doi.org/10.1098/RSPB.2018.0598>
- Pillet-Shore D (2012) Greeting: displaying stance through prosodic recipient design. *Res Lang Soc Interact* 45(4):375–398. <https://doi.org/10.1080/08351813.2012.724994>
- Pillet-Shore D (2018) How to begin. *Res Lang Soc Interact* 51(3):213–231. <https://doi.org/10.1080/08351813.2018.1485224>
- Pollick AS, De Waal FBM (2007) Ape gestures and language evolution. *Proc Natl Acad Sci* 104(19):8184–8189. <https://doi.org/10.1073/pnas.0702624104>
- Pomerantz A (1984) Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shaped. In: Atkinson JM (ed), *Structures of social action* pp. 57–101. Cambridge University Press. <https://doi.org/10.1017/CBO9780511665868.008>
- Pougnault L, Lemasson A, Mulot B, Levréro F (2021) Temporal calling patterns of a captive group of chimpanzees (*Pan troglodytes*). *Int J Primatol* 42(6):809–832. <https://doi.org/10.1007/S10764-021-00262-Y/FIGURES/6>
- Pougnault L, Levréro F, Leroux M, Paulet J, Bombani P, Dentressangle F, Lemasson A (2022) Social pressure drives “conversational rules” in great apes. *Biol Rev* 97(2):749–765. <https://doi.org/10.1111/brv.12821>
- Raam RL, Sterner KN, Noviello CM, Stewart CB, Disotell TR (2005) Catarrhine primate divergence dates estimated from complete mitochondrial genomes: Concordance with fossil and nuclear DNA evidence. *J Hum Evol* 48(3):237–257. <https://doi.org/10.1016/J.JHEVOL.2004.11.007>
- Range F, Fischer J (2004) Vocal repertoire of sooty mangabeys (*Cercocebus torquatus atys*) in the Taï National Park. *Ethology* 110(4):301–321. <https://doi.org/10.1111/J.1439-0310.2004.00973.X>
- Range F, Noë R (2002) Familiarity and dominance relations among female sooty mangabeys in the Taï National Park. *Am J Primatol* 56(3):137–153. <https://doi.org/10.1002/AJP.1070>
- Reby D, Cargnelutti B, Hewison AJM (1999) Contexts and possible functions of barking in roe deer. *Anim Behav* 57(5):1121–1128. <https://doi.org/10.1006/anbe.1998.1056>
- Richman B (1976) Some vocal distinctive features used by gelada monkeys. *J Acoust Soc Am* 60(3):718–724. <https://doi.org/10.1121/1.381144>
- Roberts SG, Torreira F, Levinson SC (2015) The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Front Psychol* 6:509. <https://doi.org/10.3389/fpsyg.2015.0509>
- Rodrigues ED, Santos AJ, Veppo F, Pereira J, Hobaiter C (2021) Connecting primate gesture to the evolutionary roots of language: A systematic review. *Am J Primatol* 83(9):e23313. <https://doi.org/10.1002/ajp.23313>
- Rossano F (2013) Gaze in conversation. In: Sidnell J, Stivers T (eds), *The handbook of conversation analysis*, pp. 308–329. Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch15>

- Rossano F, Brown P, Levinson SC (2009) Gaze, questioning, and culture. In *Cognition and Culture*, pp. 187–249. <https://doi.org/10.1017/CBO9780511635670.008>
- Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn taking for conversation. In *Studies in the Organization of Conversational Interaction*, pp. 7–55. Academic Press. <https://doi.org/10.1016/B978-0-12-623550-0.50008-2>
- Schegloff EA (2007) *Sequence organization in interaction: A primer in conversation analysis I*. Vol. 1. Cambridge University Press. <https://doi.org/10.1017/CBO9780511791208>
- Schel AM, Bono A, Aychet J, Pika S, Lemasson A (2022) Intentional gestural communication amongst red-capped mangabeys (*Cercocebus torquatus*). *Anim Cogn* 25(5):1313–1330. <https://doi.org/10.1007/s10071-022-01615-7>
