



Infrastructure of mother-infant interactions across development in chimpanzees (*Pan troglodytes*) in the wild

Bas van Boekholt^{*}, Simone Pika

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

ARTICLE INFO

Keywords:

Evolution of language
Turn-taking
Mother-infant interactions
Participation framework
Use of signals and actions
Temporal relationships

ABSTRACT

The infrastructure underlying human social interaction can be described by several characteristics, such as the exchange of signals and actions, specific temporal relationships, and the use of directed gaze and body direction. These characteristics are remarkably uniform across several different languages, cultures with some of them emerging in mother-infant interactions early in development before the onset of words. It has been suggested that distinct features underlying human social action might have preceded the evolution of language and are shared across the whole primate lineage. However, despite decades of research on nonhuman primate communication, our understanding of general characteristics underlying communicative interactions, and, more specifically, the role they play in the development of communication, remains surprisingly limited. Hence, here we aimed to gain a more comprehensive overview, by studying mother-infant interactions of one of our closest living relatives, the chimpanzee (*Pan troglodytes schweinfurthii*), living in their natural environments. Specifically, we addressed the following two research questions: 1) Which characteristics built the main infrastructure of mother-infant interactions? 2) Which factors influence the infrastructure of mother-infant interactions? To answer these questions, we observed communicative interactions of a total of 17 chimpanzee mother-infant dyads (0–5 years) in the Ngogo community, Kibale National Park, Uganda between February 2021 and February 2023 ($N = 1295$ observation hours). We specifically focused on four different contexts where interactions frequently occurred, food sharing, nursing, grooming and joint-travel, and investigated the role of demographic factors (age and sex of the infant, interactant class) and interactional factors (context, unit type, turn transition). The results showed that mother-infant interactions were characterized by an equally distributed exchange of signals and actions, showed response times ranging from zero to two seconds, and involved the establishment and maintenance of participation frameworks through high frequencies of directed gaze and body direction. There was little to no effect of age and sex of the infant, interactant class, unit type and turn transition on these characteristics. However, context had a strong influence with relative lower frequencies of signals, quicker response times, and lower frequencies of directed gaze and body direction observed in the joint-travel context. By taking a comparative developmental approach, this study highlights commonalities in the infrastructure of mother-infant interactions between humans and chimpanzees, which contribute to uncover how extinct humans might have socially interacted.

1. Introduction

1.1. Infrastructure of human interaction

Humans show considerable similarities in the way they use language in conversations across the world (Dingemans & Enfield, 2024; Dingemans et al., 2015; Stivers et al., 2009; but see, Keller et al., 2018). While initially, studies investigating these “universal” characteristics or underlying infrastructure of conversation were biased towards speech

interactions and WEIRD populations (Bard et al., 2021; Henrich, Heine, & Norenzayan, 2010), new studies expanded these findings to other populations, modalities (Casillas, de Vos, Crasborn, & Levinson, 2015; de Vos, Torreira, & Levinson, 2015), and languages (Dingemans et al., 2015; Kendrick et al., 2020; Stivers et al., 2009). Overall, these studies demonstrated that human conversations are characterized by a specific cooperative infrastructure, involving tight temporal relationships with little overlap and short gaps (Levinson, 2016; Levinson & Torreira, 2015; Sacks, Schegloff, & Jefferson, 1974; Stivers et al., 2009).

^{*} Corresponding author.

E-mail address: basvanboekholt@hotmail.com (B. van Boekholt).

<https://doi.org/10.1016/j.evolhumbehav.2025.106671>

Received 28 February 2024; Received in revised form 31 January 2025; Accepted 20 February 2025

Available online 7 March 2025

1090-5138/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Additionally, in these interactions, participants make use of directed gaze and body orientation to establish and maintain ‘participation frameworks’, with a preference for face-to-face communication (Goodwin, 2007; Kendon, 1990; Mondada, 2009). A participation framework refers to both the embodied actions and linguistic organization interactants use to determine participation and participant roles (Goffman, 1981). It is not only established before the start of an interaction but can dynamically shift and be upheld throughout the interaction (Goodwin, 2007; Goodwin & Goodwin, 2004). While these characteristics were originally described in spoken conversation (Sacks et al., 1974), recent studies showed that they seem to structure general social interactions, including other modalities (de Vos et al., 2015; Holler, 2022), and joint action (e.g., Clark, 2006; Good & Beach, 2005; Mondada, 2007).

In addition, from the moment infants start to interact with the world around them, their interactions display precursors to this infrastructure (e.g., Bateson, 1975; Caskey, Stephens, Tucker, & Vohr, 2011; Gratier et al., 2015; Kaye & Wells, 1980). For example, first vocal exchanges start within the first month of life (Dominguez, Devouche, Apter, & Gratier, 2016), and around the age of three months infants already avoid overlap (Bateson, 1975; Stern, Jaffe, Beebe, & Bennett, 1975), and produce responses within short time windows (Dominguez et al., 2016; Gratier et al., 2015; Van Egeren, Barratt, & Roach, 2001). These temporal relationships change and are adjusted over development, with full-blown adult-like response times appearing at the age of approximately six years (Hilbrink, Gattis, & Levinson, 2015; Nguyen, Versyp, Cox, & Fusaroli, 2022). Additionally, mother-infant interactions are characterized by high frequencies of directed gaze from early age (e.g., Filipi, 2009; Lavelli & Fogel, 2013; Schneider, Roemer, Northrup, & Iverson, 2022), by which participation frameworks are co-constructed by both interactants on a moment-by-moment basis (Goodwin, 2007; Goodwin & Goodwin, 2004; Rossano, 2013a). These mother-infant interactions involve the exchange of signals, defined as any behavioral act that affects the behavior of another individual indirectly (Krebs & Dawkins, 1984; Wheeler et al., 2011). They can also contain actions (e.g., Fantasia, Markova, Fasulo, Costall, & Reddy, 2015; Nomikou, Leonardi, Radkowska, Raczaszek-Leonardi, & Rohlfing, 2017; Reddy, Markova, & Wallot, 2013), defined as socially directed behaviors that result in the perceived goal through direct manipulation of another’s body or the movement of one’s own body (Fröhlich, Wittig, et al., 2016a; Halina, Rossano, & Tomasello, 2013). For instance, the goal of acquiring an object from an interactant can be reached by using the signals of holding out a hand and asking: “*Can you give that to me*”, but also by the action of grabbing the object.

Several studies emphasized that a detailed understanding of mother-infant interactions and involved infrastructure are crucial to tackle the origins and possible precursors of language acquisition, language use and cognitive mechanisms underlying language development (e.g., Abreu & Pika, 2022; Bateson, 1975; Greenfield, 1991). By focusing on these early interactions, it is also possible to study the evolutionary path of different characteristics, since those that appear earlier in ontogeny might also appear earlier in phylogeny (Bard & Leavens, 2014; Gillespie-Lynch, Greenfield, Lyn, & Savage-Rumbaugh, 2014). These ideas have been further nurtured by the “*interaction engine hypothesis*” (Levinson, 2006, 2016), which postulates that the capability of humans to socially interact, and involved skills and mechanisms, paved the way for language to evolve. While the whole “*interaction engine*” might be unique for humans, different components may be of different antiquity with precursors possibly already in existence across the whole primate lineage (Levinson, 2016; Levinson & Holler, 2014). Hence, one crucial window is the comparative approach, which examines similarities and differences between human and other animal species to then draw informed inferences about the abilities of extinct ancestors and create scenarios of language evolution (e.g., Fitch, 2005, 2017; Pika, 2015).

1.2. Comparative approach to interactions

To date, recent studies disentangling the infrastructure of non-human animal interactions investigated joint-action coordination (Bangerter, Genty, Heesen, Rossano, & Zuberbühler, 2022; Genty et al., 2020; Heesen et al., 2021), sequence organization (Mondada & Meguerditchian, 2022; Mondémé, 2023; Rossano, 2013b), repair mechanisms (Fröhlich & van Schaik, 2022; Heesen, Fröhlich, Sievers, Woensdregt, & Dingemanse, 2022), and the underlying organizational structure of social interactions — turn-taking (Fröhlich et al., 2016). Especially, cooperative turn-taking has received increasing attention with some elements of turn-taking being observed in a wide variety of species (Mondémé, 2021; Pika, Wilkinson, Kendrick, & Vernes, 2018). However, the term ‘turn-taking’ has not been consistently defined, varying from simple alternating signals to having recursive complex internal structures and interactive repair across turns (Dingemanse & Enfield, 2024; Pika et al., 2018). Following this varying use of the term ‘turn-taking’ in comparative studies, Dingemanse and Enfield (2024) suggested a conceptual distinction between two types of turn-taking: Turn-taking in the broad sense (TTB) and turn-taking in the narrow sense (TTN). TTB refers to the use of alternating signals avoiding overlap, which may be observed in a wide variety of species, while TTN is defined as involving flexible recipient selection, unfixed, recursively complex internal structure, and interactive repair across turns. To facilitate informed and meaningful cross- and within species comparisons for interactions, Pika and colleagues (2018) recently developed a comparative framework defining turn-taking as: “the orderly exchange of purely communicative signals or behaviours (e.g., peek-a-boo games in humans) between individuals characterized by principles for the co-ordination of turn transfer, which result in observable temporal regularities”. To comparatively investigate turn-taking the study introduced four key elements of human conversational turn-taking, providing insights into both the infrastructure and the content of interactions: (A) flexibility of turn-taking organization, (B) who is taking the next turn, (C) when do response turns occur, and (D) what should the next turn do. Element B, who is taking the next turn, and element C, when do response turns occur, relate to the infrastructure of interactions while element A, flexibility of turn-taking organization, and element D, what should the next turn do, discuss the sequential organization of interactions.

Another potential characteristic describing the infrastructure of interactions is the relative use of signals and actions. The majority of human communicative interactions, such as conversations, consist of signal-signal exchanges with relatively little use of actions (Sacks et al., 1974; Sidnell & Stivers, 2013). However, this might not hold for interactions of nonhuman primates (hereafter primates), which are generally of a more imperative nature with individuals trying to achieve a physical goal and manipulating recipients (Hobaiter & Byrne, 2014; Krebs & Dawkins, 1984; Pika, 2009). However, in their aim to compare primate to human interactions, the majority of comparative studies on primate communication have predominantly focused on signals and emitter-centered data (e.g., Badihi et al., 2024; Fröhlich, 2017; Hobaiter, Byrne, & Zuberbühler, 2017; Roberts, Vick, & Buchanan-Smith, 2012), ignoring the role actions might play (but see: Fröhlich, Wittig, & Pika, 2016b).

One key model species to comparatively study communication is one of our closest living relatives, chimpanzees (*Pan troglodytes*), since any similarities found are likely being due to homology, enabling insight into the capacities of our last common ancestor. To date, studies showed that chimpanzees engage in a variety of social interactions including high levels of cooperation, such as hunting (e.g., Boesch, 1994; Watts & Mitani, 2002), food sharing (e.g., Badescu et al., 2020; Wilkinson, Leudar, & Pika, 2012), and grooming (e.g., Mitani, 2009; Watts, 2000). These interactions are influenced by both demographic and social factors such as age, sex, rank and social bond (e.g., Langergraber, Mitani, & Vigilant, 2009; Newton-Fisher & Kaburu, 2017; Watts, 2000). Moreover, studies investigating communicative abilities of chimpanzees revealed

that they make use of a multifaceted repertoire of signals, involving vocalizations, oro-facial sounds, facial expressions and gestures (e.g., Bard, Maguire-Herring, Tomonaga, & Matsuzawa, 2019; Goodall, 1986; Parr, Waller, Vick, & Bard, 2007). This repertoire is used with high degrees of flexibility (e.g., Bard et al., 2019; Pika & Mitani, 2006; Plooij, 1978; Tomasello, Call, Nagell, Olguin, & Carpenter, 1994), and first-order intentionality (e.g., Call & Tomasello, 2007; Schel, Townsend, Machanda, Zuberbühler, & Slocombe, 2013), including sensitivity to attentional states and audience effects (e.g., Crockford, Wittig, Mundry, & Zuberbühler, 2012; Liebal, Call, Tomasello, & Pika, 2004). However, communicative exchanges have been very little studied although they have the crucial potential to provide us with important insights about the antiquity and involved elements of the infrastructure of human interactions.

