Compositionality in Primate Gestural Communication and Multicomponent Signal Displays

Federica Amici1,2 · Linda Oña3 · Katja Liebal1

Received: 6 December 2021 / Accepted: 24 June 2022
© The Author(s) 2022

Abstract
Compositionality is the ability to combine meaningful elements into new combinations with novel meanings, and it has long been considered one of the main hallmarks of human communication. However, very few studies have addressed the compositional aspects of communication in species other than humans, although a comparative approach is essential to understand the evolutionary origins of human compositionality. We review previous research on compositionality in the gestural communication systems of nonhuman primates, with a special focus on the multi-component aspects of compositionality. We start by discussing the importance of a comparative approach to study the evolution of human language and then compare the current state of the art on compositionality in the vocal, facial, and gestural communication systems of primates and other species. We further discuss alternative approaches to study compositionality in primates, which may help overcome some of the current methodological limitations in this research area. In particular, we 1) highlight the importance of interdisciplinary tools that facilitate the statistical identification of multicomponent and multimodal combinations of signals, 2) discuss different approaches to infer the meaning of signal combinations, with a special focus on the use of contextual cues and meta-communication, and 3) discuss temporal and intentional aspects of compositionality in primates. Finally, we outline possible lines of research for future studies in this area (e.g., more consistent use of terms across research areas, use of different methodological tools and larger datasets, inclusion of developmental approaches), which might shed light into the evolutionary origins of one of the most crucial properties of human communication.

Keywords Compositionality · Combinatoriality · Multicomponent and multimodal signal combinations · Meta-communication

Handling Editor: Joanna M. Setchell

* Federica Amici
amici@eva.mpg.de

Extended author information available on the last page of the article

Published online: 28 July 2022
Abstract in Spanish
La composicionalidad es la capacidad de combinar elementos significativos en nuevas combinaciones con nuevos significados, y desde hace tiempo se considera una de las principales marcas típicas de la comunicación humana. Sin embargo, muy pocos estudios han abordado los aspectos compositivos de la comunicación en otras especies, aunque un enfoque comparativo es esencial para entender los orígenes evolutivos de la composicionalidad humana. En este artículo revisamos las investigaciones que se han conducido sobre la composicionalidad en la comunicación gestual de los primates no humanos, con un enfoque especial en los aspectos multimodales de la composicionalidad. Empezamos discutiendo la importancia de un enfoque comparativo para estudiar la evolución del lenguaje humano, y luego comparamos nuestros conocimientos actuales sobre la composicionalidad en los sistemas de comunicación vocal, facial y gestual de los primates y otras especies. En las siguientes secciones, analizamos enfoques alternativos para estudiar la composicionalidad en los primates, que pueden ayudar a superar algunas de las limitaciones metodológicas actuales en este ámbito de investigación. En particular, (i) destacamos la importancia de las herramientas interdisciplinarias que facilitan la identificación estadística de las combinaciones multicomponentes y multimodales de señales, (ii) discutimos diferentes enfoques para inferir el significado de las combinaciones de señales, con un enfoque especial en contexto y metacomunicación, y (iii) discutimos los aspectos temporales e intencionales de la composicionalidad en primates. Por último, esbozamos posibles líneas de investigación para futuros estudios en esta área (por ejemplo, un uso más coherente de los términos en distintas áreas de investigación, el uso de diferentes herramientas metodológicas y de conjuntos de datos más grandes, la inclusión de aspectos del desarrollo), que podrían arrojar luz sobre los orígenes evolutivos de una de las propiedades más cruciales de la comunicación humana. *The translated abstract was not copy-edited by Springer Nature.*

Abstract in German
Kompositionalität ist eines der Hauptmerkmale menschlicher Sprache und bezeichnet die Fähigkeit, bedeutungsvolle Elemente zu neuen Kombinationen mit neuartigen Bedeutungen zu kombinieren. Obwohl ein vergleichender Ansatz für das Verständnis der evolutionären Ursprünge der Kompositionalität menschlicher Sprache unerlässlich ist, gibt es bisher nur sehr wenige Studien, die sich mit kompositionellen Aspekten der Kommunikation bei anderen Tierarten befasst haben. In diesem Artikel geben wir einen Überblick über die bisherige Forschung zur Kompositionalität nicht-menschlicher Primaten, wobei wir den Schwerpunkt auf gestische Kommunikationssysteme und multimodale Aspekte der Kompositionalität legen. Wir beginnen mit einer Erörterung der Bedeutung eines artvergleichenden Ansatzes für die Untersuchung der Evolution der menschlichen Sprache, und vergleichen dann den aktuellen Stand der Forschung zur Kompositionalität in den vokalen, mimischen und gestischen Kommunikationssystemen von Primaten und anderen Arten. Darüber hinaus diskutieren wir alternative Ansätze zur Untersuchung der Kompositionalität bei Primaten, die dazu beitragen können, einige der derzeitigen methodischen Beschränkungen in diesem Forschungsbereich zu überwinden. Insbesondere (i) heben wir die Bedeutung von statistischen Methoden vor, welche die Identifizierung von
aus mehreren Komponenten bestehenden sowie multimodalen Signalkombinationen erleichtern, (ii) diskutieren wir verschiedene Ansätze zur Ableitung der Bedeutung von Signalkombinationen, mit Fokus auf der Verwendung von kontextuellen Hinweisen und Metakommunikation, und (iii) diskutieren zeitliche und intentionale Aspekte der Kompositionality bei Primaten. Abschließend skizzieren wir mögliche Forschungslinien für zukünftige Studien in diesem Bereich (z.B. konsistente Verwendung von Begriffen in verschiedenen Forschungsbereichen, Verwendung verschiedener methodischer Ansätze und größerer Datensätze, Bewertung von Entwicklungsaspekten), die Licht in die evolutionären Ursprünge einer der wichtigsten Eigenschaften menschlicher Kommunikation bringen können. *The translated abstract was not copy-edited by Springer Nature.*

**Introduction**

Linguistic scholars have traditionally differentiated between two typical properties of human language: combinatoriality and compositionality. Although these terms are not used consistently, combinatoriality generally refers to the ability to combine meaningless sounds into meaningful morphemes and words, whereas compositionality usually refers to the ability to further recombine already meaningful elements (i.e., morphemes and words) into new elements with novel meanings (de Boer et al., 2012; Hockett, 1960). Therefore, whereas combinatoriality is characterized by non-meaningful components, compositionality is assumed to imply the use of meaningful elements. Both combinatoriality and compositionality convey productivity and flexibility to human languages, allowing them to become potentially open-ended despite being based on a limited number of initial components (Chomsky, 1981; Fitch, 2010; Jackendoff, 2011; Werning et al., 2012). Combinatoriality and compositionality are generally considered universal properties of natural human languages that not only characterize spoken language, but also sign language and different forms of human nonverbal communication, including the expressions of emotions via facial and bodily displays (Cavicchio et al., 2018; Sandler & Lillo-Martin, 2006).

