

SCIENTIFIC REPORTS



OPEN

Social bonding drives vocal exchanges in Bonobos

Florence Levréro¹, Sonia Tuitou¹, Julia Frédet¹, Baptiste Nairaud¹, Jean-Pascal Guéry² & Alban Lemasson³

The origin of human speech is still a hotly debated topic in science. Evidence of socially-guided acoustic flexibility and proto-conversational rules has been found in several monkey species, but is lacking in social and cooperative great apes. Here we investigated spontaneous vocal interactions within a peaceful context in captive bonobos to reveal that vocal interactions obey temporally and social rules. Dyadic vocal interactions were characterized by call overlap avoidance and short inter-call intervals. Bonobos preferentially responded to conspecifics with whom they maintained close bonds. We also found that vocal sharing rate (production rate of shared acoustic variants within each given dyad) was mostly explained by the age difference of callers, as other individual characteristics (sex, kinship) and social parameters (affinity in spatial proximity and in vocal interactions) were not. Our results show that great apes spontaneously display primitive conversation rules guided by social bonds. The demonstration that such coordinated vocal interactions are shared between monkeys, apes and humans fills a significant gap in our knowledge of vocal communication within the primate phylogeny and highlights the universal feature of social influence in vocal interactions.

The evolutionary origins of language and speech remains a fundamental question in science. In particular, whether clues to the origins of speech are present in nonhuman primate communication remains a hotly debated topic^{1–4}. Despite the diversity of social cultures and languages in humans, universal features in conversations are found across all languages, such as the avoidance of overlapping and a minimum gap between turns^{5–8}. Orderly vocal exchanges (antiphony between two or more animals or duets within male–female pairs⁹) have been found across the primate order: from lemurs¹⁰, to New World monkeys^{11–15}, Old World monkeys¹⁶ and lesser apes^{17,18}. Vocal turn-taking appears to be associated with social life and cooperation capacities^{1,2,5,9,19–21}. It is thought to maintain and reinforce social bonds between individuals (e.g. in non-human primates^{10,22}), enable the extraction of information in the absence of overlap (e.g.²³ but see²⁴) and reduce stress as in the case of social grooming²⁵.

Vocal exchange is “a characteristic communication style in which a sender produces a vocalization to address a receiver, and the receiver emits a call in response within a brief interval” (cited from²⁶). Vocal exchange patterns are influenced by social factors in non-human primates. ‘Interlocutors’ are not randomly selected, and preference is given to elders^{11,27–29}, social allies^{12,14,22} or dominant individuals^{30,31}. The attention of the audience also influences vocal outputs leading to persistence (repetition of calls) and elaboration (changes in the acoustic structure of calls) in situations where no response has been received^{32,33}. Shared primitive forms of vocal turn-taking within non-human primate species might suggest an ancient evolutionary origin^{1,34}. Surprisingly, however, studies based on great apes are scarce and controversial. No evidence of spontaneous vocal coordinated exchanges has been found in wild chimpanzees³⁵, who display complex social interactions and cooperative abilities³⁶. Indeed, Arcadi³⁵ found that chimpanzees do not “respond” to the majority of calls they heard (within 5 sec), and that instead, bonded males tend to chorus together, matching each-other’s pant hoots^{37,38}. Nevertheless, a recent study in great apes found for the first time that captive gorillas display some rule-governed call exchanges³¹. Relying on our current knowledge, vocal turn-taking is thus reported across phylogenetically distant species (monkeys and more generally in some social mammals such as African elephants³⁹, bottlenose dolphins⁴⁰, bats^{41,42}, naked mole-rats⁴³) but with some apparent discontinuities among great apes. More investigations among great ape species, our closest, highly social, relatives, are thus necessary in order to ascertain if vocal-turn taking behavior is as a result of convergent evolution (analogies as adaptations to similar social requirements) or is shared ancestry (homologies which are inheritance behaviours)³⁴.

¹Université de Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, Neuro-PSI, CNRS UMR 9197, Saint-Etienne, France. ²La Vallée des Singes, 86700, Romagne, France. ³Université de Rennes, EthoS “Ethologie Animale et Humaine,” UMR 6552 –CNRS-Université de Caen Normandie, Station Biologique, 35380, Paimpont, France. Correspondence and requests for materials should be addressed to F.L. (email: florence.levrero@univ-st-etienne.fr)

