



OPEN

# Breaking conversational rules matters to captive gorillas: A playback experiment

Loïc Pougault<sup>1,2,3</sup>✉, Florence Levréro<sup>2,4</sup>, Baptiste Mulo<sup>3</sup> & Alban Lemasson<sup>1,4</sup>

Across human cultures, conversations are regulated by temporal and social rules. The universality of conversational rules suggests possible biological bases and encourages comparisons with the communicative interactions of nonhuman animals. Unexpectedly, few studies have focused on other great apes despite evidence of proto-conversational rules in monkeys, thus preventing researchers from drawing conclusions on potential evolutionary origins of this behaviour. A previous study showed however that western lowland gorillas engage in soft call interactions that seem temporally- and socially-ruled. Indeed, interactions occurred mainly between individuals close in age who followed a preset response delay, thus preventing call overlap. Here, we experimentally investigated the presence of these rules in a captive gorilla group, using a violation-of-expectation paradigm. Head orientation responses suggest that the respect of response delay matters to subjects, but the importance of the interlocutors' age proximity appeared less clear. The intensity of the response varied with subjects' age in a context-dependent way, supporting a possible role of learning. Our findings support the growing number of studies highlighting the importance of vocal turn-taking in animals and a possible sociogenesis of this ability. The capacity to "converse" might have been a key step in the co-evolution of communication and complex sociality.

Despite the diversity of human cultures, some basic conversational rules are respected in all societies and thus appear universal<sup>1</sup>. These features are gathered in a so-called "contract of communication" that conversing interlocutors informally agree on<sup>2</sup>. The contract takes into account "context relevance" (*i.e.* the evaluation of the context as pertinent or not to initiate a conversation), "reciprocity" (*i.e.* the evaluation of both partners as valid interlocutors), and "contract-based temporal rules" (*i.e.* the respect of a reciprocal exchange of alternating, short, and flexible turns between two or more interlocutors<sup>3</sup>, and speech overlap avoidance). Several authors have questioned the possible biological bases of these conversational rules and have, for several decades, conducted cross-species comparisons to understand their origin and role across animal communication systems<sup>4-7</sup>.

To facilitate cross species investigations, authors typically distinguish "conversation-like vocal exchanges" from other vocal patterns like isolated calling (one call emitted and no other calls can be heard around, *e.g.* in red-capped mangabeys *Cercocebus torquatus*<sup>8</sup>), repeated calling (the same caller calls several times in a row, *e.g.* in Japanese macaques *Macaca fuscata*<sup>9</sup>), disorganized phonoresponses (one individual produces a call, typically an alarm call that triggers calls in an apparent chaotic way from the other group members, *e.g.* in blue monkeys *Cercopithecus mitis stuhlmanni*<sup>10</sup>), chorusing (two or more individuals overlap their emission of a given call type, *e.g.* in chimpanzees *Pan troglodytes*<sup>11</sup>) and duetting (two individuals synchronise long series of calls or songs with stereotyped temporal association: *e.g.* in birds<sup>12-14</sup>, and gibbons *Hylobates syndactylus*<sup>15</sup>). "Conversation-like vocal exchanges" are distinguished from the other types of vocal interactions because of their following specific features. First, they involve a diversity of recurrent vocal partners (which differs from duets). Second, interlocutors are typically familiar individuals belonging to a given social group and can be of any age or either sex (which differs from synchronized signalling<sup>16-20</sup>). Third, "conversation-like vocal exchanges" are not restricted to a specific context, which differs from – usually long-distance – collective communication associated with mate attraction, territory protection and environmental disturbance<sup>21</sup>, and time of the day or year (excluding morning choruses and reproductive signals<sup>22,23</sup>). Most commonly, these "conversation-like vocal exchanges" are observed

<sup>1</sup>Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000, Rennes, France.

<sup>2</sup>Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES/CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France. <sup>3</sup>ZooParc de Beauval & Beauval Nature, 41110, Saint Aignan, France. <sup>4</sup>These authors contributed equally: Florence Levréro and Alban Lemasson. ✉e-mail: [loic.pougault@univ-rennes1.fr](mailto:loic.pougault@univ-rennes1.fr)

in everyday and basically all day long short-distance communication (*i.e.* soft calls exchanged in peaceful interactions<sup>5</sup>), as in humans<sup>3,24</sup>.

The “conversation-like vocal exchanges” are temporally organized on the basis of turn-taking which refers to “the orderly exchange of purely communicative signals or behaviours between individuals characterized by principles for the coordination turn transfer, which result in observable temporal regularities” (cited from Pika and colleagues<sup>25</sup>). Communicative turn-taking has been found in a broad range of nonhuman species (*e.g.* elephants *Loxodonta africana*<sup>26</sup>; sperm-whales *Physeter macrocephalus*<sup>27</sup>; common bottlenose dolphins *Tursiops truncatus*<sup>28</sup>; starlings *Sturnus vulgaris*<sup>4</sup>; naked mole-rats *Heterocephalus glaber*<sup>29</sup>; bats *Diaemus youngi*<sup>30</sup>), including a broad range of nonhuman primates (NHP) species<sup>25,31,32</sup>. Thus, to respect turn-taking, interlocutors have to obey basic rules<sup>5</sup>. Interlocutors have coordinated roles, one of them sends a first signal and the other responds, and, when the conversation keeps going, the two interlocutors alternate. The respondent typically adjusts the timing of its vocal response by respecting a minimum silence gap (to prevent call overlap) and a maximum silence gap (to ensure call coordination). A recent study on marmosets, focusing on long periods of vocal interactions, showed such temporal coordination of vocal productions between individuals<sup>32</sup>.

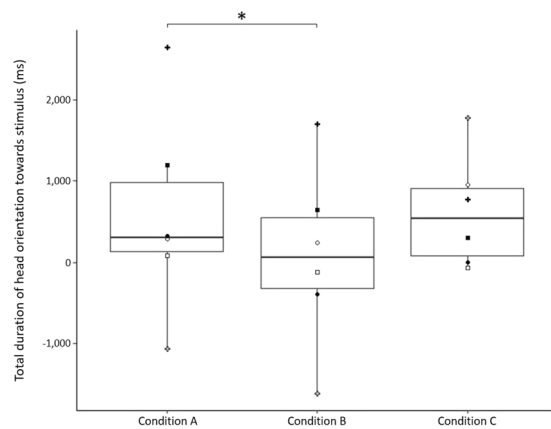
Another important rule shared by human, nonhuman primates and several other animal species is that exchanging partners are not randomly nor opportunistically chosen (bottlenose dolphins<sup>28</sup>, large-billed crows *Corvus macrorhynchos*<sup>33</sup>, meerkats *Suricata suricatta*<sup>34</sup>, sperm whales<sup>27</sup>, elephants<sup>35</sup>). Dunbar<sup>36</sup> suggested that monkey call exchanges had started to function as a mean to groom-at-distance when the size of group became too large to allocate enough time to physical grooming among all members and maintain close bonds. Indeed, “conversation-like vocal exchanges” occur mainly between two interlocutors that are determined, according to the study species, by their affiliative bonds (humans<sup>37</sup>; NHP: spider monkeys *Ateles geoffroyi*<sup>38</sup>, Japanese macaques<sup>39</sup>, pygmy marmosets *Cebuella pygmaea*<sup>6</sup>, bonobos *Pan paniscus*<sup>40</sup>) or by their respective ages (humans<sup>41</sup>; NHP: common marmosets<sup>42</sup>, western lowland gorillas *Gorilla gorilla gorilla*<sup>43</sup>, Campbell’s monkeys<sup>31</sup>). In some cases, it was even found that human and nonhuman primates advertise this preferential bond by responding to one another using a matching acoustic structure, a phenomenon known as “vocal convergence” in ethology or “vocal accommodation” in sociolinguistics (humans<sup>44</sup>; NHP: spider monkeys<sup>38</sup>, Diana monkeys *Cercopithecus diana*<sup>45</sup>).