In this paper, we review the current state of the art on compositionality in nonhuman animals, with a special focus on gestural communication and multicomponent signal displays in nonhuman primates (hereafter, primates). We trace the evolutionary origins of human compositionality in the multicomponent communication systems of other species to identify possible lines of future research based on the main limitations of current research on this topic. We especially focus on compositionality, rather than combinatoriality, because evidence of compositionality is still scant in primates, likely because compositionality implies the creation of novel meanings (through the recombination of different components) and therefore is not easy to prove. We start our review by discussing the importance of a comparative approach to the study of language evolution, providing evidence of compositionality in the vocal communication systems of nonhuman species. We then describe the current state of the art on compositionality in the vocal, facial, and gestural communication systems of primates and other species and review alternative approaches to study compositionality in primates, which may
help overcome some of the current methodological limitations in this research area. In particular, we first focus on the importance of using a multicomponent and multimodal approach to the study of compositionality. We then introduce different approaches that are used to infer the meaning of signals and signal combinations, with a special focus on the use of contextual cues and meta-communication. We further discuss temporal and intentional aspects of compositionality in primates to conclude with possible lines for future research on compositionality, especially stressing the importance of large data-sets and large-scale collaborations to overcome current limitations in the study of compositionality.

Comparative Approach to the Evolution of Human Language

According to some authors, language evolved from scratch in humans and fundamentally differs from other animals’ communication systems, so that a comparative approach would not be informative in searching for the evolutionary origins of human language (Chomsky, 1966; Scott-Phillips, 2015). However, a trait as complex as language is unlikely to have evolved in such a short time only in the human lineage (Pinker & Bloom, 1990), and it is probable that language rather built on traits already present in the common ancestors of other species, including neurobiological substrates (Arbib, 2005, 2017), anatomical structures, such as the vocal apparatus (Fitch et al., 2016; Riede et al., 2005) and cognitive skills (Seyfarth et al., 2005). Therefore, comparisons with other species, and in particular with nonhuman primates, may be very useful to identify the potential precursors to human language (Tomasello, 2008; Zuberbühler, 2005).

To identify the possible precursors to human language, however, researchers focus on different aspects of primate communication, largely depending on which communicative modality is assumed as the origin of human language (e.g., gestural, vocal or orofacial origin; Slocombe et al., 2011). As a consequence, the majority of existing research on primate communication is unimodal, and researchers usually study one specific modality in isolation, by using different theoretical and methodological approaches (Liebal et al., 2022; Slocombe et al., 2011). The majority of studies has been conducted on vocalizations, which are mostly studied in wild and captive monkeys, often with a focus on their functionally referential usage in very specific contexts. Gestures are mainly studied in great apes, mostly in captive settings, with a focus on the signaler and the intentional and flexible use of gestural repertoires across different contexts. Facial expressions are frequently studied in monkeys, also mostly in captive settings, with a focus on the adaptive function of these signals (Liebal et al., 2013; Slocombe et al., 2011; Waller et al., 2016). This means that comparative approaches to compositionality in primates are addressed differently, as researchers working on gestures, vocalizations, or facial expressions focus on different aspects of primate communication. Furthermore, our current knowledge on primate communication is still limited, because the majority of studies on this topic has been conducted on vocalizations, and very few studies have investigated
Compositionality in Primate Gestural Communication and…

the multicomponent or multimodal combination of signals (see 3.1). Moreover, studies on compositionality are largely limited to very few species. This problem affects the whole research area of communication, with 24.5% of the published unimodal studies having been conducted in chimpanzees (*Pan troglodytes*), 14.6% in rhesus macaques (*Macaca mulatta*) and 8.8% in common marmosets (*Callithrix jacchus*), so that most species are severely underrepresented (Liebal et al., 2022).

**Current State of the Art in Research on Compositionality in Primates and Other Species**

The purpose of the next two sections is to summarize the current state of the art on research into compositionality, separately for the vocal communication system and the gestural and facial communication systems of nonhuman species and, especially, primates. First, we discuss current evidence of compositionality in primate vocal communication, also referring to previous work on nonprimate species, such as birds, whose vocal systems (including aspects of compositionality) have been the focus of abundant research. Second, we review studies on compositionality in gestural and facial communication. As these studies, to our knowledge, have been conducted exclusively on primates, we will also briefly discuss research on humans for comparative purposes.

**Compositionality and Other Properties of the Vocal Communication Systems of Nonhuman Species**

For several decades, compositionality has been considered one of the hallmarks of human language, setting human communication apart from that of other animals (Hurford, 2011). In past years, however, experimental evidence has shown that several species do not only use complex repertoires to communicate with conspecifics but that their communication systems also often show properties that had long been considered to be uniquely human. Research on vocal communication in nonhuman species, in particular, has shown several similarities with humans in several important aspects of vocal communication: from vocal learning to phonology and syntax (Fishbein et al., 2019).

Several taxa, for instance, have complex sequential calls with regularly recurring patterns, hierarchical structures and/or specific ordering rules, suggesting that these properties of communication are phylogenetically widespread and likely to have emerged through convergent evolution (e.g., primates: Clarke et al., 2006; Hedwig et al., 2015; Girard-Buttoz et al., 2021; cetaceans: Shapiro et al., 2010; Allen et al., 2019; bats: Bohn et al., 2009; rock hyraxes, *Procavia capensis*: Kershenbaum et al., 2012; birds: Sasahara et al., 2012; Deslandes et al., 2014; Weiss et al., 2014; Cody et al., 2016; Hedley, 2016). Moreover, the communication systems of several nonhuman species share the same statistical patterns of human language. In line with Zipf’s law of brevity (Zipf, 1949), for example, more common vocalizations tend
to be shorter in several species (e.g., Formosan macaques, *Macaca cyclopis*: Semple et al., 2010; bat species: Luo et al., 2013; male rock hyraxes: Demartsev et al., 2019; penguins: Favaro et al., 2020). In line with Menzerath’s law (Menzerath, 1954), longer sequences of vocalizations are composed of shorter calls in geladas (*Theropithecus gelada*) (Gustison et al., 2016) and gibbons (*Nomascus nasutus* and *N. concolor*) (Huang et al., 2020). Regularities in the structure of communication systems do not necessarily provide information about the meaning of sequences and about the compositional aspects of communication but suggest that some properties that constrain human language evolution are largely shared across species, and the divide between communication in humans and other animals may be narrower than previously thought.