1.3. Infrastructure and development of chimpanzee interaction

The first pioneering study of communicative development in chimpanzees was carried out in the 1970s by Plooij (1978, 1979), inspiring subsequent studies on communicative and emotional development (Bard et al., 2014), signal acquisition (Liebal, Schneider, & Errson-Lembeck, 2019; Pika & Fröhlich, 2019), and involved cognitive complexity (Fröhlich, Wittig, & Pika, 2016a; Fröhlich, Wittig, & Pika, 2019). Concerning the onset of signaling, studies investigating chimpanzees in both captive and natural environments, showed that infants start to produce vocal signals in interactions within the first hours of their lives (Soldati et al., 2022) and gestures within their first year of life (Bard et al., 2014; Fröhlich & Pika, 2019). They show an earlier onset age of intentional communication than humans (Fröhlich et al., 2019), with a shift from more action-oriented to more signal-oriented interactions over development (Fröhlich et al., 2019; Fröhlich, Wittig, & Pika, 2016b; Plooij, 1978). However, a systematic measurement of the relative use of signals and actions over development is currently still missing.

Studies on temporal relationships in chimpanzee interactions are rare (but see: Badihi et al., 2024; Fröhlich, Kuchenbuch, et al., 2016) but these have been studied in other primate species, identifying species-specific time windows in which responses during interactions occur (squirrel monkeys, *Saimiri sciureus*: Masataka & Biben, 1987; common Marmoset, *Callithrix jacchus*: Takahashi, Narayanan, & Ghazanfar, 2013; for a review see Pika et al., 2018; Pougnauld et al., 2022), with overlapping and delayed responses resulting in different behavioral reactions (Pougnauld, Levrero, Mulo, & Lemasson, 2020; Takahashi, Fenley, & Ghazanfar, 2016). However, most studies only considered specific signals (e.g., Coo call: Koda, 2004; Phee call: Takahashi et al., 2013), or focused on a single modality (gestures: Badihi et al., 2024). Additionally, relatively little is known about whether these response times change over development.

The establishment and maintenance of participation frameworks in chimpanzee interactions has received more attention. A recent study on two communities of chimpanzees in Uganda and Côte D'Ivoire showed that infants make use of audience checking via eye gaze (Fröhlich et al., 2019), while a second study revealed that infants aged 18 to 53 months were aware of the recipient's visual orientation, with infants using a higher frequency of visual signals when within the visual field of the mother (Dafreville, Hobaiter, Guidetti, Sillam-Dusses, & Bourjade, 2021). Both mother and infant chimpanzees seemed to direct their signals by using directed gaze (Dezecache, Crockford, & Zuberbühler, 2019; Fröhlich et al., 2019; Schneider, Call, & Liebal, 2011), and body direction (Schneider et al., 2011). Moreover, chimpanzees utilized both gaze and body direction to establish a participation framework at the start of the interaction (Fröhlich, Kuchenbuch, et al., 2016). After this initial establishment of a participation framework, chimpanzees used both signals and body direction to maintain this participation framework displaying joint commitment (Bangerter et al., 2022; Heesen, Zuberbühler, et al., 2021; van Boekholt, Wilkinson, & Pika, 2024).

However, only one study investigated the use of gaze over development in zoo-based chimpanzees, showing no change in directed gaze for both mother and infant within the first year of life (Amici, Ersson-Lembeck, Holodynski, & Liebal, 2023).

1.4. Present study

Hence, to fill the existing gaps in the current body of research, the present study aimed to provide a more general understanding of the infrastructure of chimpanzee mother-infant interactions across development. We specifically addressed the two following research questions: First, which characteristics form the basic infrastructure of mother-infant interactions? To address this question, we observed mother-infant interactions in a habituated community of Eastern chimpanzees (*Pan troglodytes schweinfurthii*) living in their natural environment. We specifically focused on the distribution and relative use of signals and actions during interactions, temporal relationships in turn transitions, and the establishment and maintenance of a participation framework via gaze and body direction. If, as proposed by Levinson (2016), the turn-taking organization characterizing human social interactions is ancient and shared with our closest living relatives, the chimpanzees, we expected that mother-infant interactions would be characterized by the use of signals and actions, specific temporal relationships, and high frequencies of directed gaze and body orientation. Second, which main factors are influencing the infrastructure of mother-infant interactions? To address this question, we focused on the following demographic and interactional factors: 1) the age of the infant (0–5 years), 2) the sex of the infant, 3) interactant class (infant or mother), 4) context of the interaction (food sharing, grooming, nursing and joint-travel), 5) unit type produced (action or signal), and, 6) whether the signal or action was part of a turn transition. If development influences mother-infant interactions, we expected to find an influence of the age of the infant on the previously mentioned characteristics (relative use of signals and actions, temporal relationships in turn transitions, and the establishment and maintenance of participation frameworks). However, if these characteristics form the fundamental infrastructure of mother-infant interactions in general, we expected these characteristics to show consistency across interactions with little to no influence of either demographic or interactional factors.

2. Methods

2.1. Study site and subjects

Data on communicative interactions of Eastern chimpanzees (*Pan troglodytes schweinfurthii*) were collected in two groups of the Ngogo chimpanzee community, Kibale National Park, Uganda during two observation periods (February – September 2021, and August 2022–February 2023). Group size varied, but at the end of the observation period the groups consisted of 85 (West) and 110 (Central) individuals (Ngogo Chimpanzee Project, unpublished data). We observed the behavior of a total of 17 mother-infant dyads with infant ages ranging from 4 to 62 months (see Table 1 for detailed information about our focal individuals). Because one mother gave birth to a second infant in between the observation periods, two dyads consist of the same mother. We had access to long-term data on demography and relatedness for both groups.

2.2. Data collection

Full-day focal observations (07:30 AM – 18:30 PM) of mother-infant dyads were conducted using the rule of continuous recording when the visibility allowed (Altmann, 1974; Martin & Bateson, 1994). We maintained a record of the frequency with which a particular dyad had been observed. In situations where we could choose which of several dyads to film, we targeted those individuals previously sampled least often. In

Table 1

Information regarding, group, ID, sex, age, observation hours and video-recorded interaction of the 17 observed mother-infant dyads. The last line provides a summary for each column (P1/P2: first/second period of data collection). ^a Infant not yet born in P1. ^b Infant died during P2.

Group	Mother	Infant	Sex	Age infant P1 (in months)	Age infant P2 (in months)	Observation hours (days)	Video-recorded interactions (in hours)
West	Bacall	Sally	F	16–24	34–40	47 (18)	6.2
West	Beecher	Haldane	M	13–21	31–37	83 (27)	16.1
West	Binoche	Bly	M	N/A ^a	8–14	36 (9)	8.2
West	Carson	Oswald	M	N/A ^a	3–9	69 (19)	9.4
West	Carson	E.O.	M	38–46	56–62	140 (35)	30.8
West	Fleming	MF-DOOM	M	16–24	34–40	76 (21)	10
West	Leigh	Lootus	F	6–14	24–30	88 (25)	17.5
West	Miliah	Malaika	F	25–33	43–49	98 (29)	18.2
West	Ntwiga	Leo	M	9–17	27–33	83 (23)	19.7
West	Penelope	Lecter	M	15–23	33–39	104 (33)	30.7
West	Sabin	Louis	M	36–44	54–60	103 (25)	22.2
Central	Beryl	Lindsay	F	37–45	55–61	53 (17)	10.2
Central	Callas	Kano	M	35–43	53–59	41 (12)	8.0
Central	Christine	Isobel	F	4–12	22–24 ^b	78 (20)	9.6
Central	Emmylou	Colton	F	4–12	22–28	91 (22)	18.3
Central	Fitzgerald	Gatsby	M	28–46	56–62	70 (17)	15.5
Central	Hester	Goblin	M	16–24	34–40	34 (8)	7.5
11:6	17	17	6:11	15	17	1295 (360)	258.1

addition, due to the developmental focus of our study, we tried to follow each mother-infant dyad at least one day during a 30-day period. During focal follows, all interactions between the infant and its mother in the contexts of food sharing, grooming, joint-travel and nursing were recorded using a digital high-definition camera (Sony AX100E 4K) with an external attached unidirectional microphone (Sony shotgun mod ECM-CG60). Using this equipment allowed us to capture high-quality video footage (van Boekholt et al., 2024). The present study was purely observational and non-invasive (for more information see the Ethics statement).

2.3. Definitions

To investigate interactions and involved characteristics, we adopted the concept of “an interaction” from the research field of conversation analysis (Levinson, 2006; Sacks et al., 1974; Sidnell & Stivers, 2013), combined with approaches from comparative studies to primate communication (Pika et al., 2018; Roberts et al., 2012; Rossano, 2013b). Here, an interaction was defined as any exchange between at least two individuals, in which one individual initiated the interaction by producing one signal or action, or a sequence of signals or actions, which could be responded by a recipient with one signal or action, or a sequence of signals or actions. An interaction ended when neither interactants had produced any signal(s) or action(s) for more than thirty seconds (Koski, Sterck, de Vries, & van den Tweel, 2007; Newton-Fisher & Lee, 2011; Roberts et al., 2012; Takahashi et al., 2013). Actions were defined as any directed behavior that is mechanically effective, and results in achieving the perceived goal by direct manipulation of another’s body or the movement of one’s own body (Fröhlich, Wittig, & Pika, 2016a; Halina et al., 2013). For instance, grooming or joint-travel dorsal are examples for actions. Signals included facial expressions, gestures, oro-facial sounds, vocalizations, or combinations of them (See Table S1 for definitions). Any produced action or signal within the interaction was defined as one “unit” where unit type referred to it being either an action or a signal. As this is an exploratory holistic approach to the various characteristics describing the infrastructure of mother-infant interactions, we formulated an objective agnostic mechanistic definition for a turn, where we defined a turn as all consecutive units produced by one individual. A turn would end if either: 1) Another individual produced a unit or a sequence of units, meaning a turn transition had occurred between both interactants (Clayman, 2013; Sacks et al., 1974); 2) Neither interactants produced any unit for thirty seconds which ended the interaction retroactively, and therefore the turn as well. This definition simplifies real-world communication in which an initiating turn might be followed by a pursuing turn from the same individual, and

might not reflect a “turn” from the perspective of the animal. However, as we do not have direct access to the underlying meaning or linguistic organization of animal communication, any definition will contain a degree of simplification and arbitrariness. Defining the end of a turn by the response (or lack thereof) of an interactant allowed us to investigate the potential back-and-forth nature of interactions. Additionally, while this definition might influence metrics such as turn duration and frequency, it is independent of the metrics measured in this study, as they either were only measured between turns produced by two different individuals (response time), or were not measured on the level of singular turns but rather on the whole interaction (signal/action ratio and participation framework). The latency of turn transitions (response time) was measured from the onset of the last unit in one turn till the onset of the first unit in the next turn (Fischer, Cheney, & Seyfarth, 2000). This response time measures how promptly an individual responds in relation to the start of production. While this definition differs from the more commonly measured parameter “offset to onset” (Badihi et al., 2024; Fröhlich, Kuchenbuch, et al., 2016; Pougnauld, Lemasson, Mulot, & Levréro, 2021), it creates an equal baseline measurement over different modalities and unit types of different lengths – i.e. independent of the duration of a unit, the processing time for recognizing its onset would be roughly equal. If this processing time differs for different unit types is something we explore as well in this study. Data was collected in the contexts food sharing, nursing, grooming and joint-travel, with contexts being defined by the presence of certain context-specific behaviors (detailed in the supplementary materials, Table S2). Previous studies have shown that these contexts are successful candidates for studying frequently occurring communicative exchanges between mothers and infant (e.g., Dafreville et al., 2021; Fröhlich, Wittig, & Pika, 2016b; Hirata, 2008, 2009). One example interaction for each context can be found in Fig. 1 and the supplementary materials (Fig. S2–4), including videos of each of these interactions (Video S1–4)