It is worth noting that some conversational characteristics have also been found in nonprimate species. However, only the vocal interactions of nonhuman primates match all the criteria defining a “conversation-like vocal exchange” (see above the mentioned features). The origins of such vocal behaviours in nonhuman primates remain debated, notably because most published research on the topic has focused on monkeys, and additionally available findings in nonhuman apes (phylogenetically closer to humans) are controversial. The current state of knowledge prevents to conclude between a convergent evolution guided by the requirements of social life or a shared inheritance, suggesting an ancient mechanism which was already present in the primate lineage.

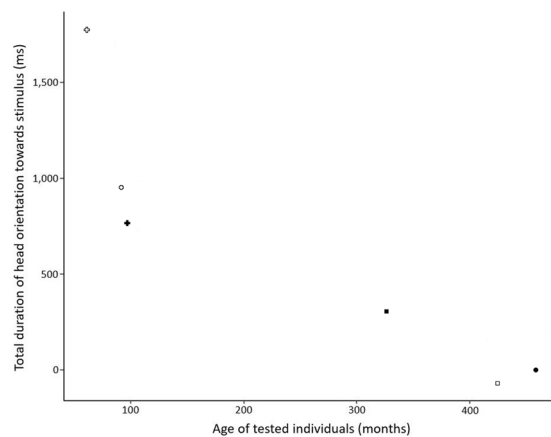
To our knowledge, three studies have focused on vocal conversational rules in nonhuman great apes, two of which found support for turn-taking and call overlap avoidance (bonobos<sup>40</sup>; gorillas<sup>43</sup>), and one that did not (chimpanzees<sup>46</sup>). Authors also questioned the fact that these abilities to converse in animals are genetically programmed or, as in humans<sup>47,48</sup>, socially acquired. Empirical evidence is currently lacking to answer this question, although the few available studies on monkeys showed that juveniles break conversational rules more often than do adults (howler monkeys *Alouatta pigra*<sup>21</sup>; Campbell’s monkeys<sup>49</sup>). Moreover, playback experiments using a violation of expectation paradigm showed that adult monkeys clearly discriminated between conversationally appropriate and inappropriate vocal exchange patterns, whereas socially inexperienced juveniles did not (Japanese macaques<sup>50</sup>; Campbell’s monkeys<sup>49</sup>). These experiments show that the audience detects the violation of expected social rules in vocal interactions between third parties<sup>49–52</sup>, suggesting that the social rules in vocal interactions are not a simple neurobiological guided behaviour but a social awareness of the respect of social rules.

Here, we aimed to progress this debate by conducting a playback experiment with captive western lowland gorillas of different ages. A recent study, ran on a captive group of western lowland gorillas housed at the ZooParc de Beauval, found that individuals do seem to engage in socially- and temporally-ruled vocal interactions<sup>43</sup>. Vocal interactions typically involve grunts, defined as soft contact calls<sup>53–55</sup> that are acoustically individually distinctive<sup>55</sup>. Grunts seem to play a key role in the coordination and peaceful activities, and thus act as a contact call<sup>53–55</sup>. During grunt exchanges, temporal rules are respected with an average “response” delay of 0.5 seconds (maximum 3 s) and with obvious call overlap avoidance<sup>43</sup>. Also, preferred interlocutors are non-randomly chosen. Grunt exchanges are mainly dyadic interactions that are most frequently observed among partners close in age, regardless of kinship<sup>43</sup>.

To assess the social relevance of these two potential conversational rules (inter-call durations and the age difference of interlocutors) for the audience in the captive population housed at the ZooParc de Beauval ( $n = 6$ ), we compared the responses of gorillas to grunt exchanges that have been built to match both the temporal and the social rules (condition A), in comparison with built grunt exchanges which respectively violate the overlap avoidance expectation (condition B) and the expectation of an age proximity between vocally exchanging partners (condition C). We also tested whether the age of the subject (perceiver) influenced its responses to these different vocal exchanges. We expected a modification of the attentional state of the audience by comparing the difference of the total duration of head orientation and the total number of occurrences of locomotion in the direction of the loudspeaker in the 60 seconds after the onset of the second vocalisation and 60 seconds before playback. We also examined the first latency to reposition the head front, corresponding to the duration of the first gaze towards the loudspeaker after the diffusion. The change of attentional state may be in both ways: an increased response towards the non-congruent pattern (implying that animals are surprised to hear the violation of the overlap avoidance and the expectation of an age proximity between interlocutors, as in Japanese macaques<sup>50</sup>), or towards the natural ones that they are exposed to in their everyday social life (implying that the audience is strongly interested by a relevant interaction including their congeners, as in Campbell’s monkeys<sup>49</sup>).



**Figure 1.** Total duration of head orientation toward the loudspeaker (in the 60 seconds ‘after’ minus 60 seconds ‘before’ playback) in the three playback conditions: “likely vocal exchange” (condition A), “overlapped vocal exchange” (condition B), “age-difference vocal exchange” (condition C). Each symbol corresponds to an individual. Box and whisker plots report the median, 25th and 75th percentiles, and the lowest and highest data, which are no more than 1.5 interquartile range from the box. \* $P < 0.05$ .

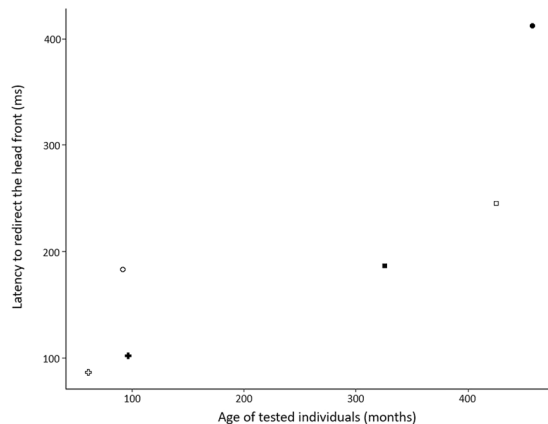


**Figure 2.** Total duration of head orientation in the “age-difference vocal exchange” (condition C) (in the 60 seconds ‘after’ minus 60 seconds ‘before’ playback) in relation to the subject’s age. Each symbol corresponds to an individual (same as in Fig. 1).