Despite abundant evidence of combinatoriality in the vocal communication systems of several nonhuman species (Townsend et al., 2018), evidence of compositionality is however scanter. Among birds, compositionality has been shown in pied babblers (*Turdoides bicolor*; Engesser et al., 2016), and in Japanese great tits (*Parus minor*), who react differently to single notes and to their combinations (Suzuki et al., 2016). Among primates, evidence of compositionality has been found in putty-nosed monkeys (*Cercopithecus nictitans*), who can combine specific alarm calls into new vocalizations that have a novel meaning (Arnold & Zuberbühler, 2006, 2008). Furthermore, Campbell’s monkeys (*Cercopithecus campbelli*) can combine meaningful vocalizations into context-specific sequences (Ouattara et al., 2009a) and modify the meaning of their alarm calls by adding a suffix to the call (Ouattara et al., 2009b). For instance, adult males combine loud calls into specific sequences that are more likely used in certain contexts (e.g., in the presence of specific predators, or other nonpredatory animals) and which may vary depending on how the information transmitted was acquired (Ouattara et al., 2009a). Moreover, chimpanzees can combine pant hoots and food calls in a context-specific way, suggesting that the new sequences may acquire a novel meaning (Leroux et al., 2021).

**Compositionality in Primate Gestural and Facial Communication**

Although the previous section suggested that some primate species appear to compositionally recombine vocalizations, there is little evidence that they also can combine nonvocal signals (i.e., gestures and facial expressions) into sequences with a novel meaning. To date, however, there is virtually no research on the sequential combination of facial expressions in nonhuman species, although such combinations are sometimes referred to as “blended displays” (i.e., displays that share characteristics with at least two other prototypical facial expressions, but have no different meaning from them; Parr et al., 2005). In chimpanzees, for instance, blended displays often are used in the same context in which one of their composing elements are typically produced, suggesting that combinations of facial expressions may mirror conflicting emotional states, rather than convey novel meanings (Parr et al., 2005).

With regards to gestural communication, primates have relatively large repertoires (Hobaiter & Byrne, 2011; Liebal et al., 2013). However, the few studies that have investigated sequences of gestures in primate communication systems have found no
evidence that sequences have a novel meaning as compared to their elements (e.g.,
western gorillas, Gorilla: Genty & Byrne, 2010; Tanner & Perlman, 2017; Sumatran
orangutans, Pongo abelii: Tempelmann & Liebal, 2012; chimpanzees: Liebal et al.,
2004; Hobaiter & Byrne, 2011). In particular, although these studies have offered
different conclusions about the emergence of gesture combinations, none of them
has found evidence that primates combine gestures into longer sequences to spe-
cifically create novel meanings. Captive chimpanzees, for example, frequently pro-
duce sequences of gestures, but the majority of these sequences consist of simple
repetitions of the same gestures, mostly tactile ones, which likely serve to increase
recipients’ responsiveness (Liebal et al., 2004). Wild chimpanzees also use long
and largely redundant sequences of rapid-fire variable gestures (Hobaiter & Byrne,
2011). However, older individuals are more likely to use single gestures, rather than
sequences, suggesting that sequences are mostly used by inexperienced individuals
to increase their chances of obtaining a response and to learn how to communicate
more effectively with conspecifics (Hobaiter & Byrne, 2011). As for chimpanzees,
Sumatran orangutans usually produce sequences of gestures mostly consisting of rep-
etitions of the same gestures (Tempelmann & Liebal, 2012). Moreover, orangutans
often continue to gesture regardless of whether the recipient responds, suggesting that
sequences of gestures in this species may be mainly triggered by emotional arousal
(Tempelmann & Liebal, 2012). Therefore, across primate species, there seem to be
no significant differences between the meaning of a gesture when produced as part of
a sequence and when used alone, casting doubt on the existence of compositionality
in primate gestural communication (Liebal et al., 2004).

New Approaches to Study Compositionality in Nonhuman Species

Why should primates be able to compositionally combine vocalizations, but not
other signals? Much of primate social and technological behaviour contains hier-
archically structured sequences of actions (Estienne et al., 2017), and primates are
surprisingly good in understanding hierarchically structured information (Watson
et al., 2020). Therefore, it is unlikely that compositionality is limited to the vocal
communication system because of cognitive constraints, and methodological issues
may better explain why there is not yet clear evidence of compositionality in primate
gestural and facial communication. Therefore, the following sections will focus on
some blind spots in comparative approaches to compositionality in nonhuman spe-
cies, specifically primates, and make some suggestions on how to address them.

(Another) Call for Multicomponent and Multimodal Approaches to Study
Compositionality in Primate Communication

One reason why primates may not show evidence of compositionality in their ges-
tures or facial expressions is that primate communication is multicomponent and
multimodal (Slocombe et al., 2011; Liebal et al., accepted), and compositionality
might occur through the recombination of different elements within and across
modalities. In particular, whereas multimodal combinations of signals include signals produced in different sensory modalities (e.g., visual, auditory, tactile) (Rowe, 1999), multicomponent combinations include multiple signals regardless of their modality (e.g., gestures and facial expressions) (Micheletta et al., 2013; Waller et al., 2013, 2022). However, very few studies have investigated whether primates compositionally recombine elements to create new meaningful elements with a multicomponent and/or multimodal approach.

A recent study on red-capped mangabeys (Cercocebus torquatus), for instance, investigated combinations of signals across different modalities (Aychet et al., 2021) and identified 424 combinations. The majority involved multiple components, often in more than one modality. Combinations differed in their complexity depending on the social context in which they were used, being more complex in playful and partially aggressive contexts compared with others (Aychet et al., 2021). Other studies have used such a multicomponent or multimodal approach but have not specifically assessed compositionality in these combinations. Some researchers have investigated how primates combine gestures with vocalizations but did not specifically assess whether these combinations elicit a different response compared with their single components (Hobaiter et al., 2017; Wilke et al., 2017). In wild chimpanzees, in particular, gesture-vocal combinations occurred relatively seldom, but they were found in all age groups (Hobaiter et al., 2017; Wilke et al., 2017). Moreover, primates often used multimodal combinations after vocal signals that failed to elicit recipients’ response (Hobaiter et al., 2017), suggesting that primates might recur to combinations of signals to increase the probability of eliciting a response, rather than to convey novel meanings. Multimodal combinations were indeed more likely to elicit a response as compared to the single vocal elements they included (Wilke et al., 2017), although in these studies there was unfortunately no assessment of the kind of responses elicited by combinations and their elements.