2.4. Video coding and reliability

For each mother-infant dyad, interactions were coded which fulfilled the following two criteria: 1) Both interactants were visible throughout the interaction; 2) the interaction was complete, including both the beginning and the end. Additionally, when data availability allowed, we balanced the number of interactions for each combination of mother-infant dyad and context (grooming/nursing context: 4–7 interactions per dyad; joint-travel context 7–8 interactions per dyad), as well created an even distribution over the available age ranges for each infant. However, due to a limited availability of food sharing interactions, all complete interactions with enough visibility were coded (4–18

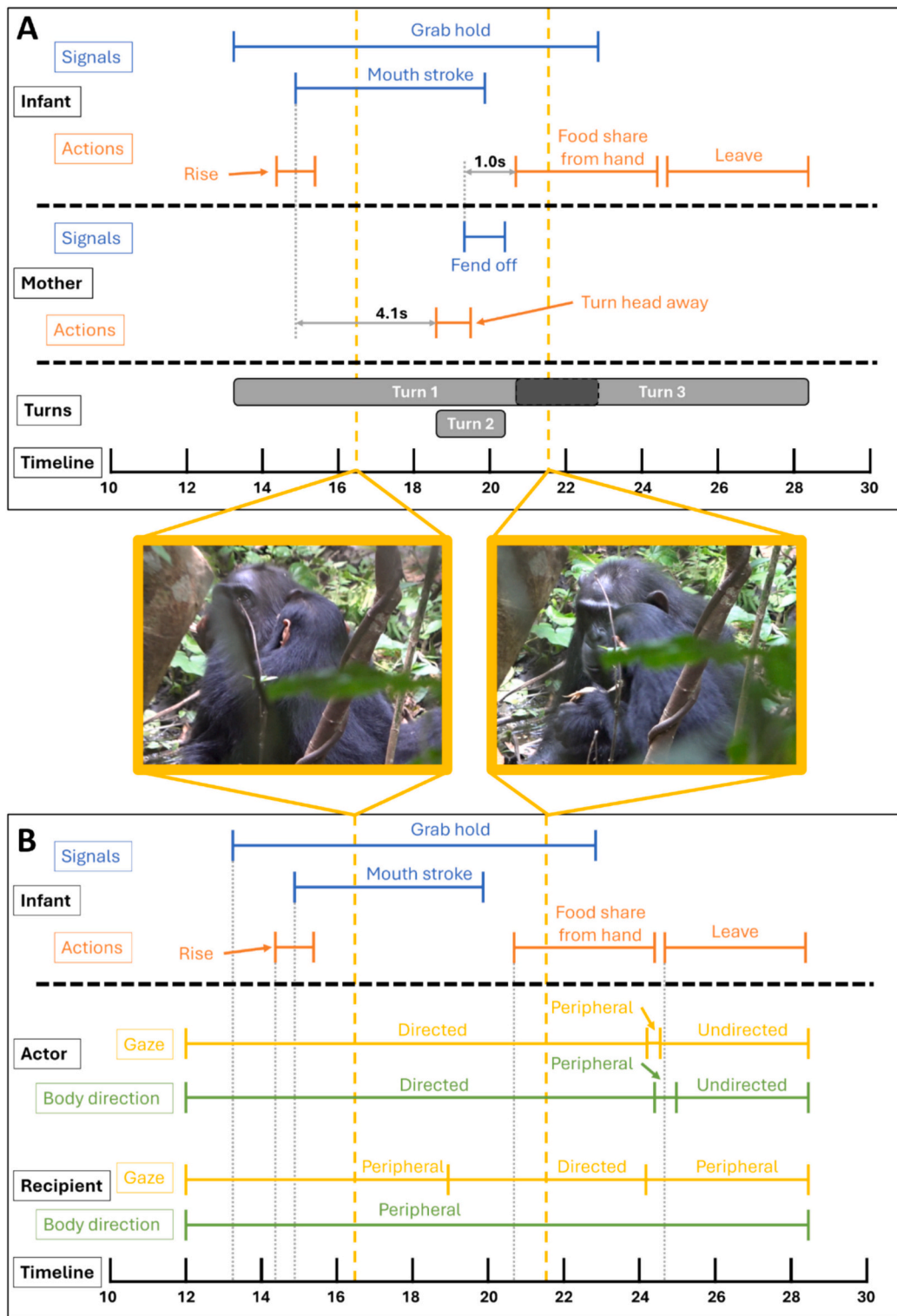


Fig. 1. Example of one food sharing interaction and the extracted characteristics measured in this study. The original video can be found in the supplementary materials (Video S1). *Relative use of signals and actions* were calculated by counting the number of actions and signals produced in one interaction, as well as counting the types of turn transitions (signal-signal, signal-action, action-signal, action-action). **A)** Example of how *temporal relationships* were extracted from the coded interactions. Response times were measured from the onset of the last unit in one turn to the onset of the first unit of the next turn produced by the other individual. **B)** Example of how *participation frameworks* were extracted from the coded interactions. Actor is the individual producing the unit whereas the recipient receives the unit. Gaze and body direction were continuously coded for both interactants. *Gaze and body direction* were established at the onset of each produced unit within an interaction for both the actor and recipient. For the *establishment of the participation framework*, only the gaze and body direction at the start of the interaction were considered.

interactions per dyad for eight dyads). This led to a total of 361 coded videos, totaling ~19 hrs of footage. We used the ELAN software (version 6.2; Lausberg & Sloetjes, 2009) to code all interactions. Behavioral definitions were based on established ethograms of chimpanzee behavior (Byrne et al., 2017; Fedurek, Slocombe, Hartel, & Zuberbuhler, 2015; Fernandez-Carriba, Loeches, Morcillo, & Hopkins, 2002; Fröhlich, Wittig, & Pika, 2016b; Genty & Fuchs, 2023; Goodall, 1989; Hobaiter & Byrne, 2011; Kersken, Gomez, Liszkowski, Soldati, & Hobaiter, 2019; Nishida, Kano, Goodall, McGrew, & Nakamura, 1999; Parr et al., 2007; Rodrigues, Santos, Hayashi, Matsuzawa, & Hobaiter, 2021; Watts, 2016; Wilke et al., 2022; Wilke et al., 2022; for the complete ethogram see Table S3), and were supplemented with new behaviors when literature was insufficient. To study the establishment and the maintenance of the participation framework, we coded the gaze and the body direction of both interactants continuously. Gaze and body direction were categorized based on the relative position of the other interactant in relation to the midline of the individual measured in degrees angle. The midline for the gaze and body direction aligned with the direction of the nose or nipples of the individual respectively. Gaze and body direction was classified in three categories: directed (≤ 45 degrees from the midline), peripheral (between 45 and 90 degrees from the midline) and undirected (>90 degrees from the midline) (Fig. S5). See Fig. 1 for an example of a coded food sharing interaction showcasing how the three measured characteristics were extracted.

Eighteen percent of the selected videos (>15 % per context) were coded by a second observer and tested using Cohen's Kappa coefficient in the "EasyDIAG" package (Holle & Rein, 2015). This software allowed to test agreement both on annotation value and temporal nature. To reach agreement, the annotation value had to be identical and the duration had to overlap by at least 60 %. Reliability agreement was thus based on a combined measurement (Heesen et al., 2021). The test revealed good agreement for actions ($n = 577$, Cohen's $\kappa = 0.71$), gaze ($n = 1046$, Cohen's $\kappa = 0.63$) and body direction ($n = 1046$, Cohen's $\kappa = 0.67$) and moderate agreement for signals ($n = 416$, Cohen's $\kappa = 0.56$) (Cicchetti, 2001; Cicchetti, Volkmar, Klin, & Showalter, 1995).

2.5. Data analyses

To tackle our first research question, concerning which characteristics construct the infrastructure of mother-infant interactions, we used descriptive statistics to describe first the ratio of signals and actions, both on an interactional level as well as surrounding turn transitions (Fig. 1). Second, after excluding outliers, we analyzed the distribution of response times measured from onset of the last unit in one turn to the onset of the first unit in the next turn displayed on a density curve (Badihi et al., 2024; Stivers et al., 2009). Additionally, we calculated a response threshold by performing the "abrupt point change test for exponentially distributed bout length" on the distribution of response times plotted on a log-survival plot (Haccou & Meelis, 1992; Koda, 2004; Sugiura, 2001; See supplementary materials for more information). Third, we measured the relative use of directed, peripheral and undirected gaze and body direction for both the actor and the recipient. We analyzed the gaze and body direction for both interactants at the start of each produced unit within an interaction (Fig. 1B). For the establishment of a participation framework, we separately considered the gaze and body direction at the moment the first unit in an interaction got produced.

To investigate the second research question, concerning which factors influenced the infrastructure of mother-infant interactions, we fitted different Generalized Linear Mixed Models (GLMM) for each main characteristic. The first model (S/A-model) focused on the use of signals and actions and used a dataset containing all units produced within the coded interactions ($n = 3655$) and was fitted with a binomial error structure with a logit link function (McCullagh & Nelder, 1989). The second model (RT-model) focused on the temporal relationships of turn transitions and used a dataset containing all turn transitions within the

coded interactions ($n = 1154$), and was fitted with a gaussian error structure with the response times log-transformed (Baayen, 2008). The third set of models (PF-models) focused on the maintenance of the participation framework, and used a dataset containing the combination of both the gaze and body direction of both the actor and recipient at the onset of each produced unit (Gaze actor $n = 3116$; Body direction actor $n = 3182$; Gaze recipient $n = 3024$; Body direction recipient $n = 3187$) and were fitted with an ordinal (i.e., cumulative logit link) error structure (Agresti, 2000). For the demographic factors, each model contained the predictor variables of infant age, infant sex and interactant class (mother or infant). For the interactional factors each model contained the predictor variable of context. Additionally, for the PF-models the variable of unit type (signal or action) was added whereas for the RT-model, we added turn-transition type (the unit types of both the last unit of the previous turn and the first unit of the next turn, e.g., action – action). Lastly, for the S/A-model and the PF-models we added the variable of turn transition (whether the produced unit was followed by a turn transition). As control variables and to avoid pseudoreplication, we added the following fixed and random effect control predictors to all models: group (fixed), dyad ID (random), interaction ID (random) and date of interaction (random). The levels of the random effect structure contained 17 dyads over 361 interactions collected on 139 different days. Apart from random intercepts, relevant random slopes were added to the random effect structure where identifiable (Barr, Levy, Scheepers, & Tily, 2013; Schielzeth & Forstmeier, 2009). The models were implemented in R (version 4.1.3; R Core Team, 2015) using the function *glmer* of the package "lme4" (S/A and RT-model; Bates, Maechler, Bolker, & Walker, 2015) and the function *clmm* of the package "ordinal" (PF-models; Christensen, 2015). To avoid 'cryptic' multiple testing we conducted one full-null model comparison (Forstmeier & Schielzeth, 2011) using a likelihood ratio test (Dobson, 2002) with the null model only containing control predictors and the random effects structure. If this test was significant, we subsequently compared the full model to reduced models lacking one test predictor using likelihood ratio tests to determine the effect of the individual test predictors (R *drop1* function with argument 'test' set to 'Chisq'). 95 % confidence intervals and fitted values were determined by means of parametric bootstrapping ($N = 100$ bootstraps). More statistical information can be found in the supplementary materials. The data associated with this research are available at https://osf.io/gf3m3/?view_only=4a9457b74557421f8a0702847815b0c3.

3. Results

During a total of 1295 hrs of observation (mean \pm sd = 76.1 ± 28.1 h/mother-infant dyad), we recorded a total of 258 hrs of video footage on mother-infant interactions (15.2 ± 7.6 h per dyad). This method resulted in 2282 recordings (food-sharing: $n = 93$; 5.5 ± 6.1 per dyad; grooming: $n = 585$; 34.5 ± 23.9 per dyad; nursing: $n = 345$; 20.3 ± 11.0 per dyad; joint-travel: $n = 1259$; 74.1 ± 42.9 per dyad). The coding of this data set resulted in a total of 361 coded interactions (food-sharing: $n = 61$; 3.6 ± 5.0 per dyad, grooming: $n = 85$; 5.0 ± 2.4 per dyad, nursing: $n = 80$; 4.7 ± 2.4 per dyad, joint-travel: $N = 135$; 7.9 ± 0.2 per dyad).