## Results

Analyses showed that individuals oriented their head towards the stimulus for a longer total time when temporal and social rules were respected (condition A) than when temporal rule was not (*i.e.* when calls were overlapping, condition B) (Wilcoxon signed rank test,  $V = 21$ ,  $P = 0.031$ , Fig. 1). However, there was no significant difference in the duration of head orientations between the likely (condition A) and the age-difference (condition C) vocal exchange conditions ( $V = 12$ ,  $P = 0.844$ ), nor between the overlapped calls and the age-difference vocal exchange conditions ( $V = 8$ ,  $P = 0.6875$ ). Moreover, no effect of condition was found for latency to redirect the head after the first gaze in direction to the loudspeaker (condition A: mean  $\pm$  SD =  $2025 \pm 1181$ msec; condition B:  $2005 \pm 72$ msec; condition C:  $4995 \pm 7256$ msec;  $V_{A/B} = 10$ ,  $P = 1$ ;  $V_{A/C} = 7$ ,  $P = 0.563$ ) nor for locomotion (condition A:  $0.22 \pm 0.29$ ; condition B:  $0.13 \pm 0.31$ ; condition C:  $0.071 \pm 0.21$ ;  $V_{A/B} = 5$ ,  $P = 0.423$ ;  $V_{A/C} = 6$ ,  $P = 0.174$ ).

The behavioural responses of the subject differed in some conditions according to its age. In the “age-difference vocal exchange” (condition C), we found that the older the subject, the shorter the head orientation duration towards the speaker (Spearman test,  $r_s = -0.94$ ,  $P = 0.017$ , Fig. 2). The total head orientation duration did not vary with subject’s age in the two other conditions ( $r_{sA} = 0.31$ ,  $P = 0.564$ ;  $r_{sB} = 0.03$ ,  $P = 1$ ). Likewise, when focusing on the latency to redirect the head after the first gaze directed to the speaker, we found that the older the subject, the longer the first head orientation in the “likely vocal exchange” (condition A) ( $r_s = 0.94$ ,  $P = 0.017$ , Fig. 3), but not in the two other conditions ( $r_{sB} = 0.49$ ,  $r_{sC} = -0.49$ ,  $P = 0.356$  in both cases). However, the age of the subject did not predict locomotion ( $r_{sA} = -0.76$ ,  $P = 0.08$ ,  $r_{sB} = -0.13$ ,  $P = 0.805$ ,  $r_{sC} = 0.34$ ,  $P = 0.512$ ).



**Figure 3.** Latency to redirect the head front in the “likely vocal exchange” (condition A) in relation to the subject’s age. Each symbol corresponds to an individual (same as in Fig. 1).

## Discussion

The violation-of-expectation paradigm is typically used in research on human and nonhuman animals to evaluate their ability to perceive certain rules<sup>56</sup>. Duration of post-playback head orientation responses is a common measure to assess a subject’s hearing general interest in a presented stimulus among both nonhuman<sup>57,58</sup> and human<sup>59</sup>, notably newborn<sup>60</sup> primates. In our study, results from the total duration of head orientation suggest that the gorilla’s audience pays attention to the vocal interactions among its group members and is more attentive to a “likely-vocal exchange” (condition A) than to an “overlapped vocal exchange” (condition B). In such “violation of attempt” setup, it is often predicted that unexpected stimuli would cause stronger responses than expected ones<sup>61</sup>. Indeed, nonhuman primates have shown more head orientations in response to a violation of the original pattern in several playback experiments<sup>62,63</sup>. Although our results may appear surprising in light of such studies, numerous authors have also reported an opposite effect of attempt violation on subjects’ responses. For instance, human infants tested in cross-modal studies show a familiarity/consistency preference<sup>64,65</sup>. The same is true for nonhuman primates. When exposing Japanese macaques to videos of conspecific facial expressions during vocal emissions, alongside the playback of a vocalisation which only matches one of the videos, subjects preferentially look at the congruent one (*i.e.* when vocalisation and video matched)<sup>66</sup>. A stronger attentional state towards an expected vocal sequence pattern rather than an unexpected one was also found in common marmosets<sup>67</sup>. Thus, the direction of the subjects’ attention bias in such experimental paradigms may be stimulus- or context-dependent. Interestingly, Lemasson and colleagues, who aimed answering the same research question as ours, obtained the same pattern in the Campbell monkeys they tested<sup>49</sup>.

In our study, the fact that the subjects paid a differential attention to the different test conditions (namely a strongest attention towards the “likely-vocal exchange” than the “overlapped vocal exchange”) suggests that the pattern was perceived and that temporal rules could be relevant to them. In the “overlapped vocal exchanges”, gorillas may not be able to extract all the relevant information, and notably not easily identify the callers, which can explain the weak responses of the audience in this condition. This can also be due to the fact that individuals evaluated the opportunity to participate in the ongoing simulated vocal exchange and lost interest quickly when they realised that the exchange was inappropriate, as suggested by Lemasson *et al.*<sup>49</sup> after similar results obtained in adult Campbell’s monkeys. We cannot totally rule out other possible alternative interpretations (although not mutually exclusive), notably that the lack of attention towards the “overlapped vocal exchange” (condition B) could be due to the fact that this stimulus is shorter in duration (so with a smaller quantity of “acoustic” information). However, the short duration hypothesis is unlikely for two reasons. First, because the difference in mean duration of the stimuli respecting temporal rules (*i.e.* conditions A and C) and those not respecting them (*i.e.* condition B) is very small (see methods). Second, because we found no significant difference between conditions B and C, although their stimuli differ in duration. This also reinforces our idea that the possible message confusion associated with call overlapping is not the core origin of the attentional bias observed above. The avoidance of call overlap could be a relevant social rule in the interactions of western gorillas, but did not elicit an attentional response strong enough to induce locomotion toward the loudspeaker. This is also true for the other conditions and not surprising given the fact that the simulated vocal exchanges involved soft contact calls emitted during peaceful activities<sup>53–55</sup>. Anyhow, such importance of the temporal organization of vocal interactions has already been highlighted in several monkey species<sup>25,49</sup>. In Campbell’s monkeys and Japanese macaques respectively, individuals that violate the turn-taking rule are typically young and inexperienced juveniles<sup>49</sup> or low-ranking (socially isolated) adult males<sup>68</sup>. In most human cultures, conversational overlap denotes a conversation failure and can lead to the end of an exchange<sup>3</sup>. In fact, vocal overlapping is perceived as a serious impoliteness in many traditional human societies<sup>69</sup>. In birds as well, overlap increases with aggressiveness and may lead to the end of an exchange (for a review see<sup>4</sup>). For example, in black capped chickadees (*Poecile atricapillus*), dominant males overlap more than do subordinates<sup>70,71</sup> and this overlapping increases arousal in overlapped individuals<sup>72,73</sup>.