Other researchers investigated multicomponent combinations in chimpanzees, focusing on the combination of gestures and facial expressions in a semi-wild setting (Oña et al., 2019). In this study, different combinations elicited different responses: adding a specific facial expression to a gesture modified the likelihood of a specific response to occur (Oña et al., 2019). In particular, the bared-teeth face was linked to an increase in recipients’ affiliative behaviour when it was combined with one gesture type (i.e., stretched arm gesture), but not with a different gesture type (i.e., bent arm gesture). Thus, this study did not conclude that the different facial and gestural components have specific meanings and that new meanings are consistently created through their combination (i.e., compositionality), but it suggested that facial expressions may fulfil the important function of modifying following or co-occurring gestures—a property that the authors referred to as “componentiality” (Oña et al., 2019).

**Defining the Meaning of Signals: Challenges and New Approaches**

A second reason why primates may be able to compositionally combine vocalizations, but there is no clear evidence that also may be compositionally combine other
signals, might be that, at least in gestural studies, the “meaning” of signals has been traditionally inferred from the context in which signals are used (Liebal & Oña, 2018). Inferring the meaning of a gesture from the context in which it takes place, however, might be misleading, not only because some gestures are used in a variety of contexts, but also because, even within the same context, gestures may convey very different information (Hobaiter & Byrne, 2014). Therefore, a more promising approach may be to infer the meaning of gestures by monitoring the response they elicit from their recipients (Genty et al., 2009; Hobaiter & Byrne, 2014). When recipients react to a signal by producing nonaversive behaviour and signallers stop gesturing, the recipients’ behaviour may be considered to approximate the signallers’ goal and thus the meaning of the signal itself (Hobaiter & Byrne, 2014). Studying compositionality in gestural communication with this more precise operational definition of “meaning” might lead to different results.

Gestures can be combined in many different ways, and used in a variety of contexts, so that vast datasets are required to obtain a comprehensive perspective on the possible meaning of gestures and their multicomponent combinations. Moreover, many combinations of signals are likely to contain elements of persistence and emphasis (Liebal et al., 2004), simply changing the intensity of the information conveyed. Therefore, in many cases, combinations of signals may trigger the same response as the single signals, despite having a different meaning (at least in terms of intensity). Large datasets including further measures of recipient’s response may help detecting these subtler changes in the meaning of signals. When the combination of signals alters the meaning of the single signals only in term of intensity, for example, this change may be captured by measuring recipients’ reaction time.

Even the identification of single gestures is problematic. In chimpanzees, for instance, the number of different gesture types produced varies from more than one hundred (Roberts et al., 2014) to less than 30 gestures (Pika et al., 2005), depending on how gestures are operationally defined. These top-down classifications are largely based on researchers’ intuition, but finer-grained distinctions based on bottom-up approaches that focus on the different forms of single gestures may reveal a more complex picture (Bard et al., 2019). For instance, have we really identified the meaningful units of gestural communication, or are current gestures themselves combinations of independent actions that convey meaning to the recipient? All our findings, so far, are conditional on the assumption that we have properly identified primate signals—something we cannot unfortunately be sure of.

**Context: From Gestural Meaning to Meaning Modulator**

With regard to meaning, the context also might play a crucial role, as one signal may convey different meanings depending on the context in which it is used. In chimpanzees, for instance, the bared teeth face modulated recipients’ response to different arm gestures, but only during affiliative events (Oña et al., 2019). Therefore, it is possible that individuals may attribute a different meaning to the same signals also depending on the context in which they are used. In bonobos (Pan paniscus), for instance, gestures can have different meanings depending on the specific context in
which they are produced, whereas they do not appear to acquire a novel meaning when combined with other gestures (Graham et al., 2020).

However, context has been often used as a proxy of meaning in several studies, but rarely used as a possible modifier of signal meaning. Moreover, the study of compositionality is often limited to very few contexts. Several studies, for instance, have investigated how certain signals acquire a novel meaning in the playful context, where individuals exchange unpredictable behavioural patterns that may easily escalate into overt aggression, so that clearly conveying the playful meaning through specific signals may be highly adaptive (Bateson, 1955; Burghardt, 2005; Pellis & Pellis, 2009; Spinka et al., 2016). However, signals may be compositionally combined in other contexts (e.g., aggression or feeding), and they may alter the meaning of behavioural elements that would otherwise trigger different responses, well beyond the play context. Unfortunately, data in contexts other than play are extremely scant in literature, likely because it is difficult to obtain high-quality information about multicomponent signals in sufficient contexts, especially in the wild. Therefore, future studies on compositionality should systematically account for the several contexts in which single elements and their combinations are produced.

**Meta-Communication as a Complimentary Approach to Study Compositionality in Primate Communication**

Meta-communication is typically used to refer to those signals (“secondary communication”) that alter the meaning of other behavioural elements (“primary communication”; Bateson, 1955; Mitchell, 1991, for a detailed discussion about the use of this term). Meta-communicative signals, for instance, are those that convey a playful meaning to behaviours that otherwise belong to the aggressive repertoire of a species, such as hitting or wrestling, or that are generally used in other functional contexts (Bateson, 1956; Bekoff & Allen, 1998; Fagen, 1981; Pellis & Pellis, 2009). In contrast to compositional combinations, meta-communicative signals add an extra layer to the communicative interaction, by communicating the true intention of an action or signal, which is produced in parallel (e.g., a play face signalling the playful intention before producing a potentially agonistic hit gesture). However, as for compositional combinations, signals and behavioural elements acquire a different meaning when they are combined compared with when they are independently produced. Therefore, meta-communication may prove an important complimentary concept for studying compositionality, and more specifically, to investigate the change of a signal’s meaning when combined with another and different signal type.