3.1. Characteristics of mother-infant interactions

3.1.1. Relative use of signals and actions

Over a total of 361 interactions mother and infant produced 1704 actions (46.7 %; mean = 51.6 per individual) and 1943 signals (53.3 %; mean = 58.9 per individual). The infants produced more units than mothers (infants $n = 2410$, mean = 141.8; mothers $n = 1237$, mean = 77.3), but the ratio between actions and signals were comparable for both (action / signal infants = 48.6 % / 51.4 %; action / signal mother = 43.0 % / 57.0 %). Three interactions contained only signals (0.83 %), 46 interactions contained only actions (12.74 %), and 312 interactions contained a combination of actions and signals (86.43 %) containing

zero to twenty-one turn transitions. Focusing on the different types of turn transitions within interactions, there was a similar distribution of the different types (action-action turn transition = 337 (29.2 %); action-signal turn transition = 292 (25.4 %); signal-action turn transition = 225 (19.5 %); signal-signal turn transition = 300 (25.9 %)).

3.1.2. Temporal relationships

In total, the analyzed dataset contained 1.154 turn transitions over 361 interaction (3.20 ± 3.44 turn transitions per interaction). Response times, which were measured from the onset of the last unit in one turn till the onset of the first unit in the next turn, ranged from 0.01 till 188.20 s with a median of 1.98 s, a unimodal peak at 0.94 s and a mean of 4.28 s (Fig. 2). The abrupt point change test revealed a response threshold of three seconds (Fig. S1). More than half of the response times fell within this response threshold (695/1154 = 60.23 %).

3.1.3. Establishment and maintenance of participation framework

For the majority of the interactions, the gaze and body direction of the actor at the start of the interaction was respectively directed (gaze = 67.6 %; body direction = 73.3 %), followed by peripheral (gaze = 23.6 %; body direction = 18.5 %) and undirected (gaze = 8.8 %; body direction = 8.2 %). In comparison, the gaze and body direction of the recipient at the start of the interaction was less directed (gaze = 43.5 %; body direction 56.5 %), more peripheral (gaze = 41.3 %; body direction = 28.0 %) and more undirected (gaze = 15.2 %; body direction = 16.0 %). Regarding the maintenance of participation frameworks, the gaze of the actor and body direction of both the actor and recipient were mainly directed throughout interactions (gaze actor = 78.9 %; body direction actor = 84.9 %; body direction recipient = 71.0 %), followed by peripheral (gaze actor = 15.2 %; body direction actor = 9.8 %; body direction recipient = 17.4 %) and undirected (gaze actor = 5.8 %; body direction actor = 5.3 %; body direction recipient = 11.6 %). Recipients directed their gaze to the actor in roughly half of the interactions (49.58 %), followed by peripheral (34.4 %) and undirected (15.9 %).

3.2. Influence of demographic and interactional factors

3.2.1. Relative use of signals and actions

The full model (S/A-model) was highly significant compared with the null model, revealing a significant effect of at least one of the predictor variables (likelihood ratio test: $\chi^2 = 35.691$, p -value <0.001). The only predictor variable which showed a significant effect was context, where signals were more likely to occur in the food sharing context

compared to the other contexts (Table S4, Fig. 3). Actions were more likely to occur in the joint-travel context compared to the other contexts. The grooming and nursing contexts did not significantly differ from each other with regards to their relative signal and action use (Table S4, Fig. 3).

3.2.2. Temporal relationships

The full model (RT-model) was highly significant compared with the null model revealing a significant effect of at least one of the predictor variables (likelihood ratio test: $\chi^2 = 39.062$, p -value <0.001). The only predictor variable which showed a significant effect was context, where response times were quicker in the joint-travel context compared to the other three contexts. These did not significantly differ from each other (Table S5, Fig. 4).

3.2.3. Maintenance of participation framework

In all four models, the full model was highly significant compared to the null model (likelihood ratio test: Gaze actor: $\chi^2 = 75.176$, p -value <0.001; Body direction actor: $\chi^2 = 75.983$, p -value <0.001; Gaze recipient: $\chi^2 = 94.732$, p -value <0.001; Body direction recipient: $\chi^2 = 85.090$, p -value <0.001). All predictor variables, except sex, showed a significant effect in at least one model. The age of the infant had an effect on gaze of the actor and the body direction of both the actor and recipient. Here, interactions with older infants had lower frequencies of directed gaze of the actor and directed body direction of both interactants (Table S6, Fig. 5). Interactant class was significant in all models, with the mother showing lower frequencies of directed gaze and body direction compared to the infant. This was regardless of whether the mother was the actor or the recipient (Table S6, Fig. S6). Context was significant in all models, with lower frequencies of directed gaze and body direction of both the actor and recipient in the joint-travel context (Table S6, Fig. S7). Moreover, there were higher frequencies of directed body direction of both the actor and recipient in the nursing context compared to the food sharing and joint-travel context (Table S6, Fig. S7). Unit type was significant in the gaze and body direction of the actor, with actions being accompanied by higher frequencies of directed gaze and body direction compared to signals (Table S6, Fig. S8). Lastly, turn transition was significant in the gaze and body direction of the recipient Here, units that were part of a turn transition (meaning it was the last unit in a turn and followed by a response from the other individual) were accompanied by higher frequencies of directed gaze and body direction of the recipient compared to units not being part of a turn

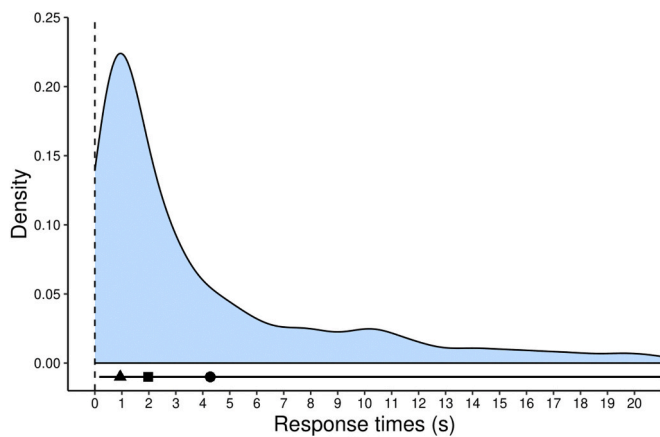


Fig. 2. Density plot of the response times measured from the onset of the last unit in one turn till the onset of the first unit in the next turn. Data include the 95 % interquartile ranges of response times ($N = 1038$). The highest modality ($\blacktriangle = 0.94$ s), median ($\blacksquare = 1.98$ s) and mean ($\bullet = 4.28$ s) are displayed below the graph. To increase visibility, the graph is cut off after 20 s (range = 0.01–188.20 s).

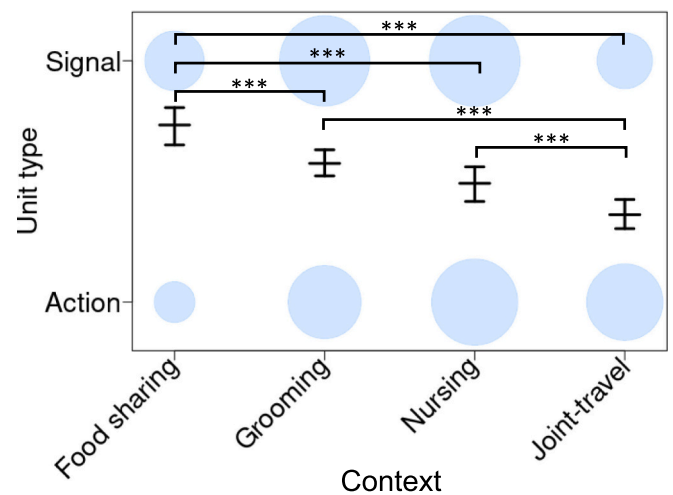


Fig. 3. Effect of context on the use of signals and actions ($*** \leq 0.001$). The area of the dots corresponds to the sample size. The lines represent the fitted model with error bars based on 95 % confidence intervals with all other covariates and factors centred to a mean of zero.

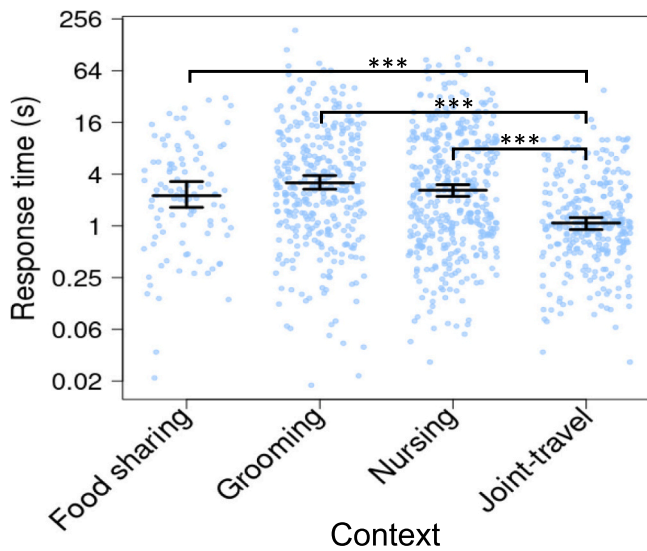


Fig. 4. Effect of context on the duration of response times. Response times on the y-axis are placed on a logarithmic scale. Dots represent individual data points. The lines represent the fitted model with error bars based on the confidence intervals with all other covariates and factor centred to a mean of zero.

transition. This means, that the unit was followed by at least one unit produced by the same individual (within-turn) (Table S6, Fig. S9).

4. Discussion

The aim of this study was to gain a better understanding of the infrastructure underlying mother-infant interactions in chimpanzees living in their natural environments across development. We investigated the following two research questions: First, which characteristics construct the basic infrastructure of mother-infant interactions in chimpanzees? Second, which main factors influence the infrastructure of mother-infant interactions?

Overall, the results showed that mother-infant interactions were characterized by an equal production of signals and actions, response times distributed around zero to two seconds, and high frequencies of directed gaze and body direction, both in general and at the initiation of an interaction. There was no effect of the demographic factors age and sex of the infant on the relative frequency of signals and actions and response times, meaning these characteristics did not show any change over development. However, the interactional factor context had an effect, with interactions in the food sharing context involving a relative higher frequency of signals compared to all other investigated contexts. In addition, interactions in the joint-travel context showed a relative higher frequency of actions compared to the other contexts. Response times in the joint-travel context were significantly quicker compared to