The importance of age proximity between interlocutors for a gorilla audience remains unclear. Indeed, the age proximity or the age distance between interlocutors did not elicit different responses at the group level. One possible explanation is that breaking the social (age difference) rule is twice more common (15%) than breaking the

temporal (overlap) rule (8%) in gorillas and thus less surprising. The systematic focus on the silverback adult male in all broadcast vocal exchanges could have however induced a bias of general interest (*i.e.* silverback male vocalizations would systematically attract the attention of his groupmates whichever the context they are produced in), given that the role of the adult male is central in the social cohesion and coordination of the group<sup>74,75</sup>. This dominant male indeed develops tight bonds with each individual and is a key interlocutor for all group members. This supposed social rule (based on observations of naturally produced vocal exchanges<sup>43</sup>) cannot be considered as mandatory at this stage given that it does not elicit an unanimous behavioural response to violation.

Interestingly, the behavioural responses of the audience varied between individuals according to their age. The total duration of head orientation shows that the stronger attentional state of young gorillas toward the age-difference vocal interaction (*i.e.* condition C) is not trivial, and could reveal their interest for an eventual interaction involving another young individual and the adult male. That could explain why we did not find similar correlations in “likely vocal exchange” and “overlapped vocal exchange” conditions (*i.e.* conditions A and B) involving only adult individuals. This result suggests that, even if not a mandatory rule, gorillas may have expectations in this regard. Moreover, when we considered the first latency to reposition the head front as an indicator of the level of interest towards the stimulation, older gorillas paid more attention to exchanges between interlocutors close in age respecting common response delay (condition A) than did younger gorillas. Older gorillas may be socially more experienced than younger ones and more reliably perceive if a vocal interaction is relevant, as the “likely vocal exchange (*i.e.* condition A), or not, as the “overlapped vocal exchange” and the “age-difference vocal exchange” (*i.e.* conditions B and C), explaining the absence of correlation in conditions B and C. Nonetheless, these findings support the prediction, earlier suggested in monkeys<sup>49</sup>, that appropriate rules of conversing may be socially learned. Differences between adult and young individuals regarding the appropriateness of the context of call emission have already been reported in various primate species (see food calling in tamarins *Saguinus oedipus*<sup>76</sup>, and alarm calling in vervet monkeys *Chlorocebus pygerythrus*<sup>77</sup>). Some studies have also showed vocal development in the ability to interact properly with congeners (see vocal synchronisation in gibbons *Hylobates agilis agilis*<sup>78</sup>, as well as call matching and choice of preferred interlocutors in spider-monkeys<sup>38</sup>). Moreover, two playback studies have confirmed that conversational rules can be detected by adults and not by juveniles (turn-taking in Campbell’s monkeys<sup>49</sup>, and call matching in Japanese macaques<sup>50</sup>).

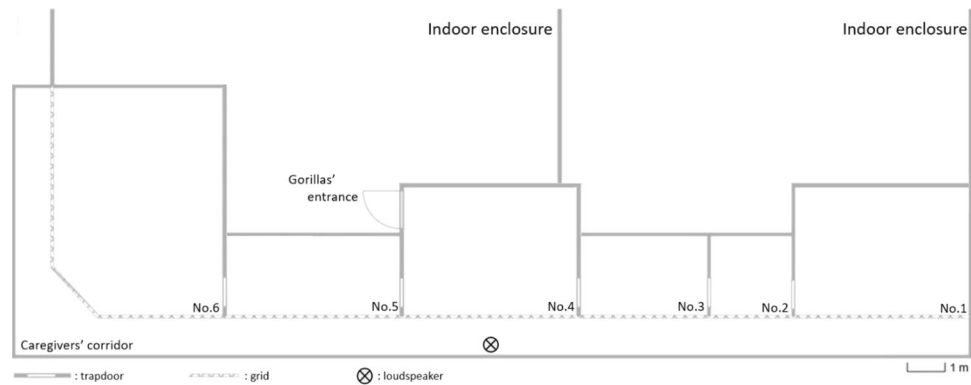
We acknowledge that our study is just one step towards a more advanced understanding of vocal interactions in great apes, and given the small sample size, some factors other than age might play a role. Indeed, despite all the advantages that the captivity offers to the study of vocal communication, gorillas cannot be isolated during experiments because of the risk of intense stress. Despite the randomisation of subject and condition orders and despite the fact that playbacks were time-spaced (over two months), a habituation effect cannot be discarded. Lastly, our result showing that young individuals responded strongly towards the “age-difference situation” (*i.e.* condition C) may suggest that the attentional response between each condition was driven by the age of individuals. Then, this difference of attentional state of the youngsters towards simulated vocal exchanges involving individuals of their own age range could induce a bias, mainly in the “overlapped vocal exchange” condition which include only adult individuals. To address this possible bias, a supplementary condition may be considered in future comparable studies. It would be similar to the “overlapped vocal exchange” but involving a young individual as first emitter who vocally interacts with the adult male, as the simulated vocal exchanges played back in this study (condition C). Our results are nevertheless promising and seem pointing out the possible existence of conversational rules managing vocal interactions in western lowland gorillas<sup>43</sup>. Future studies may consider investigating multiple and larger groups with a greater age-category range, and may test for potential intergroup variations. In the same vein, Levréro and colleagues<sup>40</sup> highlighted that bonobos are able to spontaneously display primitive conversation rules guided by social bonds, and Lameira and colleagues<sup>79,80</sup> showed that orang-utans (*Pongo pygmaeus*) have the cognitive abilities to produce spontaneous vocal exchanges. Only one study in apes, conducted in the field on male chimpanzees, revealed that the production of long calls (*i.e.* pant-hoots) is non-temporally organized (*i.e.* emissions are rather isolated calls without consecutive calls from other group members, or cases of synchronous choruses<sup>46</sup>). A recent study however has showed that chimpanzees have the ability to engage in gestural turn-taking<sup>81</sup>. All these findings suggest that all great apes possess the cognitive skills to engage in a “conversation-like” interaction. These results are promising and deserve further comparative work.

In conclusion, these recent studies on great apes highlight how vocal interactions are governed by social rules in our closest relatives, as they are across human cultures. Our findings help to fill the gap between nonhuman and human primate communication, particularly in terms of our understanding of factors that shape vocal interactions. The increase of the complexity of social interactions (from birds to mammals including monkeys, nonhuman apes and humans) may be at the origin of the most complex rules governing human communication. Parallels between human language and nonhuman primate vocal communication appear multifaceted, *e.g.*, information coding in acoustic structure, syntactic structure and socially-determined plasticity<sup>49,82,83</sup>. Here, a vast research area is opening into the developmental acquisition of conversational rules. The function, as well as the content of information in vocal exchanges remains unclear and still needs to be explored more systematically. Yet what is clear, is that the capacity to interact vocally with others and to respect ruled vocal interactions could have been an important step in the evolution of the vocal communication of primates<sup>36</sup>. This ability appeared before the capacity to articulate sounds in speech and to learn new acoustic structures, and is thus likely rooted deep in the animal lineage.