In primates, a variety of signals might serve a meta-communicatory function, including facial expressions, body gestures and vocalizations (Bekoff, 1972, 1995; Yanagi & Berman, 2014). Rhesus monkeys, for instance, use specific body movements (e.g., crouch and stare, gambolling, roll-onto-back-and-stare) to signal the onset of social play, and/or facilitate and moderate its occurrence (Sade, 1973; Yanagi & Berman, 2014; also see Petrů et al., 2008). Similarly, several primate species are thought to initiate social play by slapping, touching, or pawing at conspecifics, whereas looking between the legs or arm extensions may be used to maintain social
Compositionality in Primate Gestural Communication and…

play (Tomasello & Call, 1997). However, most research has focused on the meta-communicatory function of the play face (or relaxed open mouth-face)—a facial expression that primates often use in the context of play (van Hooff, 1972; Palagi, 2008; Spinka et al., 2016). Play faces, in particular, appear to clarify the playful meaning of behaviours that may otherwise appear agonistic (Bekoff & Allen, 1998; Palagi, 2008, 2009; Pellis & Pellis, 1996), likely preventing aggressive escalations among players (Waller & Dunbar, 2005). Indeed, play faces more frequently occur during rough or contact play, such as wrestling (Chevalier-Skolnikoff, 1974; Demuru et al., 2015; Fedigan, 1972; Palagi, 2007; Palagi & Paoli, 2007), and they often are associated with longer play bouts (Palagi, 2007; Spinka et al., 2016; Waller & Cheery, 2012) and a higher number of players (Palagi, 2008). Therefore, play faces appear to serve a crucial communicative role during social play in primates, because they seem to convey the sender’s intention before producing a second behaviour (such as hitting or chasing), which could potentially also be used in agonistic interactions. Although they do not necessarily change, but disambiguate a signal’s or behaviour’s meaning, such meta-communicative signals may be a crucial element in compositional combinations of multicomponent elements.

**Temporal Aspects of Compositionality**

There also is no consensus in the literature on how temporally close two elements should be, to be considered part of the same combination. Some authors consider signals to be elements of a combination when they are produced simultaneously, within or across modalities (Liebal et al., 2010; Partan, 2002). However, there is no agreement with regard to whether simultaneously produced signals have to overlap completely, or whether the second signal might follow with some delay. Similarly, for sequential combinations, intervals between two consecutive signals vary from one (Genty & Byrne, 2010) to several seconds (Liebal et al., 2004). With regard to meta-communicative signals, most authors suggest that they do not need to be produced right before the behavioural elements whose meaning they should modify, but they may be produced simultaneously or immediately afterwards (Beresin & Farley-Rambo, 2018; Fedigan, 1972; Palagi, 2008; Pellis & Pellis, 1996; Schwartzmann, 1979; Spinka et al., 2016). Therefore, play faces may repeatedly occur during the play bout to continuously assert the playful intent of behavioural patterns that happen before, during or after the play face, maintaining and prolonging social play rather than initiating it (Fagen, 1981; Palagi, 2007, 2008; Pellis & Pellis, 1996; Waller & Cheery, 2012; Wright et al., 2018; Yanagi & Berman, 2014). In vocal communication, in contrast, the elements of a combination cannot be produced simultaneously, but they must be in close temporal proximity to compositionally acquire a new meaning, and they also must be ordered in specific ways, because only certain sequences of elements trigger a novel response (Ouattara et al., 2009a; Suzuki et al., 2016).

For gestural sequences, researchers have usually established a priori temporal timeframes in which gestures have to occur to be considered part of the same sequence. In chimpanzees, for instance, gestures were considered to be part of the same sequence if they were directed in the same context to the same recipient, and
they were not separated by more than 5 seconds from each other (Liebal et al., 2004). In gorillas, gestures were similarly considered to belong to the same sequence if they were separated from each other by no more than 1 second (Genty & Byrne, 2010). This approach provides a clear temporal timeframe in which to analyse gestural sequences, but the decision of this timeframe may be arbitrary. Other researchers have therefore used a different approach, considering a sequence to be formed by all the gestures performed by an individual toward a recipient until the signaler stops gesturing and/or the recipient responds (Hobaiter & Byrne, 2011). However, this approach may be problematic, because it fails to operationally differentiate between gestural sequences that result from a failure to elicit a response, and those that may have been compositionally recombined to convey a novel meaning.

A better approach may therefore be to statistically identify combinations of signals by extracting clusters from gestural sequences, based on the strength of association between the single elements (Kershenbaum et al., 2016). To do that, it may be possible to focus on the signals produced by the same individual toward the same recipient, which are not mere repetitions of the same signal (Liebal et al., 2004). Then, cluster algorithms may be used to detect temporal dependencies between specific signals, selecting clusters of elements that are more likely to occur together within the communication flow. For example, cluster sequence mining may account for the order of signal occurrences and the distribution of time intervals, generating and evaluating candidate clusters with hierarchical clustering without setting a priori temporal constraints (Fukui et al., 2019). By categorizing signals into clusters, researchers might identify regularities in primate communication systems and then test whether such clusters are compositional combinations of signals that have acquired a novel meaning. As for single gestures, the meaning of signal combinations may be inferred from the response given by the recipient (Hobaiter & Byrne, 2014) and then compared with that elicited by its single elements. In this way, it may be possible to assess whether recipients provide different responses to combinations compared with their elements and thus whether combinations are produced compositionally.

Similar approaches have been recently used to investigate combinations of signals in mangabeys and chimpanzees. In mangabeys, for instance, the authors used sequence analyses to identify possible multicomponent combinations of signals based on dissimilarity measures and then used network analyses to identify the typical combinations produced by mangabeys (Aychet et al., 2021). In this way, they could identify eight major kinds of combinations, which included signals across multiple modalities. In chimpanzees, the authors used collocation analyses to calculate the probability of signal co-occurrence (i.e., whether pant hoots and food calls co-occurred with a higher frequency as compared to combinations of either signal with other calls; Leroux et al., 2021). The authors showed that pant hoot and food calls, which are meaningful elements in the vocal communication systems of chimpanzees, were more likely to be combined together. These combinations were more likely to happen in specific contexts (i.e., when higher-ranking individuals were present and the food patches were larger; Leroux et al., 2021).
**Intentional Production of Compositional Structures/Combinations**

There is currently no research on whether combinations, including gestures, are planned, voluntarily produced means of communication, in which signallers intentionally combine elements to convey new meanings. Intentionality is a key feature of human language, and intentional production is an inherent part of the gesture definition in comparative communication research, although this aspect has been considered to a much lesser extent in facial and vocal research (Liebal et al., 2013). Therefore, gestures are usually coded by default as intentionally produced means of communication that are flexibly adjusted to the recipient’s behaviour, even when they may not always be produced intentionally, whereas vocalizations and facial expressions are often presented as merely involuntary, spontaneous expressions of emotional states (Chevalier-Skolnikoff, 1974; Scopa & Palagi, 2016; Tomasello, 2008). Therefore, different approaches in which all signals can include an emotional and an intentional component, independently of the signal modality, may be important (Graham et al., 2019; Heesen et al., 2021).