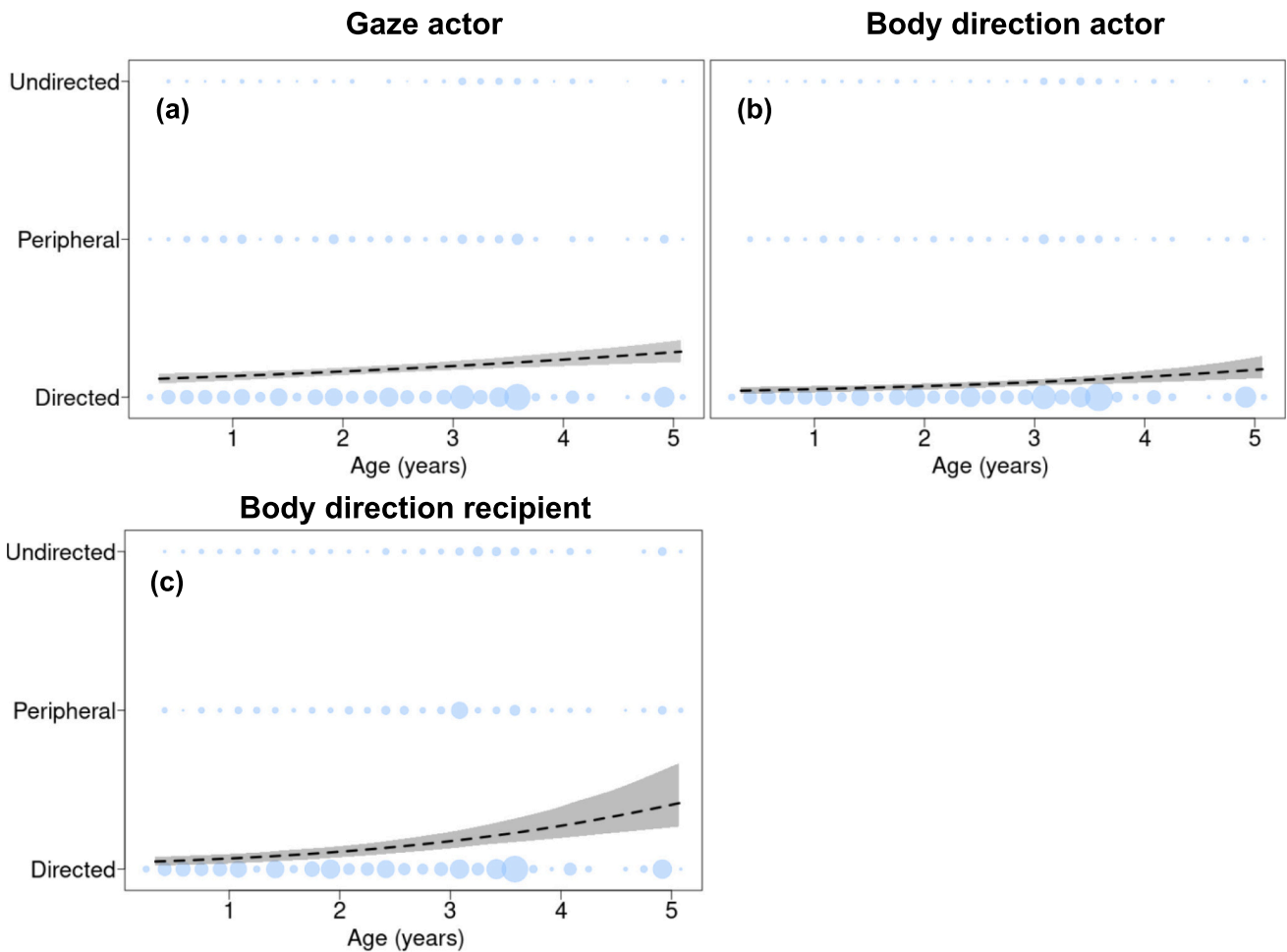


Fig. 5. Effect of age on the levels of (a) gaze actor, (b) body direction actor, and (c) body direction recipient. Depicted are data points for each age binned per two months where the area of the circles corresponds to the sample size. The dashed line represents the fitted model with the grey area based on the confidence intervals with all other covariates and factors centred to a mean of zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the other three investigated contexts. Moreover, our results showed small effects of age, context, interactant class, unit type and the presence of a turn transition on the frequencies of directed gaze and body direction used to maintain participation frameworks. Signals produced in interactions in the joint-travel context by mothers with older infants showed the lowest frequencies of directed gaze and body direction of the actor. The lowest frequencies of directed gaze and body direction of the recipient were found in interactions in the joint-travel context with older infants, when there was no turn transition, and the mother was the recipient. Additionally, higher frequencies of directed body direction of both the actor and the recipient were found in interactions in the nursing context compared to the other contexts. In the next sections, we will discuss these results in more detail for each characteristic separately.

4.1. Relative use of signals and actions

We found that mothers and infants of the Ngogo population rarely engaged in social interactions that consisted of purely signal or purely action exchanges, with more than 86 % of interactions containing both signals and actions. Our results suggest that both actions and signals play an important role in communicative exchanges in chimpanzees. However, with previous studies investigating interactions in primates either focusing on only purely signal-signal exchanges (Badihi et al., 2024; Pougnault et al., 2021) or not distinguishing between signals and actions (Fröhlich, Kuchenbuch, et al., 2016), it is difficult to carry out quantitative direct comparisons. Our results thus show that in contrast to adult human communicative interactions (Sacks et al., 1974), the basic infrastructure of mother-infant interactions of chimpanzees seems to be differently designed, involving predominantly an action-signal turn-taking system which is complemented by signal-signal or action-action systems. These interactions might be more akin to interactions between mothers and prelinguistic infants which can exhibit a mixture of both signals and actions (e.g., Fantasia et al., 2015; Nomikou et al., 2017; Reddy et al., 2013).

Concerning turn transitions, our results showed that signal-signal turn transitions were as frequent as all other possible turn transitions. This suggests that while signal-only interactions were rare, within an interaction bouts of signal-signal exchanges occurred as frequently as any other types of exchanges (Badihi et al., 2024). Future studies could investigate when exactly these signal-signal exchanges occur, representing a potential precursor to human conversations.

The relative frequency of signals and actions produced by chimpanzee infants and mothers within the analyzed interactions was consistent over the studied developmental time-window with no effect of infants' age. This finding contrasts with other studies on communicative development of chimpanzees reporting a shift from action-oriented to signal-oriented interactions over development (Fröhlich & Pika, 2019; Fröhlich, Wittig, & Pika, 2016b; Plooi, 1978). For instance, Plooi (1978) described a developmental sequence in chimpanzees at Gombe, Tanzania, with a chimpanzee infant initially using actions to "defend" himself from his mother tickling him, but later using similar movements as signals to initiate tickling. Another study of Fröhlich et al. (2019; 2016b) at Tai, Ivory coast, and Kanyawara, Uganda, demonstrated a decrease of action use and increase of visual gesture use in joint-travel initiations over development. There may be several explanations for these contrasting results. First, the deployment of signals and actions during interactions may mirror community differences in communication (Pika & Deschner, 2019). For instance, Fröhlich, Wittig, and Pika (2016b) showed a community difference in the use of tactile gestures. However, since all observations were made within the same community, we cannot verify this idea.

Second, the difference could be due to our study investigating interactions in several contexts, while Fröhlich, Wittig, and Pika (2016b) focused solely on the joint-travel context. We found that the studied mother-infant dyads produced relative more actions in the joint-travel context compared to the other three investigated contexts. This

increase in action use in joint-travel could be explained due to the relative higher "urgency" of this context compared to the other contexts (Fröhlich, Wittig, & Pika, 2016b; Tomasello & Zuberbühler, 2002). In chimpanzee communities it is not uncommon for infants to be killed or wounded by within-group fighting or displays (Goodall, 1986; Lowe, Hobaite, Asiimwe, Zuberbühler, & Newton-Fisher, 2020; Wilson et al., 2014). If a within-group fight or threatening situation arises, a mother would want to quickly "collect" her infant and jointly leave. This type of interaction would require a quick resolve and coordination of both participants. Because, by definition, actions immediately achieve their goals themselves (Halina et al., 2013), in "urgent" contexts, individuals may decide to interact using actions – e.g., putting the infant in a ventral position (Fröhlich, Wittig, & Pika, 2016b) – rather than producing more time-costly signals that require the recipient to respond – e.g., perform an extend hand gesture to urge the infant to climb ventrally. Additional evidence for this external pressure for a quicker resolve was shown by the faster response times observed in the joint-travel context. The other three analyzed contexts contained in general relative relaxed and peaceful activities (Goodall, 1986). One explanation for the higher frequency of signals in the food sharing context could be due to conflicting interests of both interactants. Grooming and nursing interactions seem to be beneficial for one individual, while not having considerably costs for the other individual (Goodall, 1986; Watts, 2000). Food sharing interactions, in contrast, entail a cost for the mother as the infant tries to acquire food from the mother, thereby often directly decreasing the mother's nutritional intake (Jaeggi & Van Schaik, 2011). This conflict of interest increases the uncertainty of the outcome of the interaction, which can lead to more complex forms of communication (Grampp et al., 2023; Peckre, Kappeler, & Fichtel, 2019). A similar effect can be seen in consortships, in which adult male chimpanzees produce more gestural signals compared to other contexts (Hobaite, 2011; Hobaite & Byrne, 2012).

4.2. Temporal relationships

The distribution of response times between turns revealed that most responses ranged between zero and two seconds. Additionally, using an abrupt change point analysis (Koda, 2004; Sugiura, 2001), a response threshold of three seconds was found. Given that we measured response times from onset to onset, the fastest response times (e.g., <300 ms) in our dataset are unlikely to reflect actual responses, as these would surpass a minimum required processing time. This significant part of our dataset (~10 %) could be explained by either the interactant responding to a unit prior to the last unit, or, similar to human interaction, by the interactant predicting the upcoming unit and preparing a response accordingly. Future studies could distinguish between these two explanations by investigating if more "predictable" turn transitions (having a higher occurrence in the dataset) lead to quicker response times. The results showed that, apart from context, none of the chosen predictors influenced response times. Therefore, it seemed that even at relatively early infant ages, both mother and infant chimpanzees showed similar response times. Additionally, these results suggest that the temporal characteristic of interactions does not change for the first five years of life. This finding, which is in line with results of a study by Takahashi and colleagues (2016) showing no change in parent-infant inter-call intervals in common marmosets during their first two months of development. However, this finding is in contrast to studies on mother-infant vocal exchanges in humans (Gratier et al., 2015; Hilbrink et al., 2015), which reported response times of 600 milliseconds at the age of three months that initially increased but then decreased and only reached adult-like response times until the age of six years. However, a recent meta-analysis based on 26 different studies on turn-taking interactions, showed that evidence for this change in response times was mixed and might be inconclusive (Nguyen et al., 2022). Since in the present study we used a different measurement for response times, any direct comparisons are not possible. There was no effect of unit

transition type, meaning that the response times were independent of whether both interactants responded to actions or signals with an action or signal themselves. This finding suggests that similar to human conversations (Levinson & Torreira, 2015; Stivers et al., 2009), the timing of turn-taking seems to be independent of modality. This would allow these results to be compared with other studies specifically focusing on signal-signal exchanges (e.g., Badihi et al., 2024; Pognault et al., 2021; Takahashi et al., 2013).

4.3. Establishment and maintenance of participation framework

Our results showed relatively high frequencies of both directed gaze and body direction for the actor, as well as the recipient, both in general and specifically at the start of an interaction. These findings suggest that both interactants mutually oriented themselves towards each other, establishing a participation framework, that was then upheld throughout the interaction. Due to the physical nature of body directions, it had less freedom to vary through some parts of the interaction, as some units required a directed body direction (e.g., joint travel ventral, climb dorsal, full embrace). However, in all models, both gaze and body direction showed the same direction, highlighting the importance of both gaze and body direction to interactions. The only exception was the higher frequencies of directed body direction, but not gaze, in the nursing context compared to the other contexts. However, this could be explained by the physical arrangement of nursing, requiring a mutually directed body direction of both mother and infant.

While our results showed that many factors influenced the frequencies of gaze and body direction, most effects did not fall in the realm of biological relevance as frequencies of directed gaze and body direction remained high throughout. One example of this was the effect of age. When the age of infants increased, the gaze and body direction of both interactants shifted away from directed. However, this shift over the total age range of the infant (4–62 months) was on average from 94.5 % directed to 74.5 % directed. This small but significant shift away from directed gaze was also found in a study on captive chimpanzees, as well as mothers and infants from WEIRD cultures (Amici et al., 2023). As unit type and turn transition led to even smaller changes, we will not discuss what might have caused those shifts. However, one significant difference was found between mothers and their infants. Independent of her role as actor or recipient, mothers showed lower frequencies of directed gaze and body direction compared to the infant. One explanation for this finding could be due to the mothers expanded social environment. Chimpanzee infants are dependent on their mother throughout their childhood (Brundl et al., 2020; Whiten & van de Waal, 2018), and the infant's focus, including its gaze and body direction, will be more focused towards its mother rather than its social environment (Amici et al., 2023). This explanation is also supported by the higher frequencies of directed gaze and body direction observed in younger infants, as they are even more dependent on their mothers. In contrast, mothers actively monitored their surroundings to protect their infant and, therefore, showed lower frequencies of directed gaze and body direction. Future studies could test this prediction by investigating the influence of party composition on gaze and body direction of mothers and infants. Specifically, it would be expected that the presence of adult males, who might be perceived as a potential threat to young dependent infants (Goodall, 1986; Lowe et al., 2020; Wilson et al., 2014), would lead to a decrease of directed gaze and body direction for the mother.