## Methods

**Compliance with ethical standards.** All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The study has been conducted in accordance with the current laws in France (agreement with the 2010/63/UE). Animal husbandry and care were under the management of the animal caregivers of the ZooParc de Beauval, France. Our study was based on non-invasive observations and





**Figure 4.** Diagram of the layout of the cages.

Subject name (birth date)	Age-sex category <sup>a</sup>	Mother's name (all fathered by the adult male)	Cage number (date and new cage number if changed during the experimental period)	Number of calls used in built vocal exchanges
Inge (02/03/1980)	Adult female		No. 1	4
Kabinda (10/12/1982)	Adult female		No. 5	4
Sheila (18/02/1991)	Adult female		No. 6	4
Mapenzi (14/04/2010)	Subadult male	Kabinda	No. 3	0
Kuimba (19/08/2010)	Subadult female	Tamarilla <sup>†</sup>	No. 2	2
Mayélé (21/03/2013)	Juvenile female	Kabinda	No. 5 (05/11/2018, No. 3)	2

**Table 1.** Tested individuals in the payback experiment. <sup>a</sup>Age-sex categories defined according to Gatti and colleagues<sup>89</sup>; <sup>b</sup>Loudspeaker location; <sup>†</sup>Dead individual (10/02/2017).

was approved by the ethics committee “Comité Rennais d’Ethique en matière d’Expérimentation Animale” (*i.e.* Rennes Ethical comity for experiments using animals; CNREEA 2018041710224608-APAFIS#14920).

**Animals and housing conditions.** The sample comprised twelve captive-born western lowland gorillas, including one adult silverback male, four adult females, and seven (four males and three females) immature individuals. The composition of the captive group reflects that of wild groups<sup>84,85</sup>. All adult females were unrelated. The adult male had sired all the immature individuals of the group.

The animals were housed at the ZooParc de Beauval (France) in an enriched building of 720 m<sup>2</sup> (6 m high, with glass walls, wired roof, straw litter and wood perches, as well as temperature and humidity control), connected to several adjacent cages totalling 59 m<sup>2</sup> (average size: 5 m<sup>2</sup>) with visual and auditory access between them. Cages were aligned from right to left (*i.e.* from No. 1 to No. 6), and cage No. 4, of the silverback male, was central (Fig. 4). Depending on the weather conditions, group had access to a 3,500 m<sup>2</sup> outdoor enclosure (with grass, bush, and wood perches). Gorillas shared the enclosure with eight patas monkeys (*Erythrocebus patas*).

Gorillas were trained to enter the cages daily for breakfast and medical-training between 8:30 and 10:30 a.m., returning afterward to the large enclosure in which they spent the rest of the day and night. While they were free to go in any cages, each gorilla used to settle every day in a determined cage (one individual per cage, except for mother and offspring who may stay together). Once they were in their usual preferred cage, caregivers closed the trapdoors between each cage in order to avoid food theft (see Table 1 for cage occupation).

**Vocal recordings.** Between April and July 2018, we recorded 474 spontaneous grunts, produced by the twelve individuals during 33 breakfast sessions for a total of 27 hours of recording, using all occurrence sampling<sup>86</sup>. We pooled together all possible grunt structures (*i.e.* atonal and tonal grunts, single and double grunts<sup>53,54</sup>). Recordings were done at an average distance of 2 meters using a Marantz PMD660 handheld digital audio recorder (Marantz, Japan) (sample rate 44.1 kHz, resolution 32bits, wav. format) connected to a directional Sennheiser MKH70-1 microphone (Sennheiser, Germany). We identified 125 vocal interactions, based on Lemasson and colleagues’ definition<sup>43</sup> for gorillas (*i.e.* series of calls emitted by different callers with an inter-call duration of less than 3 seconds). Vocal interactions typically involved callers from different cages (usually only two callers: 77%) and non-overlapped exchanged calls (92%). The adult male contributed 70% of these vocal interactions and was significantly more often a respondent than an initiator (Binomial test,  $N_{1st\ position} = 20$ ,  $N_{2nd\ position} = 41$ ,  $p = 0.01$ ). He responded more often to adult females (85% of response) than to young individuals (*i.e.* sub-adults, juvenile and infants, 15%).

**Playback experiments.** Playback experiments took place between September and December 2018. Participating gorillas never experienced any playback experiment in the past. Here, we tested their responses to

the broadcast of three types of vocal exchanges, *i.e.* one control condition, one violating the temporal expectation and one violating the social expectation:

- Condition A, “likely vocal exchange”: we built a control dyadic vocal exchange (as described in Lemasson *et al.*<sup>41</sup>) between two individuals close in age (*i.e.* both adults), that respected the average inter-call duration of 500 milliseconds (ms) (see Supplementary Fig. S1);
- Condition B, “overlapped vocal exchange”: we built an unlikely vocal exchange that did not obey the temporal turn-taking rule, with 50% of overlap between two vocalisations<sup>87</sup> emitted by two adults (see Supplementary Fig. S2);
- Condition C, “age-difference vocal exchange”: we built an unlikely vocal exchange, with an inter-call duration of 500 ms, that did not obey the social rule with two callers of distant ages (*i.e.* initiator: sub-adult or juvenile, mean age  $\pm$  SD =  $6.39 \pm 1.33$  years; respondent: adult male, 27 years; see Supplementary Fig. S3).

Vocal exchange acoustic stimuli were created using the vocalisations recorded during the aforementioned observations, discarding non-exchanged grunts, and using PRAAT and Goldwave softwares. Calls of initiators and respondents were systematically concatenated to respect their original position from dyadic natural vocal interactions previously recorded. Even the control dyadic “likely vocal exchanges” were built to prevent any bias due to artificial concatenation of calls. To match the most frequent vocal events occurring in natural conditions, we decided to use the adult male as the respondent in all acoustic stimuli and different females as initiators. To prevent kin-bias, subjects only heard unrelated female initiators. Thus, a given adult female gorilla was allotted to a given subject in conditions A and B. To prevent pseudo-replication, a given call exemplar was used only once. Given that spontaneous vocal interactions (*i.e.* abovementioned recordings) contain an equal number of grunts and double grunts, we used 50% of grunts and 50% of double grunts. The total durations of the played back stimuli in conditions A and C (excluding inter-call durations) were respectively  $163 \pm 26$  ms and  $146 \pm 30$  ms, and  $125 \pm 36$  ms in condition B.