Importantly, although researchers have proposed different criteria to identify intentional communication, there is yet no agreement on which and how many of these criteria need to be met to consider a signal as being intentional, nor is there consistency on how these criteria are applied across modalities. These criteria, for instance, include the social use of signals (i.e., signals are always produced in the presence of a recipient), sensitivity to recipients’ attentional state (in the case of visual signals), as well as persistence and/or elaboration (i.e., signals are repeated until the recipient produces a response, and might be even elaborated to elicit the recipient’s response in case of initial failure). However, these criteria have not yet been used to study compositionality, and to our knowledge intentionality is not usually considered a necessary prerequisite for compositionality, neither in human literature nor in literature on other species (Arnold & Zuberbühler, 2006, 2008; Engesser et al., 2016; Ouattara et al., 2009a, 2009b; Suzuki et al., 2016). Therefore, we still do not know whether primates combine different signals across modalities to create novel meanings, and whether these combinations are characterized by a set of markers indicating their intentional use.

**Future Directions**

Very few studies have addressed the compositional aspects of communication in species other than humans. These few studies have mostly focused on vocal communication, so we know very little about compositionality in primate gestural communication systems. Future research should consider several aspects to systematically address this issue. First, there is no consensus on how compositionality should be operationalized in nonhuman species, and there are often important differences with the approach used in humans. Terms, for instance, are not consistently applied across studies, and they may have a very different meaning from the ones traditionally used in human research. Addressing these
issues will be necessary to set the basis for more systematic comparative work on compositionality.

Second, the study of compositionality and meta-communication in primate gestural communication systems may benefit from the use of other methodological approaches. Candidate combinations of gestures, for instance, should be ideally identified with statistical approaches that allow a more objective clustering of signals. Because primate communication is multicomponent and multimodal (Slocombe et al., 2011), studies of compositionality should assess combinations including several gestures, gestures with facial expressions, and gestures with vocalizations. Moreover, the meaning of single gestures, facial expressions, vocalizations, and their combinations could be inferred from the response they elicit, rather than from the context in which they are produced (Hobaiter & Byrne, 2014). Combinations also should be studied across different contexts to assess specifically whether context modulates recipients’ response to gestural combinations (Oña et al., 2019). Furthermore, it would be interesting to study the specific characteristics of the combinations produced by primates, especially focusing on the presence of markers that suggest their intentional use, on the temporal distance to which the single elements are produced, and on the order in which the single elements are produced, as different orders of elements may clearly affect the meaning of combinations (Townsend et al., 2018). This implies the availability of very large datasets with a variety of primate signals and multicomponent combinations in many contexts and calls for international long-term collaborations with standardized, methodological approaches.

Finally, another aspect to consider is how compositionality in gestural communication might emerge through development. By observing primates at different ages (i.e., from infancy to adulthood) and by longitudinally following some individuals, it would be possible to analyse how compositionality and its different properties emerge through development and whether they follow the same developmental trajectories as in humans. Moreover, it would be interesting to explore the social factors that likely facilitate the emergence of compositionality in primate communication systems. For instance, it is possible that gestures (and their combinations), being more frequently used in affiliative social contexts (Pika et al., 2005), are more likely produced by individuals and groups that have better access to resources and need to invest less time in non-affiliative activities (e.g., higher-ranking individuals, captive groups). However, it also is possible that gestures (and their combinations) are more common across individuals living in socially more complex groups (e.g., in the wild), because they may need to use a wider variety of communicatory signals to fulfil their social needs (Freeberg et al., 2012). Future systematic studies on the gestural communication systems of primates may not only provide the first clear evidence of compositionality in their gestural communication but also shed light into the evolutionary origins of one of the most crucial properties of human communication.
Acknowledgments The authors thank the editors and reviewers for the extensive and constructive feedback throughout the review process and the organizers of this special issue for their invitation to contribute. They are grateful to Blanca Striegler for help with the bibliographic search and also acknowledge support from the Max Planck Society.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Note As this study is a review of literature, no ethical permission was required.

Data Availability As this study is a review of literature, no data were used and no data are available.

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

References

Allen, J. A., Garland, E. C., Dunlop, R. A., & Noad, M. J. (2019). Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song. Proceedings of the Royal Society B: Biological Sciences, 286, 20192014.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. Behavioral and Brain Sciences, 28, 105–167.

Arbib, M. A. (2017). Toward the language-ready brain: Biological evolution and primate comparisons. Psychonomic Bulletin & Review, 24, 142–150.

Arnold, K., & Zuberbühler, K. (2006). Semantic combinations in primate calls. Nature, 441, 303.

Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. Current Biology, 18, 202–203.

Aychet, J., Blois-Heulin, C., & Lemasson, A. (2021). Sequential and network analyses to describe multiple signal use in captive mangabeys. Animal Behaviour, 21, 00038R.

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, 535–550.

Bateson, G. (1955). A theory of play and fantasy. In I. G. Bateson (Ed.), Steps to an ecology of mind (pp. 177–193). Ballantine Books.

Bateson, G. (1956). The message “This is play.” In B Schaffner (Ed.), Group processes: Transactions of the second conference (pp. 145-242). Josiah Macy Jr. Foundation.

Bekoff, M. (1972). The development of social interaction, play, and metametacommunication in mammals: An ethological perspective. The Quarterly Review of Biology, 47, 412–434.

Bekoff, M. (1995). Play signals as punctuation: The structure of social play in canids. Behaviour, 132, 419–429.

Bekoff, M., & Allen, C. (1998). Intentional communication and social play: How and why animals negotiate and agree to play. In M. Bekoff & J. A. Byers (Eds.), Animal play: Evolutionary, comparative, and ecological perspectives (pp. 97–114). Cambridge University Press.
F. Amici et al.

Beresin, A. R., & Farley-Rambo, K. (2018). Play signals, play moves: A gorilla critique of play theory. International Journal of Play, 7, 322–337.

de Boer, B., Sandler, W., & Kirby, S. (2012). New perspectives on duality of patterning: Introduction to the special issue. Language and Cognition, 4, 251–259.

Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and stereotypy of free-tailed bat songs. PLoS One, 4, e6746.

Burger, G. M. (2005). The genesis of animal play: Testing the limits. The MIT Press.