- Schel AM, Townsend SW, Machanda Z, Zuberbühler K, Slocombe KE (2013) Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE* 8(10):e76674. <https://doi.org/10.1371/journal.pone.0076674>
- SC Levinson (2019) Interactional foundations of language: the interaction engine hypothesis. <https://doi.org/10.7551/mitpress/9780262043093.003.0001>
- Sherman PW (1977) Nepotism and the Evolution of Alarm Calls: Alarm calls of Belding's ground squirrels warn relatives, and thus are expressions of nepotism. *Science* 197(4310):1246–1253
- Smith DJ (2020) Turn-taking in office. *Const Polit Econ* 31(2):205–226. <https://doi.org/10.1007/s10602-020-09308-4>
- Snowdon CT, Cleveland J (1984) "Conversations" among pygmy marmosets. *Am J Primatol* 7:15–20. <https://doi.org/10.1002/AJP.1350070104>
- Soltis J, Leong K, Savage A (2005) African elephant vocal communication I: antiphonal calling behaviour among affiliated females. *Anim Behav* 70(3):579–587. <https://doi.org/10.1016/j.anbehav.2004.11.015>
- Sprecher S, Treger S (2015) The benefits of turn-taking reciprocal self-disclosure in get-acquainted interactions. *Pers Relat* 22(3):460–475. <https://doi.org/10.1111/PERE.12090>
- Steiper ME, Young NM (2006) Primate molecular divergence dates. *Mol Phylogenet Evol* 41(2):384–394. <https://doi.org/10.1016/J.MPEV.2006.05.021>
- Stivers T, Enfield NJ, Brown P, Englert C, Hayashi M, Heinemann T et al (2009) Universals and cultural variation in turn-taking in conversation. *Proc Natl Acad Sci* 106(26):10587–10592. <https://doi.org/10.1073/PNAS.0903616106>
- Takahashi DY, Fenley AR, Ghazanfar AA (2016) Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philos Trans R Soc B Biol Sci* 371(1693):20150370. <https://doi.org/10.1098/rstb.2015.0370>
- Takahashi DY, Narayanan DZ, Ghazanfar AA (2013) Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr Biol* 23(21):2162–2168. <https://doi.org/10.1016/j.cub.2013.09.005>
- Terleph T, Saralamba C, Reichard UH (2022) Long-distance vocal signaling in white-handed gibbons (*Hylobates lar*). *Int J Primatol* 43(5):965–986. <https://doi.org/10.1007/s10764-022-00312-z>
- van Boekholt, B. (2024). *Turn-taking in mother-infant interactions across development in chimpanzees (Pan troglodytes schweinfurthii) in the wild* (Doctoral dissertation). Universität Osnabrück. <https://doi.org/10.48693/590>
- van Boekholt B, Bosshard AB, Pika S (2025) Sequence organization of mother–infant interactions in chimpanzees (*Pan troglodytes schweinfurthii*) in the wild. *Proceedings of the Royal Society B: Biological Sciences*, 292(2059) <https://doi.org/10.1098/rspb.2025.2271>
- van Boekholt B, Pika S (2025) Infrastructure of mother-infant interactions across development in chimpanzees (*Pan troglodytes*) in the wild. *Evol Hum Behav* 46(2):106671. <https://doi.org/10.1016/j.evolhumbehav.2025.106671>
- Vatanen A (2018) Responding in early overlap: Recognitional onsets in assertion sequences. *Res Lang Soc Interact* 51(2):107–126. <https://doi.org/10.1080/08351813.2018.1413894>
- Watts DP (2000) Grooming between male chimpanzees at Ngogo, Kibale National Park. I. Partner number and diversity and grooming reciprocity. *Int J Primatol* 21(2):189–210. <https://doi.org/10.1023/A:1005469302911>
- Wells KD (1977) Territoriality and male mating success in the green frog (*Rana clamitans*). *Ecology* 58(4):750–762. <https://doi.org/10.2307/1936211>
- Zuberbühler K (2005) The phylogenetic roots of language: evidence from primate communication and cognition. *Curr Dir Psychol Sci* 14(3):126–130. <https://doi.org/10.1111/j.0963-7214.2005.00357.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.