Experiments took place during the usual breakfast in the corridor facing the cages. Sounds were played using a TAG Premio8.2 loudspeaker connected to the Marantz player at 60 dB SPL. The intensity of 60 dB corresponds to the grunts recorded in this group in similar conditions at an average distance of 2 meters. This intensity threshold matches the intensity used in other nonhuman primate studies (*e.g.* Japanese macaques<sup>50</sup>, bonobos<sup>88</sup>). The loudspeaker was systematically placed before the arrival of gorillas in the cages at 1 meter in front of cage No.4 (of the silverback male, see Fig. 4) and could not be seen by the other gorillas. The experimenter (LP) sat in front of the tested animal and fed it routinely with fruits, dry-food and lastly vegetables (always in this order), through the wire mesh. The camera was positioned behind the experimenter, above his head. The breakfast typically lasted around 30 minutes and the playback occurred at mid-term. The experimenter offered food to the subject in order to control the central position of his body and head at T0 (*i.e.* playback time). Before broadcasting the stimuli, some conditions had been respected to make the situation as plausible as possible: (*i*) both initiators and respondents used in a stimulated vocal exchange on the same side of the loudspeaker, both at the opposite side of the tested individual, and were not visible from the subject point of view, (*ii*) no vocalisation was heard within the two minutes preceding the playback. To prevent any possible influence from the experimenter, the experimenter’s gaze had to focus on the gorilla’s chest during the entire feeding period. To habituate animals to the experimental setting, we ran ‘silent sessions’ during which the full equipment was set up to mimic the conditions of real playback sessions but without any sound broadcast. Twenty days before starting the first playback session, we performed ‘silent sessions’ once a day, every day. We then performed randomly ‘silent sessions’ throughout the entire duration of the experiment. To limit habituation, no more than one subject was tested per day and each subject was tested at average intervals of 27 days (min: 3, max: 73 days apart). This variation is due to the fact that we had to stop our experiments on two occasions, after 11 and 28 days respectively, because of on-going conflicts in the group or construction work in the enclosure.

Six individuals were tested in the playback paradigm (Table 1). The three infants were excluded from the study as they were still dependent on their mothers. The silverback male was not tested as he was involved in the creation of all acoustic stimuli. We discarded one adult female and one subadult female, because they were too stressed during the experiments, either by the presence of the experimenter or by a conflict with other group members. Each tested gorilla heard each of three distinct conditions, resulting in 18 (3  $\times$  6) stimuli in total. Subject order and condition order were both randomized.

**Recording of behavioural responses.** We filmed the tested gorilla’s behaviours before and after playback. All videos were analysed in slow motion (scale of millisecond) using Kinovea 80.8.15 software for coding, while the experimenter was ‘blind’ to the experimental condition. The behavioural responses measured were the difference in the total duration of head orientation and in the total number of occurrences of locomotion (already observed in bonobos towards similar stimuli<sup>87</sup>) in the direction of the loudspeaker before and after the playback. The time windows used for the measurements were 60 seconds before and 60 seconds after the onset of the second vocalisation (*i.e.* T0 = onset of male calling, as this is the moment when subjects typically detect the breaking of the rule, as in Bouchet and colleagues<sup>50</sup>). This correction was chosen because contact call exchanges are not rare events that would typically trigger a strong surprise, because inter-cage gaze exchanges are frequent at breakfast time and lastly because subjects differed in their baseline interest towards their social environment; we thus wanted to capture not only the degree of interest towards the playback area but more importantly the increase/decrease of interest towards this area, as in other previous studies<sup>57,90,91</sup>. We also measured the first latency to reposition the head front (*i.e.* facing front the experimenter), corresponding to the duration of the first gaze towards the loudspeaker (starting the measure at T0 if the individual turned its head before the onset of the second vocalisation).

Normality and homogeneity of variance were assessed by inspecting residuals following Shapiro-Wilk W tests. Given the non-normally distributed data, we ran non-parametric tests<sup>92</sup> with a significance threshold of 0.05. The difference of total duration of head orientation, of locomotion occurrences, and the latencies to reposition the head between likely (condition A) and unlikely vocal exchanges (conditions B and C) were tested using exact Wilcoxon signed-rank tests. Spearman's correlation tests were then run to investigate the possible influence of the age of the subjects on their responses to each situation. Data were analysed in R Studio0.99.903 (R version3.3.1: R Core Team2016).

## Data availability

The data generated and analysed during the current study are available from the corresponding author on reasonable request.

Received: 23 September 2019; Accepted: 7 April 2020;