Cavicchio, F., Dachkovsky, S., Leemor, L., Shamay-Tsoory, S., & Sandler, W. (2018). Compositionality in the language of emotion. PLoS One, 13, e0201970.

Chevalier-Skolnikoff, S. (1974). The primate play face: A possible key to the determinants and evolution of play. Rice University Studies, 60, 9–29.

Chomsky, N. (1966). Cartesian linguistics: A chapter in the history of rationalist thought. Cambridge University Press.

Chomsky, N. (1981). Knowledge of language: Its elements and origins. Philosophical Transactions of the Royal Society B, 295, 223–234.

Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. PLoS One, 1, e73.

Cody, M. L., Stabler, E., Sánchez Castellanos, H. M., & Taylor, C. E. (2016). Structure, syntax and ‘small-world’ organization in the complex songs of California thrashers (Toxostoma redivivum). Bioacoustics, 25, 41–54.

Demartsev, V., Gordon, N., Barocas, A., Bar-Ziv, E., Ilany, T., Goll, Y., Ilany, A., & Geffen, E. (2019). The ‘law of brevity’ in animal communication: Sex-specific signaling optimization is determined by call amplitude rather than duration. Evolution Letters, 3, 623–634.

Demuru, E., Ferrari, P. F., & Palagi, E. (2015). Emotionality and intentionality in bonobo playful communication. Animal Cognition, 18, 333–344.

Deslandes, V., Faria, L. R., Borges, M. E., & Pie, M. R. (2014). The structure of an avian syllable syntax network. Behavioural Processes, 106, 53–59.

Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. Proceedings of the National Academy of Sciences USA, 113, 5976–5981.

Estienne, V., Stephens, C., & Boesch, C. (2017). Extraction of honey from underground bee nests by central African chimpanzees (pan troglodytes troglodytes) in Loango National Park, Gabon: Techniques and individual differences. American Journal of Primatology, 79.

Fagen, R. (1981). Animal play behavior. Oxford University Press.

Favaro, L., Gamba, M., Cresta, E., Fumagalli, E., Bandoli, F., Pilenga, C., Isaja, V., Mathevon, N., & Reby, D. (2020). Do penguins’ vocal sequences conform to linguistic laws? Biology Letters, 16, 20190589.

Fedigan, L. (1972). Social and solitary play in a colony of vervet monkeys (Cercopithecus aethiops). Primates, 13, 347.

Fishbein, A. R., Fritz, J. B., Idsardi, W. J., & Wilkinson, G. S. (2019). What can animal communication teach us about human language? Philosophical Transactions of the Royal Society B, 375, 20190042.

Fitch, W. T. (2010). The evolution of language. Cambridge University Press.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. Science Advances, 2, e1600723.

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. Philosophical Transactions of the Royal Society B, 367, 1785–1801.

Fukui, K., Okada, Y., Satoh, K., & Numao, M. (2019). Cluster sequence mining from event sequence data and its application to damage correlation analysis. Knowledge-Based Systems, 179, 136–144.

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (Gorilla gorilla): Repertoire, intentionality and possible origins. Animal Cognition, 12, 527–546.

Genty, E., & Byrne, R. W. (2010). Why do gorillas make sequences of gestures? Animal Cognition, 13, 287–301.

Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M., & Crockford, C. (2021). Chimpanzees use numerous flexible vocal sequences with more than two vocal units: A step towards language? BioRxiv, 2021(02), 03.429517.

Graham, K. E., Furuichi, T., & Byrne, R. W. (2020). Context, not sequence order, affects the meaning of bonobo (pan paniscus) gestures. Gesture, 19, 335–364.

Graham, K. E., Wilke, C., Lahiff, N. J., & Slocombe, K. E. (2019). Scratching beneath the surface: Intentionality in great ape signal production. Philosophical Transactions of the Royal Society B, 375, 20180403.
Compositionality in Primate Gestural Communication and…

Gustison, M. L., Semple, S., Ferrer-I-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath’s linguistic law. *Proceedings of the National Academy of Sciences, 113*, 2750–2758.

Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015). Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition, 18*, 423–435.

Heesen, R., Sievers, C., Gruber, T., & Clay, Z. (2021). Primate communication: Affective, intentional, or both? In B. L. Schwartz & M. J. Beran (Eds.), *Primate cognitive studies*. Cambridge University Press.

Huang, M., Ma, H., Ma, C., Garber, P. A., & Fan, P. (2020). Male gibbon loud morning calls conform to Zipf’s law of brevity and Menzerath’s law: Insights into the origin of human language. *Animal Behaviour, 160*, 145–155.

Hobaiter, C., & Byrne, R. W. (2011). Serial gesturing by wild chimpanzees: Its nature and function for communication. *Animal Cognition, 14*, 827–838.

Hobaiter, C., & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology, 24*, 1596–1600.

Liebal, K., Waller, B. M., Burrows, A., & Slocombe, K. (2013). Primate communication: A multimodal approach. Cambridge University Press.

Luo, B., Jiang, T., Liu, Y., Wang, J., Lin, A., Wei, X., & Feng, J. (2013). Brevity is prevalent in bat short-range communication. *Journal of Comparative Physiology A, 199*, 325–333.

Menzerath, P. (1954). *Die architektonik des deutschen wortschatzes*. F. Dümmel.

Mitchell, R. W. (1991). Bateson concept of metacommunication in play. *New Ideas in Psychology, 9*, 73–87.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell’s monkeys use affixation to alter call meaning. *PLoS One, 4*, e7808.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell’s monkeys concatenate vocalizations into context-specific sequences. *Proceedings of the National Academy of Sciences USA, 106*, 22026–22031.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell’s monkeys use affixation to alter call meaning. *PLoS One, 4*, e7808.

Palagi, E. (2007). Play at work: Revisiting data focusing on chimpanzees (pan troglodytes). *Journal of Anthropological Sciences, 85*, 63–81.

Palagi, E. (2008). Sharing the motivation to play: The use of signals in adult bonobos. *Animal Behaviour, 75*, 887–896.
Palagi, E. (2009). Adult play fighting and potential role of tail signals in ring-tailed lemurs (Lemur catta). *Journal of Comparative Psychology, 123*, 1–9.

Palagi, E., & Paoli, T. (2007). Play in adult bonobos (Pan paniscus): Modality and potential meaning. *American Journal of Physical Anthropology, 134*, 219–225.

Parr, L. A., Cohen, M., & de Waal, F. (2005). Influence of social context on the use of blended and graded facial displays in chimpanzees. *International Journal of Primatology, 26*, 73–103.