4.4. Limitations

The aim of this study was to pinpoint the characteristics that form the main infrastructure of chimpanzee mother-infant interactions. To do this, we analyzed interactions of 17 dyads in a community living in its natural environment across four different contexts. While we selected contexts known to be crucial candidates for communicative exchanges, we excluded the context of play, in which complex communication can

occur (e.g., Fröhlich, Wittig, & Pika, 2016a; Tomasello et al., 1994). The exclusion of the playing context was chosen due to its dynamic nature making it difficult to objectively classify the mechanical effectiveness of a unit (van Boekholt et al., 2024), which inhibited us to distinguish between actions and signals.

Even though we discussed characteristics of general interactions, all data originated from a single community of one subspecies of chimpanzee. Studies have shown that communication can differ over populations (e.g., Bard et al., 2021; Wilke, Lahiff, Badihi, et al., 2022; for a review see: Pika & Deschner, 2019). This limitation can be found across the field of evolution and human behavior (Bard et al., 2021; Henrich et al., 2010; Nielsen, Haun, Kartner, & Legare, 2017). Therefore, generalizations beyond this community should be made with caution (Simons, Shoda, & Lindsay, 2017). However, within this community of chimpanzees, we controlled for potential effect of dyad and group. Future studies collecting data from multiple communities and subspecies could confirm or refute if these characteristics are comparable across chimpanzees populations and subspecies, or whether they represent group specific patterns.

4.5. Conclusion

In this study, we explored whether turn-taking is an ancient mechanism already present in nonhuman primates. We found that chimpanzee mother-infant interactions were characterized by an equal distribution of actions and signals, response times averaging around one second, and the establishment and maintenance of participation frameworks through high frequencies of directed gaze and body direction. These characteristics seem to be independent of distinct demographic and interactional factors, with the exception of the influence of context. While some theories suggest that language arose in cooperative, possibly prosocial, contexts that require complex coordination (e.g., Pika & Bugnyar, 2011; Tomasello, 2008), this study further emphasized that also other evolutionary pressures, such as those concerning infant survival, influence communicative exchange and thus possibly communicative interactional structures (Fröhlich, Wittig, & Pika, 2016b). At the same time, greater communicative complexity – i.e. higher signal production – may be the result of contexts characterized by conflicts of interest and uncertain outcomes, such as food sharing (Grampp et al., 2023; Peckre et al., 2019), rather than cooperation. This suggests that in addition to the need to cooperate, the need to “convince” the other party and negotiate, might also be a crucial factor driving the evolution of communication (Krebs & Dawkins, 1984; Pika & Fröhlich, 2019).

By comparing communicative interactions of chimpanzees to human social actions during conversations, we conclude that chimpanzees of our study population did not engage in signal-signal ‘conversations’. However, within interactions, signal-signal turn transitions appeared as frequent as all other possible turn transitions (e.g., action-signal). This could mean that the precursor of human conversational turn-taking was a complex system composed of an interplay of signals and actions, similarly to interactions of adult caretakers with prelinguistic human infants. Subsequently, humans might have specialized in purely signal-signal turn-taking, that appear later in development. However, as some of the chosen operationalizations of metrics differed from commonly used definitions in human literature, as well as human turn-taking encompassing more aspects than the interactional infrastructure measured in this study (e.g., sequence organization, flexibility of turn size and order; Pika et al., 2018; Sacks et al., 1974) direct comparisons between human turn-taking and our results are limited. Nonetheless, considering turn-taking in the broad sense (TTB), we find similarities between mother-infant interactions in humans and chimpanzees in the form of high frequencies of directed gaze and body direction. Additionally, these interactional characteristics appear early in development, barely changing with aging infants, highlighting the fundamental nature of this infrastructure of mother-infant interactions. Future studies

should try to uncover if this infrastructure of interaction is present and comparable in other populations, species and age-classes across the primate lineage (Pika et al., 2018). Taking this comparative holistic approach will help to unravel the way ancient humans socially interacted, and if this formed the blueprint from which language evolved.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.evolhumbehav.2025.106671>.

Ethics statement

The present study was purely observational and non-invasive. All applicable national, and/or institutional guidelines for the care and use of animals were followed. The study was classified as non-animal experiment and did not require any approval from a relevant ethics committee in accordance with the German Animal Welfare Act of 25th May 1998, Section V, Article 7. All observers adhered to a strict hygiene protocol, which included a seven-day quarantine as well as the wearing of face masks when encountering chimpanzees. All observations were made from a distance of at minimum seven meters, in order to avoid disease transmission from humans to chimpanzees (Köndgen et al., 2008; Leendertz et al., 2004) as well as to not disturb the chimpanzees in their natural behavior. The study followed the legal requirements of the state of Uganda and was approved by the by the Ugandan Wildlife Authority, and the Ugandan National Council for Science and Technology. Moreover, the study followed the recommendations of the ‘Animals (Scientific Procedures) Act 1986’, as stated by the government of the United Kingdom, and the principles of “Ethical Treatment of Non-Human Primates”, as published by the American Society of Primatologists.

Funding

BvB was supported by a Leakey Foundation Research grant. This work was supported by an EU-Consolidator grant (772000, TurnTaking) of the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme awarded to SP.

CRedit authorship contribution statement

Bas van Boekholt: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Simone Pika:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no potential competing interests.

Acknowledgments

We are very grateful to K.E. Langergraber, J.C. Mitani, and D.P. Watts for allowing us to collect data and observe the behavior of the wonderful chimpanzees of the Ngogo community. We thank the Ugandan Wildlife Authority (UWA), the Ugandan National Council of Science and Technology (UNCST) and the Makerere University for permission to work at the Makerere University Biological Field Station (MUBFS). For invaluable support at the Ngogo camp and in the field, we thank S. Angedakin, C. Aliganyira, C. Birungi, I. Clark, D. Kalunga, D. Kanweri, B. Kamugyisha, R. Katumba, K. Lee, G. Mbabazi, S. Namaganda, L. Ndan-gizi, S. Ramirez Amaya, A. Sandel, and A. Tumusiime. For help with coding we want to extend thanks to M. Sarria-Mosquera and S. Roske. For advice on statistical analyses, we thank R. Mundry and E. Willems. For thought-provoking discussions, invaluable feedback, and administrative help, we thank the members of the Comparative BioCognition research group.

Data availability

The data associated with this research are available at OSF: https://osf.io/gfdm3/?view_only=4a9457b74557421f8a0702847815b0c3

References

- Abreu, F., & Pika, S. (2022). Turn-taking skills in mammals: A systematic review into development and acquisition. *Frontiers in Ecology and Evolution*, 10. <https://doi.org/10.3389/fevo.2022.987253>
- Agresti, A. (2000). *Categorical data analysis* (Second edition ed.). John Wiley Sons, Inc.
- Altmann, J. (1974). Observational study of behaviour: Sampling methods. *Behaviour*, 49(3–4), 227–267.
- Amici, F., Ersson-Lembeck, M., Holodynski, M., & Liebal, K. (2023). Face to face interactions in chimpanzee (*Pan troglodytes*) and human (*Homo sapiens*) mother-infant dyads. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 378(1875), Article 20210478. <https://doi.org/10.1098/rstb.2021.0478>
- Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge University Press.
- Badescu, I., Sciotte, P., Sandel, A. A., Desruelle, K. J., Curteanu, C., Watts, D. P., & Sellen, D. W. (2020). Premasticated food transfer by wild chimpanzee mothers with their infants: Effects of maternal parity, infant age and sex, and food properties. *Journal of Human Evolution*, 143, Article 102794. <https://doi.org/10.1016/j.jhevol.2020.102794>
- Badihi, G., Graham, K. E., Grund, C., Safryghin, A., Soldati, A., Donnellan, E., ... Hobaiter, C. (2024). Chimpanzee gestural exchanges share temporal structure with human language. *Current Biology*, 34(14), R673–R674. <https://doi.org/10.1016/j.cub.2024.06.009>
- Bangerter, A., Genty, E., Heesen, R., Rossano, F., & Zuberbühler, K. (2022). Every product needs a process: Unpacking joint commitment as a process across species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1859), Article 20210095. <https://doi.org/10.1098/rstb.2021.0095>
- Bard, K. A., Dunbar, S., Maguire-Herring, V., Veira, Y., Hayes, K. G., & McDonald, K. (2014). Gestures and social-emotional communicative development in chimpanzee infants. *American Journal of Primatology*, 76(1), 14–29. <https://doi.org/10.1002/ajp.22189>
- Bard, K. A., Keller, H., Ross, K. M., Hewlett, B., Butler, L., Boysen, S. T., & Matsuzawa, T. (2021). Joint attention in human and chimpanzee infants in varied socio-ecological contexts. *Monographs of the Society for Research in Child Development*, 86(4), 7–217. <https://doi.org/10.1111/mono.12435>
- Bard, K. A., & Leavens, D. A. (2014). The importance of development for comparative primatology. *Annual Review of Anthropology*, 43, 183–200. <https://doi.org/10.1146/annurev-anthro-102313-030223>
- Bard, K. A., Maguire-Herring, V., Tomonaga, M., & Matsuzawa, T. (2019). The gesture “Touch”: Does meaning-making develop in chimpanzees’ use of a very flexible gesture? *Animal Cognition*, 22(4), 535–550. <https://doi.org/10.1007/s10071-017-1136-0>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, 263(1), 101–113.
- van Boekholt, B., Wilkinson, R., & Pika, S. (2024). Bodies at play: The role of intercorporeality and bodily affordances in coordinating social play in chimpanzees in the wild. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1206497>
- Boesch, C. (1994). Cooperative hunting in wild chimpanzees. *Animal Behaviour*, 48(3), 653–667.
- Brundl, A. C., Tkaczynski, P. J., Nohon Kohou, G., Boesch, C., Wittig, R. M., & Crockford, C. (2020). Systematic mapping of developmental milestones in wild chimpanzees. *Developmental Science*, e12988. <https://doi.org/10.1111/desc.12988>
- Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition*, 20(4), 755–769. <https://doi.org/10.1007/s10071-017-1096-4>
- Call, J., & Tomasello, M. (Eds.). (2007). *The gestural communication of apes and monkeys*. Lawrence Erlbaum Associates.
- Casillas, M., de Vos, C., Crasborn, O., & Levinson, S. C. (2015). The perception of stroke-to-stroke turn boundaries in signed conversation. In , 2015. *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 315–320). CogSci.
- Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2011). Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics*, 128(5), 910–916. <https://doi.org/10.1542/peds.2011-0609>
- Christensen, R. H. B. (2015). *Ordinal—Regression Models for Ordinal Data. R Package Version*, 28, 2015.
- Cicchetti, D. V. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 695–700. <https://doi.org/10.1076/jcen.23.5.695.1249>
- Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate

- procedures. *Child Neuropsychology*, 1(1), 26–37. <https://doi.org/10.1080/09297049508401340>
- Clark, H. (2006). Social actions, social commitments. In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 126–150). Berg.
- Clayman, S. E. (2013). Turn-constructional units and the transition-relevance place. *The Handbook of Conversation Analysis*, 151–166.
- Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, 22(2), 142–146. <https://doi.org/10.1016/j.cub.2011.11.053>
- Dafreville, M., Hobaiter, C., Guidetti, M., Sillam-Dusses, D., & Bourjade, M. (2021). Sensitivity to the communicative partner's attentional state: A developmental study on mother-infant dyads in wild chimpanzees (*Pan troglodytes schweinfurthii*). *American Journal of Primatology*, 83(12), Article e23339. <https://doi.org/10.1002/ajp.23339>
- Dezecache, G., Crockford, C., & Zuberbühler, K. (2019). The development of communication in alarm contexts in wild chimpanzees. *Behavioral Ecology and Sociobiology*, 73(8), 104. <https://doi.org/10.1007/s00265-019-2716-6>
- Dingemanse, M., & Enfield, N. J. (2024). Interactive repair and the foundations of language. *Trends in Cognitive Sciences*, 28(1), 30–42. <https://doi.org/10.1016/j.tics.2023.09.003>
- Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., ... Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS One*, 10(9), Article e0136100. <https://doi.org/10.1371/journal.pone.0136100>
- Dobson, A. J. (2002). *An introduction to generalized linear models*. Chapman & Hall/CRC.
- Dominguez, S., Devouche, E., Apter, G., & Gratier, M. (2016). The roots of turn-taking in the neonatal period. *Infant and Child Development*, 25(3), 240–255. <https://doi.org/10.1002/icd.1976>
- Fantasia, V., Markova, G., Fasulo, A., Costall, A., & Reddy, V. (2015). Not just being lifted: Infants are sensitive to delay during a pick-up routine. *Frontiers in Psychology*, 6, 2065. <https://doi.org/10.3389/fpsyg.2015.02065>
- Fedurek, P., Slocombe, K. E., Hartel, J. A., & Zuberbühler, K. (2015). Chimpanzee lip-smacking facilitates cooperative behaviour. *Scientific Reports*, 5, Article 13460. <https://doi.org/10.1038/srep13460>
- Fernandez-Carriba, S., Loeches, A., Morcillo, A., & Hopkins, W. D. (2002). Asymmetry in facial expression of emotions by chimpanzees. *Neuropsychologia*, 40(9), 1523–1533. [https://doi.org/10.1016/S0028-3932\(02\)00028-3](https://doi.org/10.1016/S0028-3932(02)00028-3)
- Filipi, A. (2009). *Toddler and parent interaction: The organisation of gaze, pointing and vocalisation* (Vol. 192). John Benjamins Publishing.
- Fischer, J., Cheney, D. L., & Seyfarth, R. M. (2000). Development of infant baboons' responses to graded bark variants. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 267, 2317–2321. <https://doi.org/10.1098/rspb.2000.1285>
- Fitch, W. T. (2005). The Evolution of Language: A Comparative Review. *Biology & Philosophy*, 20(2–3), 193–203. <https://doi.org/10.1007/s10539-005-5597-1>
- Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin and Review*, 24(1), 3–33. <https://doi.org/10.3758/s13423-017-1236-5>
- Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, 65(1), 47–55. <https://doi.org/10.1007/s00265-010-1038-5>
- Fröhlich, M. (2017). Taking turns across channels: Conversation-analytic tools in animal communication. *Neuroscience and Biobehavioral Reviews*, 80, 201–209. <https://doi.org/10.1016/j.neubiorev.2017.05.005>
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, 6, Article 25887. <https://doi.org/10.1038/srep25887>
- Fröhlich, M., & Pika, S. (2019). Gestural usage and development in two chimpanzee groups of different subspecies (*Pan troglodytes verus*/*P.t. schweinfurthii*). In *The Chimpanzees of the Tai Forest* (pp. 422–439). <https://doi.org/10.1017/9781108674218.027>
- Fröhlich, M., & van Schaik, C. P. (2022). Social tolerance and interactional opportunities as drivers of gestural redos in orang-utans. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1859), Article 20210106. <https://doi.org/10.1098/rstb.2021.0106>
- Fröhlich, M., Wittig, R. M., & Pika, S. (2016a). Play-solicitation gestures in chimpanzees in the wild: Flexible adjustment to social circumstances and individual matrices. *Royal Society Open Science*, 3(8), Article 160278. <https://doi.org/10.1098/rsos.160278>
- Fröhlich, M., Wittig, R. M., & Pika, S. (2016b). Should I stay or should I go? Initiation of joint travel in mother–infant dyads of two chimpanzee communities in the wild. *Animal Cognition*, 19(3), 483–500. <https://doi.org/10.1007/s10071-015-0948-z>
- Fröhlich, M., Wittig, R. M., & Pika, S. (2019). The ontogeny of intentional communication in chimpanzees in the wild. *Developmental Science*, 22(1), Article e12716. <https://doi.org/10.1111/desc.12716>
- Genty, E., & Fuchs, M. (2023). GAPS: A Coding Scheme for Great Apes Signals in ELAN (Version v1.0.2). <https://greatapesgestures.github.io/>.
- Genty, E., Heesen, R., Guéry, J.-P., Rossano, F., Zuberbühler, K., & Bangerter, A. (2020). How apes get into and out of joint actions. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, 21(3), 353–386. <https://doi.org/10.1075/is.18048.gen>
- Gillespie-Lynch, K., Greenfield, P., Lyn, H., & Savage-Rumbaugh, S. (2014). Gestural and symbolic development among apes and humans: Support for a multimodal theory of language evolution. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01228>, 1664–1078.
- Goffman, E. (1981). Footing. In E. Goffman (Ed.), *Forms of talk* (pp. 124–159). University of Pennsylvania Press.
- Good, J. S., & Beach, W. A. (2005). Opening up gift-openings: Birthday parties as situated activity systems. *Text - Interdisciplinary Journal for the Study of Discourse*, 25(5), 565–593. <https://doi.org/10.1515/text.2005.25.5.565>
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behaviour*. Belknap Press of Harvard University Press.
- Goodall, J. (1989). *Glossary of chimpanzee behaviors*. Jane Goodall Institute.
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse and Society*, 18, 53–73.
- Goodwin, C., & Goodwin, M. H. (2004). Participation. In A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 222–244). Blackwell Publishing. <https://doi.org/10.1002/9780470996522.ch20>
- Gramp, M., Samuni, L., Girard-Buttoz, C., Leon, 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*). *Royal Society Open Science*, 10(11), Article 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. *Frontiers in Psychology*, 6(1167), 236–245. <https://doi.org/10.3389/fpsyg.2015.01167>
- Greenfield, P. (1991). Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioural and Brain Sciences*, 14(4), 531–595.
- Haccou, P., & Meelis, E. (1992). *Statistical analysis of behavioural data: An approach based on time-structured models*. Oxford University Press.
- Halina, M., Rossano, F., & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal Cognition*, 16(4), 653–666. <https://doi.org/10.1007/s10071-013-0601-7>
- Heesen, R., Bangerter, A., Zuberbühler, K., Iglesias, K., Neumann, C., Pajot, A., ... Genty, E. (2021). Assessing joint commitment as a process in great apes. *iScience*, 24(8), Article 102872. <https://doi.org/10.1016/j.isci.2021.102872>
- Heesen, R., Fröhlich, M., Sievers, C., Woensdregt, M., & Dingemanse, M. (2022). *Coordinating social action: A primer for the cross-species investigation of communicative repair*. <https://doi.org/10.1098/rstb.2021.0110>
- Heesen, R., Zuberbühler, K., Bangerter, A., Iglesias, K., Rossano, F., Pajot, A., ... Genty, E. (2021). Evidence of joint commitment in great apes' natural joint actions. *Royal Society Open Science*, 8(12), Article 211121. <https://doi.org/10.1098/rsos.211121>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, 33(2–3), 61–83. discussion 83–135. <https://doi.org/10.1017/S0140525X0999152X>
- Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: A longitudinal study of mother–infant interaction. *Frontiers in Psychology*, 6(1492), 12. <https://doi.org/10.3389/fpsyg.2015.01492>
- Hirata, S. (2008). Communication between mother and infant chimpanzees and its role in the evolution of social intelligence. In S. Itakura, & K. Fujita (Eds.), *Origins of the social mind: Evolutionary and developmental views* (pp. 21–38). Japan: Springer. https://doi.org/10.1007/978-4-431-75179-3_2
- Hirata, S. (2009). Chimpanzee social intelligence: Selfishness, altruism, and the mother–infant bond. *Primates*, 50(1), 3–11. <https://doi.org/10.1007/s10329-008-0122-1>
- Hobaiter, C. (2011). *Gestural communication in wild chimpanzees* St. Andrews.
- Hobaiter, C., & Byrne, R. W. (2012). Gesture use in consortship: Wild chimpanzees' use of gesture for an 'evolutionarily urgent' purpose. In S. Pika, & K. Liebal (Eds.), *Developments in primate gesture research* (pp. 129–146). John Benjamins Publishing Company.
- Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, 14(5), 747–767. <https://doi.org/10.1007/s10071-011-0409-2>
- Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, 71(96), 13. <https://doi.org/10.1007/s00265-017-2325-1>
- Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology*, 24(14), 1596–1600. <https://doi.org/10.1016/j.cub.2014.05.066>
- Holle, H., & Rein, R. (2015). EasyDIAG: A tool for easy determination of interrater agreement. *Behavior Research Methods*, 47(3), 837–847. <https://doi.org/10.3758/s13428-014-0506-7>
- Holler, J. (2022). Visual bodily signals as core devices for coordinating minds in interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1859), Article 20210094. <https://doi.org/10.1098/rstb.2021.0094>
- Jaeggi, A. V., & Van Schaik, C. P. (2011). The evolution of food sharing in primates. *Behavioral Ecology and Sociobiology*, 65(11), 2125–2140. <https://doi.org/10.1007/s00265-011-1221-3>
- Kaye, K., & Wells, A. J. (1980). Mothers' jiggling and the burst—Pause pattern in neonatal feeding. *Infant Behavior and Development*, 3, 29–46. [https://doi.org/10.1016/0163-6383\(80\)80005-1](https://doi.org/10.1016/0163-6383(80)80005-1)
- Keller, H., Bard, K., Morelli, G., Chaudhary, N., Vicedo, M., Rosabal-Coto, M., ... Gottlieb, A. (2018). The Myth of Universal Sensitive Responsiveness: Comment on Mesman et al. (2017). *Child Development*, 89(5), 1921–1928. <https://doi.org/10.1111/cdev.13031>
- Kendon, A. (1990). *Spatial organization in social encounters: The F-formation system. Conducting interaction: Patterns of behavior in focused encounters*.
- Kendrick, K. H., Brown, P., Dingemanse, M., Floyd, S., Gipper, S., Hayano, K., ... Levinson, S. C. (2020). Sequence organization: A universal infrastructure for social action. *Journal of Pragmatics*, 168, 119–138. <https://doi.org/10.1016/j.pragma.2020.06.009>