Published online: 24 April 2020

## References

1. Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci.* **106**, 10587–10592 (2009).
2. Ghiglione, R. *L'homme communiquant.* (1986).
3. Sacks, H., Schegloff, E. A. & Jefferson, G. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language (Baltim.)* **50**, 696–735 (1974).
4. Henry, L., Craig, A. J. F. K., Lemasson, A. & Hausberger, M. Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Front. Psychol.* **6**, 1416 (2015).
5. Lemasson, A., Ouattara, K., Petit, E. J. & Zuberbühler, K. Social learning of vocal structure in a nonhuman primate? *BMC Evol. Biol.* **11**, 1–7 (2011).
6. Snowdon, C. T. & Cleveland, J. "Conversations" among pygmy marmosets. *Am. J. Primatol.* **7**, 15–20 (1984).
7. Sugiura, H. & Masataka, N. Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). in *Current topics in primate vocal communication* (eds. Zimmermann, E., Newman, J. D. & Jürgens, U.) 121–140, <https://doi.org/10.1007/978-1-4757-9930-9> (1995).
8. Bouchet, H., Pellier, A. S., Blois-Heulin, C. & Lemasson, A. Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *Am. J. Primatol.* **72**, 360–375 (2010).
9. Koda, H. Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* **141**, 1279–1296 (2004).
10. Papworth, S., Böse, A.-S., Barker, J., Schel, A.-M. & Zuberbühler, K. Male blue monkeys alarm call in response to danger experienced by others. *Biol. Lett.* **4**, 472–475 (2008).
11. Clark, A. P. & Wrangham, R. W. Acoustic analysis of wild chimpanzee pant hoots: Do Kibale Forest chimpanzees have an acoustically distinct food arrival pant hoot? *Am. J. Primatol.*, <https://doi.org/10.1002/ajp.1350310203> (1993).
12. Dahlin, C. R. & Benedict, L. Angry birds need not apply: A perspective on the flexible form and multifunctionality of avian vocal duets. *Ethology* **120**, 1–10 (2014).
13. Hall, M. L. A Review of Vocal Duetting in Birds. *Adv. Study Behav.* **40**, 67–121 (2009).
14. Hall, M. L. A review of hypotheses for the functions of avian duetting. *Behav. Ecol. Sociobiol.* **55**, 415–430 (2004).
15. Geissmann, T. & Orgeldinger, M. The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim. Behav.* **60**, 805–809 (2000).
16. Forester, D. C. & Harrison, W. K. The Significance of Persistent Vocalisation By the Spring Peeper, *Pseudacris Crucifer* (Anura: Hylidae). *Behaviour* **103**, 1–15 (2008).
17. Bailey, W. J. Insect duets: Underlying mechanisms and their evolution. *Physiol. Entomol.* **28**, 157–174 (2003).
18. Okobi, D. E. J., Banerjee, A., Matheson, A. M. M., Phelps, S. M. & Long, M. A. Motor cortical control of vocal interaction in neotropical singing mice. *Sci.* **363**, 979–983 (2019).
19. Clutton-Brock, T. H. & Albon, S. D. The Roaring of Red Deer and the Evolution of Honest Advertisement. *Behaviour* **69**, 145–170 (1979).
20. Reby, D. & McComb, K. Vocal Communication and Reproduction in Deer. *Adv. Study Behav.* **33**, 231–264 (2003).
21. Briseño Jaramillo, M., Guyot-Biquand, V., Estrada, A. & Lemasson, A. Vocal repertoire of free-ranging black howler monkeys' (*Alouatta pigra*): call types, contexts and sex-related contributions. *Am. J. Primatol.* **79**, 22630–10 (2017).
22. Charlton, B. D. *et al.* Perception of male caller identity in koalas (*Phascolarctos cinereus*): Acoustic analysis and playback experiments. *Plos One* **6**, e20329 (2011).
23. Gros-Louis, J. & Mitani, J. Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**, 1041–1064 (1998).
24. France, E. F., Anderson, A. H. & Gardner, M. The impact of status and audio conferencing technology on business meetings. *Int. J. Hum. Comput. Stud.* **54**, 857–876 (2001).
25. Pika, S., Wilkinson, R., Kendrick, K. H. & Vernes, S. C. Taking turns: bridging the gap between human and animal communication. *Proc. R. Soc. B* **285**, 1–9 (2018).
26. Leong, K. M., Ortolani, A., Burks, K. D., Melleni, J. D. & Savage, A. Quantifying Acoustic and Temporal Characteristics of Vocalizations for a Group of Captive African Elephants *Loxodonta Africana*. *Bioacoustics Int. J. Anim. Sound its Rec.* **13**, 213–231 (2003).
27. Schulz, T. M., Whitehead, H., Gero, S. & Rendell, L. Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Anim. Behav.* **76**, 1977–1988 (2008).
28. Janik, V. M. Whistle Matching in Wild Bottlenose Dolphins (*Tursiops truncatus*). *Science (80-)*. **289**, 1355–1357 (2000).
29. Yoshida, S. & Okanoya, K. Evolution of turntaking: a bio-cognitive perspective. *Cogn. Stud.* **12**, 153–165 (2005).
30. Carter, G. G., Skowronski, M. D., Faure, P. A. & Fenton, B. Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim. Behav.* **76**, 1343–1355 (2008).
31. Lemasson, A., Gandon, E. & Hausberger, M. Attention to elders' voice in non-human primates. *Biol. Lett.* **6**, 325–328 (2010).
32. Takahashi, D. Y., Narayanan, D. Z. & Ghazanfar, A. A. Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Curr. Biol.* **23**, 2162–2168 (2013).
33. Kondo, N., Watanabe, S. & Izawa, E. I. A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithol. Sci.* **9**, 83–91 (2010).
34. Reber, S. A., Townsend, S. W. & Manser, M. B. Social monitoring via close calls in meerkats. *Proc. R. Soc. B Biol. Sci.* **280** (2013).
35. Soltis, J., Leong, K. & Savage, A. African elephant vocal communication I: Antiphonal calling behaviour among affiliated females. *Anim. Behav.* **70**, 579–587 (2005).
36. Dunbar, R. *Grooming, gossip and the evolution of language.* (Faber & Faber, 1996).