Partan, S. R. (2002). Single and multichannel signal composition: Facial expressions and vocalizations of rhesus macaques (Macaca mulatta). *Behaviour, 139*, 993–1027.

Pellis, S. M., & Pellis, V. C. (1996). On knowing it’s only play: The role of play signals in play fighting. *Aggression and Violent Behavior, 1*, 249–268.

Pellis, S. M., & Pellis, V. C. (2009). The playful brain: Venturing to the limits of neuroscience. *Oneworld.*

Petru, M., Špinka, M., Lhota, S., & Šípek, P. (2008). Head rotations in the play of Hanuman langurs (Semnopithecus entellus): Description and analysis of function. *Journal of Comparative Psychology, 122*, 9–18.

Pika, S., Liebal, K., Call, J., & Tomasello, M. (2005). The gestural communication of apes. *Gesture, 5*, 39–54.

Pinkser, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

Riede, T., Bronson, E., Hatzikirou, B., & Zuberbühler, K. (2005). The production mechanisms of Diana monkey alarm calls: Morphological data and a model. *Journal of Human Evolution, 48*, 85–96.

Roberts, A. I., Roberts, S. G. B., & Vick, S. I. (2014). The repertoire and intentionality of gestural communication in wild chimpanzees. *Animal Cognition, 17*, 317–336.

Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour, 58*, 921–931.

Sade, D. S. (1973). An ethogram of rhesus monkeys: I. antithetical contrasts in posture and movement. *American Journal of Physical Anthropology, 38*, 537–542.

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press.

Sasahara, K., Cody, M. L., Cohen, D., & Taylor, C. E. (2012). Structural design principles of complex bird songs: A network-based approach. *PLoS One, 7*, e44436.

Schwartzmann, H. B. (1979). *Transformations*. The anthropology of children’s play: Springer.

Scopa, C., & Palagi, E. (2016). Mimic me while playing! Social tolerance and rapid facial mimicry in macaques (Macaca tonkeana and Macaca fuscata). *Journal of Comparative Psychology, 130*, 153–161.

Scott-Phillips, T. (2015). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.

Semple, S., Hsu, M. J., & Agoramooorthy, G. (2010). Efficiency of coding in macaque vocal communication. *Biology Letters, 6*, 469–471.

Seyfarth, R. M., Cheney, D. L., & Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends in Cognitive Sciences, 9*, 264–266.

Shapiro, A. D., Tyack, P. L., & Seneff, S. (2010). Comparing call-based versus subunit-based methods for categorizing Norwegian killer whale, Orcinus Orca, vocalizations. *Animal Behaviour, 81*, 377–386.

Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: The need for multimodality in primate communication research. *Animal Behaviour, 81*, 919–924.

Spinka, M., Palckova, M., & Rehakova, M. (2016). Metacommunication in social play: The meaning of aggression-like elements is modified by play face in Hanuman langurs (Semnopithecus entellus). *Behaviour, 153*, 795–818.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications, 7*, 10986.

Tanner, J. E., & Perlman, M. (2017). Moving beyond ‘meaning’: Gorillas combine gestures into sequences for creative display. *Language & Communication, 54*, 56–72.

Tempelmann, S., & Liebal, K. (2012). Spontaneous use of gesture sequences in orangutans: A case for strategy? *Gesture Studies, 6*, 73–92.

Tomasello, M. (2008). *Origins of human communication*. MIT Press.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford University Press.

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

Waller, B. M., & Cheery, L. (2012). Facilitating play through communication: Significance of teeth exposure in the gorilla play face. *American Journal of Primatology, 74*, 157–164.
Waller, B. M., & Dunbar, R. I. M. (2005). Differential behavioural effects of silent bared teeth display and relaxed open mouth display in chimpanzees (*Pan troglodytes*). *Ethology, 111*, 129–142.

Waller, B. M., Liebal, K., Burrows, A. M., & Slocombe, K. E. (2013). How can a multicomponent approach to primate communication help us understand the evolution of communication? *Evolutionary Psychology, 11*, 147470491301100.

Waller, B. M., Kavanagh, E., Micheletta, J., Clark, P. R., & Whitehouse, J. (2022). The face is central to primate multicomponent signals. *International Journal of Primatology.*

Waller, B. M., Whitehouse, J., & Micheletta, J. (2016). Rethinking primate facial expression: A predictive framework. *Neuroscience & Biobehavioral Reviews, 82*, 13–21.

Watson, S. K., Burkart, J. M., Schapiro, S. J., Lambeth, S. P., Mueller, J. L., & Townsend, S. W. (2020). Nonadjacent dependency processing in monkeys, apes, and humans. *Science. Advances, 6*, eabb0725.

Weiss, M., Hultsch, H., Adam, I., Scharff, C., & Kipper, S. (2014). The use of network analysis to study complex animal communication systems: A study on nightingale song. *Proceedings of the Royal Society B: Biological Sciences, 281*, 20140460.

Werning, M., Hinzen, W., & Machery, E. (2012). *The Oxford handbook of compositionality*. Oxford University Press.

Wilke, C., Kavanagh, E., Donnellan, E., Waller, B. M., Machanda, Z. P., & Slocombe, K. E. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, *Pan troglodytes schweinfurthii*. *Animal Behaviour, 123*, 305–316.

Wright, K. R., Mayhew, J. A., Sheeran, L. K., Funkhouser, J. A., Wagner, R. S., Sun, L. X., & Li, J. H. (2018). Playing it cool: Characterizing social play, bout termination, and candidate play signals of juvenile and infant Tibetan macaques (*Macaca thibetana*). *Zoological Research, 18*, 272–283.

Yanagi, A., & Berman, C. M. (2014). Body signals during social play in free-ranging rhesus macaques (*Macaca mulatta*): A systematic analysis. *American Journal of Primatology, 76*, 168–179.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Hafner Publishing Company.

Zuberbühler, K. (2005). The phylogenetic roots of language: Evidence from primate communication and cognition. *Current Directions in Psychological Science, 14*, 126–130.

---

**Authors and Affiliations**

**Federica Amici** 1,2, † · **Linda Oña** 3 · **Katja Liebal** 1

1 Life Sciences, Institute for Biology, Human Biology and Primate Cognition, Leipzig University, Leipzig, Germany

2 Max Planck Institute for Evolutionary Anthropology, Department of Human Behavior, Ecology and Culture, Leipzig, Germany

3 Max Planck Institute for Human Development, Max Planck Research Group ‘Naturalistic Social Cognition’, Berlin, Germany