- Kersken, V., Gomez, J. C., Liszkowski, U., Soldati, A., & Hobaiter, C. (2019). A gestural repertoire of 1- to 2-year-old human children: In search of the ape gestures. *Animal Cognition*, 22(4), 577–595. <https://doi.org/10.1007/s10071-018-1213-z>
- Koda, H. (2004). Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour*, 141(10), 1279–1296.
- Köndgen, S., Kühl, H., N'Goran, P. K., Walsh, P. D., Schenk, S., Ernst, N., ... Leendertz, F. H. (2008). Pandemic human viruses cause decline of endangered great apes. *Current Biology*, 18, 260–264. <https://doi.org/10.1016/j.cub.2008.01.012>
- Koski, S., Sterck, E., de Vries, H., & van den Tweel, S. (2007). What to do after a fight? The determinants and inter-dependency of post-conflict interactions in chimpanzees. *Behaviour*, 144(5), 529–555. <https://doi.org/10.1163/156853907780713082>
- Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs, & N. B. Davies (Eds.), *Behavioural ecology. An evolutionary approach* (pp. 380–402). Blackwell Scientific Publications.
- Langergraber, K., Mitani, J. C., & Vigilant, L. (2009). Kinship and social bonds in female chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, 71(10), 840–851. <https://doi.org/10.1002/ajp.20711>
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41(3), 841–849. <https://doi.org/10.3758/BRM.41.3.841>
- Lavelli, M., & Fogel, A. (2013). Interdyad differences in early mother–infant face-to-face communication: Real-time dynamics and developmental pathways. *Developmental Psychology*, 49(12), 2257–2271. <https://doi.org/10.1037/a0032268>
- Leendertz, F. H., Ellerbrok, H., Boesch, C., Couacy-Hymann, E., Matz-Rensing, K., Hakenbeck, R., ... Pauli, G. (2004). Anthrax kills wild chimpanzees in a tropical rainforest. *Nature*, 430(6998), 451–452. <https://doi.org/10.1038/nature02722>
- Levinson, S. C. (2006). On the human “interaction engine”. In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Berg.
- Levinson, S. C. (2016). Turn-taking in human communication – Origins and implications for language processing. *Trends in Cognitive Sciences*, 20(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 369(20130302), 9. <https://doi.org/10.1098/rstb.2013.0302>
- Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6(731), 10–26. <https://doi.org/10.3389/fpsyg.2015.00731>
- Liebal, K., Call, J., Tomasello, M., & Pika, S. (2004). To move or not to move. How apes adjust to the attentional state of others. *Interaction Studies*, 5(2), 199–219. <https://doi.org/10.1075/is.5.2.03lie>
- Liebal, K., Schneider, C., & Errson-Lembeck, M. (2019). How primates acquire their gestures: Evaluating current theories and evidence. *Animal Cognition*, 22(4), 473–486. <https://doi.org/10.1007/s10071-018-1187-x>
- Lowe, A. E., Hobaiter, C., Asiimwe, C., Zuberbühler, K., & Newton-Fisher, N. E. (2020). Intra-community infanticide in wild, eastern chimpanzees: A 24-year review. *Primates*, 61(1), 69–82. <https://doi.org/10.1007/s10329-019-00730-3>
- Martin, P. R., & Bateson, P. P. G. (1994). *Measuring behaviour: An introductory guide* (2 ed.). Cambridge University Press.
- Masataka, N., & Biben, M. (1987). Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. *Behaviour*, 101(4), 311–319.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall.
- Mitani, J. C. (2009). Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evolutionary Anthropology*, 18(5), 215–227.
- Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9(2), 194–225. <https://doi.org/10.1177/1461445607075346>
- Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41(10), 1977–1997. <https://doi.org/10.1016/j.pragma.2008.09.019>
- Mondada, L., & Meguerditchian, A. (2022). Sequence organization and embodied mutual orientations: Openings of social interactions between baboons. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1859), Article 20210101. <https://doi.org/10.1098/rstb.2021.0101>
- Mondémé, C. (2021). Why study turn-taking sequences in interspecies interactions? *Journal for the Theory of Social Behaviour*, 52(1), 67–85. <https://doi.org/10.1111/jtsb.12295>
- Mondémé, C. (2023). Sequence organization in human–animal interaction. An exploration of two canonical sequences. *Journal of Pragmatics*, 214, 73–88. <https://doi.org/10.1016/j.pragma.2023.06.006>
- Newton-Fisher, N. E., & Kaburu, S. S. K. (2017). Grooming decisions under structural despotism: The impact of social rank and bystanders among wild male chimpanzees. *Animal Behaviour*, 128, 153–164. <https://doi.org/10.1016/j.anbehav.2017.04.012>
- Newton-Fisher, N. E., & Lee, P. C. (2011). Grooming reciprocity in wild male chimpanzees. *Animal Behaviour*, 81(2), 439–446. <https://doi.org/10.1016/j.anbehav.2010.11.015>
- Nguyen, V., Versyp, O., Cox, C., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the development of turn taking in adult-child vocal interactions. *Child Development*, 93(4), 1181–1200. <https://doi.org/10.1111/cdev.13754>
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, 162, 31–38. <https://doi.org/10.1016/j.jecp.2017.04.017>
- Nishida, T., Kano, T., Goodall, J., McGrew, W. C., & Nakamura, M. (1999). Ethogram and ethnography of Mahale chimpanzees. *Anthropological Science*, 107(2), 141–188.
- Nomikou, I., Leonardi, G., Radkowska, A., Raczaszek-Leonardi, J., & Rohlfing, K. J. (2017). Taking up an active role: Emerging participation in early mother–infant interaction during peekaboo routines. *Frontiers in Psychology*, 8, 1656. <https://doi.org/10.3389/fpsyg.2017.01656>
- Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion*, 7(1), 172–181. <https://doi.org/10.1037/1528-3542.7.1.172>
- Peckre, L., Kappeler, P. M., & Fichtel, C. (2019). Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behavioral Ecology and Sociobiology*, 73(1). <https://doi.org/10.1007/s00265-018-2605-4>
- Pika, S. (2009). Our grooming cousins: Providing the link to declarative signalling? *Studies in Communication Sciences*, 9(1), 73–102.
- Pika, S. (2015). Gestural communication in nonhuman species. In R. Scott, & S. Kosslyn (Eds.), *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource* (pp. 1–11). Wiley Online Library.
- Pika, S., & Bugnyar, T. (2011). The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nature Communications*, 2(560), 5. <https://doi.org/10.1038/ncomms1567>
- Pika, S., & Deschner, T. (2019). A new window onto animal culture the case of chimpanzee gesturing. *Gesture*, 18(2–3), 239–260. <https://doi.org/10.1075/gest.19012.pik>
- Pika, S., & Fröhlich, M. (2019). Gestural acquisition in great apes: The social negotiation hypothesis. *Animal Cognition*, 22(4), 551–565. <https://doi.org/10.1007/s10071-017-1159-6>
- Pika, S., & Mitani, J. C. (2006). Referential gestural communication in wild chimpanzees (*Pan troglodytes*). *Current Biology*, 16(6), R191–R192.
- Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B: Biological Sciences*, 285(20180598), 9. <https://doi.org/10.1098/rspb.2018.0598>
- Plooi, F. X. (1978). Some basic traits of language in wild chimpanzees? In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language* (pp. 111–131). Academic Press.
- Plooi, F. X. (1979). How wild chimpanzee babies trigger the onset of mother–infant play - and what the mother makes of it. In M. Bullowa (Ed.), *Before speech: The beginning of interpersonal communication* (pp. 223–243). Cambridge University Press.
- Pougnault, L., Lemasson, A., Mulot, B., & Levréro, F. (2021). Temporal calling patterns of a captive group of chimpanzees (*Pan troglodytes*). *International Journal of Primatology*, 42(6), 809–832. <https://doi.org/10.1007/s10764-021-00262-y>
- Pougnault, L., Levréro, F., Leroux, M., Paulet, J., Bombani, P., Dentressangle, F., ... Lemasson, A. (2022). Social pressure drives “conversational rules” in great apes. *Biological Reviews of the Cambridge Philosophical Society*, 97(2), 749–765. <https://doi.org/10.1111/brv.12821>
- Pougnault, L., Levréro, F., Mulot, B., & Lemasson, A. (2020). Breaking conversational rules matters to captive gorillas: A playback experiment. *Scientific Reports*, 10(1), 6947. <https://doi.org/10.1038/s41598-020-63923-7>
- R Core Team, T. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS One*, 8(6), Article e65289. <https://doi.org/10.1371/journal.pone.0065289>
- Roberts, A. I., Vick, S.-J., & Buchanan-Smith, H. M. (2012). Usage and comprehension of manual gestures in wild chimpanzees. *Animal Behaviour*, 84(2), 459–470. <https://doi.org/10.1016/j.anbehav.2012.05.022>
- Rodrigues, E. D., Santos, A. J., Hayashi, M., Matsuzawa, T., & Hobaiter, C. (2021). Exploring greetings and leave-takings: Communication during arrivals and departures by chimpanzees of the Bossou community, Guinea. *Primates*. <https://doi.org/10.1007/s10329-021-00957-z>
- Rossano, F. (2013a). Gaze in Conversation. In J. Sidnell, & T. Stivers (Eds.), *The handbook of conversation analysis* (p. 825). Wiley-Blackwell.
- Rossano, F. (2013b). Sequence organization and timing of bonobo mother–infant interactions. *Interaction Studies*, 14(2), 160–189. <https://doi.org/10.1075/is.14.2.02ros>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50(4), 696–735.
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS One*, 8(10), Article e76674. <https://doi.org/10.1371/journal.pone.0076674>
- Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology*, 20(2), 416–420. <https://doi.org/10.1093/beheco/arn145>
- Schneider, C., Call, J., & Liebal, K. (2011). Onset and early use of gestural communication in nonhuman great apes. *American Journal of Primatology*, 73, 1–12. <https://doi.org/10.1002/ajp.21011>
- Schneider, J. L., Roemer, E. J., Northrup, J. B., & Iverson, J. M. (2022). Dynamics of the dyad: How mothers and infants co-construct interaction spaces during object play. *Developmental Science*, Article e13281. <https://doi.org/10.1111/desc.13281>
- Sidnell, J., & Stivers, T. (2013). *The handbook of conversation analysis*. Wiley-Blackwell.
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Soldati, A., Muhumuza, G., Dezeceache, G., Fedurek, P., Taylor, D., Call, J., & Zuberbühler, K. (2022). The ontogeny of vocal processes: Insights from a newborn Wild chimpanzee (*Pan troglodytes schweinfurthii*). *International Journal of Primatology*, 44(1), 116–139. <https://doi.org/10.1007/s10764-022-00321-y>
- Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother–infant dyad. *Annals of the New York Academy of Sciences*, 263, 89–100. <https://doi.org/10.1111/j.1749-6632.1975.tb41574.x>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... Levinson, S. C. (2009). Universals and cultural variation in turn-taking in

- conversation. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Sugiura, H. (2001). Vocal exchange of coo calls in Japanese macaques. *Primate Origins of Human Cognition and Behavior*, 135–154.
- Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society Series B*, 371(1693), Article 20150370. <https://doi.org/10.1098/rstb.2015.0370>
- Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Current Biology*, 23, 2162–2168. <https://doi.org/10.1016/j.cub.2013.09.005>
- Tomasello, M. (2008). *Origins of human communications*. MIT Press.
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35(2), 137–154.
- Tomasello, M., & Zuberbühler, K. (2002). Primate vocal and gestural communication. In M. Bekoff, C. Allen, & G. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives in animal cognition* (pp. 293–299). MIT Press.
- Van Egeren, L. A., Barratt, M. S., & Roach, M. A. (2001). Mother-infant responsiveness: Timing, mutual regulation, and interactional context. *Developmental Psychology*, 37(5), 684–697. <https://doi.org/10.1037/0012-1649.37.5.684>
- de Vos, C., Torreira, F., & Levinson, S. C. (2015). Turn-timing in signed conversations: Coordinating stroke-to-stroke turn boundaries. *Frontiers in Psychology*, 6(268), 13. <https://doi.org/10.3389/fpsyg.2015.00268>
- Watts, D. P. (2000). Grooming between male chimpanzees at Ngogo, Kibale National Park, Uganda. I. Partner number and diversity and reciprocity. *International Journal of Primatology*, 21(2), 189–210.
- Watts, D. P. (2016). Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: Variation, social learning, and possible functions. *Primates*, 57, 61–72. <https://doi.org/10.1007/s10329-015-0497-8>
- Watts, D. P., & Mitani, J. C. (2002). Hunting behavior of chimpanzees at Ngogo, Kibale National Park, Uganda. *International Journal of Primatology*, 23(1), 1–28.
- Wheeler, B. C., Searcy, W. A., Christiansen, M. H., Corballis, M. C., Fischer, J., Grüter, C., ... Wild, M. (2011). Communication. In R. Menzel, & J. Fischer (Eds.), *Animal thinking: Contemporary issues in comparative cognition* (pp. 187–205). MIT Press.
- Whiten, A., & van de Waal, E. (2018). The pervasive role of social learning in primate lifetime development. *Behavioral Ecology and Sociobiology*, 72(5), 80. <https://doi.org/10.1007/s00265-018-2489-3>
- Wilke, C., Lahiff, N. J., Badihi, G., Donnellan, E., Hobaiter, C., Machanda, Z. P., ... Slocombe, K. E. (2022). Referential gestures are not ubiquitous in wild chimpanzees: Alternative functions for exaggerated loud scratch gestures. *Animal Behaviour*, 189, 23–45. <https://doi.org/10.1016/j.anbehav.2022.04.007>
- Wilke, C., Lahiff, N. J., Sabbi, K. H., Watts, D. P., Townsend, S. W., & Slocombe, K. E. (2022). Declarative referential gesturing in a wild chimpanzee (*Pan troglodytes*). *Proceedings of the National Academy of Sciences of the United States of America*, 119(47), Article e2206486119. <https://doi.org/10.1073/pnas.2206486119>
- Wilkinson, R., Leudar, I., & Pika, S. (2012). Requesting behaviours within episodes of active sharing. A new look on chimpanzee signalling. In S. Pika, & K. Liebal (Eds.), *Developments in primate gesture research* (pp. 199–221). John Benjamins Publishing Company.
- Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., ... Koops, K. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature*, 513(7518), 414–417.