37. Cutting, A. L. & Dunn, J. Conversations with siblings and with friends: Links between relationship quality and social understanding. *Br. J. Dev. Psychol.* **24**, 73–87 (2006).
38. Briseno-Jaramillo, M. *et al.* Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behav. Ecol. Sociobiol.* **72**, 192 (2018).
39. Arlet, M., Jubin, R., Masataka, N. & Lemasson, A. Grooming-at-a-distance by exchanging calls in non-human primates. *Biol. Lett.* **11**, 20150711–20150714 (2015).
40. Levréro, F. *et al.* Social bonding drives vocal exchanges in Bonobos. *Sci. Rep.* **9**, 711 (2019).
41. Bascom, W. R. The principle of seniority in the social structure of the Yoruba. *Am. Anthropol.* **44**, 37–46 (1942).
42. Chen, H. C., Kaplan, G. & Rogers, L. J. Contact calls of common marmosets (*Callithrix jacchus*): Influence of age of caller on antiphonal calling and other vocal responses. *Am. J. Primatol.* **71**, 165–170 (2009).
43. Lemasson, A., Pereira, H. & Levréro, F. Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *J. Comp. Psychol.* **132**, 141–151 (2018).
44. Giles, H., Coupland, N. & Coupland, J. Accomodation Theory: Communication, Context, and Consequence. *Context. Accommod. Dev. Appl. Socioling.* 1–68 (1991).
45. Candiotti, A., Zuberbühler, K. & Lemasson, A. Convergence and divergence in Diana monkey vocalizations. *Biol. Lett.* **8**, 382–385 (2012).
46. Arcadi, A. C. Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *J. Hum. Evol.* **39**, 205–223 (2000).
47. Ginsburg, G. P. & Kilbourne, B. K. Emergence of vocal alternation in mother-infant interchanges. *J. Child Lang.* **15**, 221–235 (1988).
48. Stern, D. N., Jaffe, J., Beebe, B. & Bennett, S. L. Vocalizing in unison and in alternation: Two modes of communication within the mother-infant dyad. *Ann. N. Y. Acad. Sci.* **263**, 89–100 (1975).
49. Lemasson, A. *et al.* Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Sci. Rep.* **1**, (2011).
50. Bouchet, H., Koda, H. & Lemasson, A. Age-dependent change in attention paid to vocal exchange rules in Japanese macaques. *Anim. Behav.* **129**, 81–92 (2017).
51. Cheney, D. L., Seyfarth, R. M. & Silk, J. B. The Responses of Female Baboons (*Papio cynocephalus ursinus*) to Anomalous Social Interactions: Evidence for Causal Reasoning? *J. Comp. Psychol.* **109**, 134–141 (1995).
52. Wittig, R. M., Crockford, C., Seyfarth, R. M. & Cheney, D. L. Vocal alliances in Chacma baboons (*Papio hamadryas ursinus*). *Behav. Ecol. Sociobiol.* **61**, 899–909 (2007).
53. Hedwig, D., Robbins, M. M., Mundry, R., Hammerschmidt, K. & Boesch, C. Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour* **151**, 1091–1120 (2014).
54. Hedwig, D., Mundry, R., Robbins, M. M. & Boesch, C. Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Anim. Cogn.* **18**, 423–435 (2015).
55. Salmi, R., Hammerschmidt, K. & Doran-Sheehy, D. M. Individual distinctiveness in call types of wild western female gorillas. *PLoS One* **9**, e101940 (2014).
56. Zuberbühler, K. & Wittig, R. M. Field experiments with non-human primates: A tutorial. in *Field and laboratory methods in primatology: A practical guide* (eds. Setchell, J. M. & Curtis, D. J.) 207–224 (Cambridge University Press, 2011).
57. Lemasson, A., Palombit, R. A. & Jubin, R. Friendships between males and lactating females in a free-ranging group of olive baboons (*Papio hamadryas anubis*): Evidence from playback experiments. *Behav. Ecol. Sociobiol.* **62**, 1027–1035 (2008).
58. Levréro, F. *et al.* Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nat. Commun.* **6**, 7609 (2015).
59. Hayashi, A., Tamekawa, Y. & Kiritani, S. Developmental Change in Auditory Preferences for Speech Stimuli in Japanese Infants. *J. Speech, Lang. Hear. Res.* **44**, 1189–1200 (2001).
60. Kinzler, K. D., Dupoux, E. & Spelke, E. S. The native language of social cognition. *Proc. Natl. Acad. Sci.* **104**, 12577–12580 (2007).
61. Lewis, M. & Goldberg, S. The Acquisition and Violation of Expectancy: An Experimental Paradigm. *J. Exp. Child Psychol.* **7**, 70–80 (1969).
62. Saffran, J. *et al.* Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* **107**, 479–500 (2008).
63. Wilson, B. *et al.* Auditory artificial grammar learning in Macaque and Marmoset monkeys. *J. Neurosci.* **33**, 18825–18835 (2013).
64. Meltzoff, A. N. & Borton, R. W. Intermodal matching by human neonates. *Nature* **282**, 403–404 (1979).
65. Kuhl, P. K. & Meltzoff, A. N. The bimodal perception of speech in infancy. *Science* (80-). **218**, 1138–1141 (1982).
66. Ghazanfar, A. A. & Logothetis, N. K. Facial expressions linked to monkey calls. *Nature* **423**, 937–938 (2002).
67. Reber, S. A. *et al.* Common marmosets are sensitive to simple dependencies at variable distances in an artificial grammar. *Evol. Hum. Behav.* **40**, 214–221 (2019).
68. Lemasson, A. *et al.* Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291 (2013).
69. Calame-Griaule, G. *Ethnologie et langage. la parole chez les dogon.* (1965).
70. Baker, T. M., Wilson, D. R. & Mennill, D. J. Vocal signals predict attack during aggressive interactions in black-capped chickadees. *Anim. Behav.* **84**, 965–974 (2012).
71. Ficken, M. S., Ficken, R. W. & Witkin, S. R. Vocal Repertoire of the Black-Capped Chickadee. *Auk* **95**, 34–48 (1978).
72. Dabelsteen, T., McGregor, P. K., Holland, J., Tobias, J. A. & Pedersen, S. B. The signal function of overlapping singing in male robins. *Anim. Behav.* **53**, 249–256 (2003).
73. Mennill, D. J. & Ratcliffe, L. M. Overlapping and matching in the song contests of black-capped chickadees. *Anim. Behav.* **67**, 441–450 (2004).
74. Harcourt, A. H. Social relationships between adult male and female mountain gorillas in the wild. *Anim. Behav.* **27**, 325–342 (1979).
75. Watts, D. P. Gorilla social relationships: A comparative overview. in *Gorilla biology: A multidisciplinary perspective* 302–327 (Cambridge University Press, 2002).
76. Cheney, D. L. & Seyfarth, R. M. *How monkeys see the world.* (University of Chicago Press, 1990).
77. Seyfarth, R. M. & Cheney, D. L. Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640–1658 (1986).
78. Koda, H. *et al.* Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. *PLoS One* **8** (2013).
79. Lameira, A. R., Maddieson, I. & Zuberbühler, K. Primate feedstock for the evolution of consonants. *Trends Cogn. Sci.* **18**, 60–62 (2014).
80. Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A. & Shumaker, R. W. Vocal fold control beyond the species-specific repertoire in an orang-utan. *Sci. Rep.* **6**, 30315 (2016).
81. Fröhlich, M. *et al.* Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci. Rep.* **6**, 1–14 (2016).
82. Bouchet, H., Coye, C. & Lemasson, A. Le Langage est-il le Propre de l’Homme? Apports des Etudes sur les Primates Non Humains. *Tétralogiques* **21**, 87–133 (2016).
83. Zuberbühler, K. The phylogenetic roots of language: Evidence from primate communication and cognition. *Curr. Dir. Psychol. Sci.* **14**, 126–130 (2005).
84. Robbins, M. M. *et al.* Social structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*). *Am. J. Primatol.* **64**, 145–159 (2004).
85. Parnell, R. J. Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *Am. J. Primatol.* **56**, 193–206 (2002).

86. Altmann, J. Observational Study of Behavior: Sampling Methods. *Behaviour* **49**, 227–267 (1974).
87. Levréro, F. *et al.* Bonobos converse and pay attention to breaking conversational rules. in *Behaviour 2017 - 35th International Ethological Conference* (ed. Instituto Universitário de Ciências Psicológicas, S. e da V. (ISPA); A. for the S. of A. B. (ASAB)) hal-01576988 (2017).
88. Keenan, S. *et al.* Enduring voice recognition in bonobos. *Sci. Rep.* **6**, (2016).
89. Gatti, S., Levréro, F., Ménard, N. & Gautierhion, A. Population and group structure of western Lowland Gorillas (*Gorilla gorilla gorilla*) at Lokoué, Republic of Congo. *Am. J. Primatol.* **63**, 111–123 (2004).
90. Palombit, R. A., Seyfarth, R. M. & Cheney, D. L. The adaptive value of ‘friendships’ to female baboons: experimental and observational evidence. *Anim. Behav* **54**, 599–614 (1997).
91. Lemasson, A., Boutin, A., Boivin, S., Blois-Heulin, C. & Hausberger, M. Horse (*Equus caballus*) whinnies: a source of social information. *Anim. Cogn.* **12**, 693–704 (2009).
92. Siegel, S. & Castellan, N. J. *Nonparametric statistics for the behavioral sciences* (2nd ed.) (1988).

## Acknowledgements

We are deeply grateful to the ZooParc de Beauval and the gorilla caregivers for their valuable assistance. We also thank Katarzyna Pisanski and Melissa Sebilliau for correcting our English. This work was financially supported by a CIFRE grant (#2018\_00074) funded by the Association Beauval Nature pour la Conservation et la Recherche, Centre National de la Recherche Scientifique, Rennes 1 University and Lyon/Saint-Etienne University.

## Author contributions

L.P., F.L., B.M. and A.L. conceived, designed and coordinated the study. L.P. collected data and carried out the statistical analyses. L.P., F.L. and A.L. interpreted and drafted the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-63923-7>.

**Correspondence** and requests for materials should be addressed to L.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020