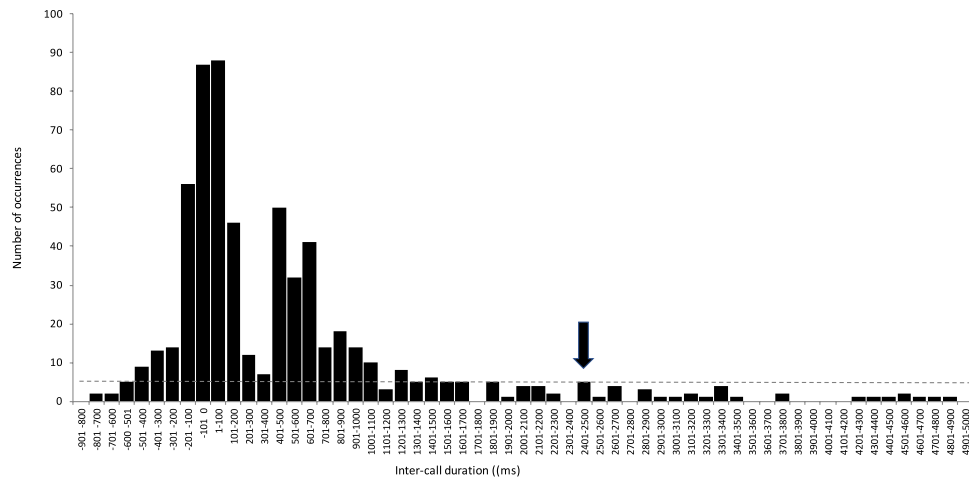


Figure 1. Distribution of inter-call durations between different consecutive callers recorded from 70 independent focal samples of 10 minutes each ($n = 779$ inter-call durations). Grey dotted line: baseline of the occurrence of inter-call durations. Black arrow: threshold indicating the maximum response delay of vocal exchanges (i.e. when the number of occurrences starts to be above the baseline level).

Social influences on the acoustic patterns of calls can also be assessed by examining the rate of vocal sharing, defined as the production rate of vocal ‘variants’ between individuals at a given time²⁶. The study of vocal sharing in nonhuman primates provides insight into vocal flexibility, and particularly into the flexibility of call production and call use. In several species, it has been shown that a given individual can produce several acoustic variants (i.e. stereotypical patterns of frequency modulation) of the same affiliative call type at a given stage of its life, with some variants being potentially shared with preferred group members e.g.^{26,44}. Vocal sharing between affiliative or kin partners has been identified in mouse lemurs⁴⁵, pygmy marmosets⁴⁶, Wied’s black-tufted-ear marmoset⁴⁷, Campbell’s monkeys^{26,48}, siamang gibbons^{49,50} and recently in captive chimpanzees⁵¹. However, what adaptive benefits are conferred by vocal sharing remains an open question. The most parsimonious explanation is that it functions to reinforce social bonding, facilitating navigation (mate identification and reunion, joint territory defense) and aiding social integration (gibbons^{49,50}, marmosets^{52,53}, chimpanzees^{38,51,54} and humans^{55,56}).

In this study, we attempted to identify features that are already known to characterize both monkey call exchanges and human conversations in a great ape species, the bonobo (*Pan paniscus*). Bonobos live in dense social network with a fission-fusion social system³⁷ and present high cognitive and cooperative abilities e.g.⁵⁸. As with chimpanzees, the vocal mode of communication plays an important role in bonobos due to their forest habitat. Their vocal repertoire is graded^{59,60}, meaning that call sub-structures are extremely variable. Surprisingly, to the best of our knowledge, neither the temporal organization of vocal interactions nor the extent of vocal sharing, have previously been investigated in bonobos. Here, to plug this gap, we examined the temporal organization of vocal interactions and the social factors that influence vocal sharing in a captive group of bonobos. We expected to observe some level of vocal turn-taking in specific contexts and to identify social factors that control the patterns of vocal exchanges. More specifically, we predicted more frequent vocal interactions and a higher rate of vocal exchanges between preferred social partners.

Results

Occurrence of ‘vocal responses’. We studied whether call utterances within the group occurred randomly or whether they were temporally organized with a given call being followed by a so-called ‘vocal response’³⁷. In peaceful contexts (see Methods), we found that calls often elicited a vocal response from other group members. Vocal responses were, however, not systematic and three calling patterns were observed: overlapped calling, successive calling and isolated calling (Fig. 1). Overlapped calling consisted in calls emitted simultaneously by different individuals (i.e. inter-call durations were negative, meaning that the onset of the call “response” started before the offset of the first call and consequently the two calls partially overlapped) (Fig. 1). Successive calling, sometimes referred to antiphony, consisted of a temporal synchronization with no overlap of the call utterances between two consecutive callers (i.e. inter-call durations were positive and short, meaning that there was a short silence between the two consecutive calls) (Fig. 1; see also a typical sequence of vocal exchanges in the Supporting Audio File). Isolated calling consisted of calls emitted independently from the call utterances of other group members (i.e. inter-call durations were above 2500 ms, which has been defined as the baseline threshold) (see Fig. 1).

In total we recorded 779 call intervals within 70 ten-minute sequences. We recorded 48.8% of successive calling (designated as vocal exchanges hereafter; 385/779 call intervals), 26.4% of isolated calls (206/779) and 24.1% overlapped calling (188/779). Vocal exchanges thus followed two basic temporal rules: (1) relatively rare call overlap, since 67.2% of inter-call durations were positive ($n = 385/573$; Fig. 1), and (2) short response delay (<2500 ms). Interestingly, Fig. 1 shows a bimodal distribution with a first peak including both overlapped calls (between -200 ms and 0 ms) and positive and very short call intervals (between 1 ms and 200 ms). The second peak of the bimodal distribution (Fig. 1) shows that positive inter-call durations were usually above 400 ms

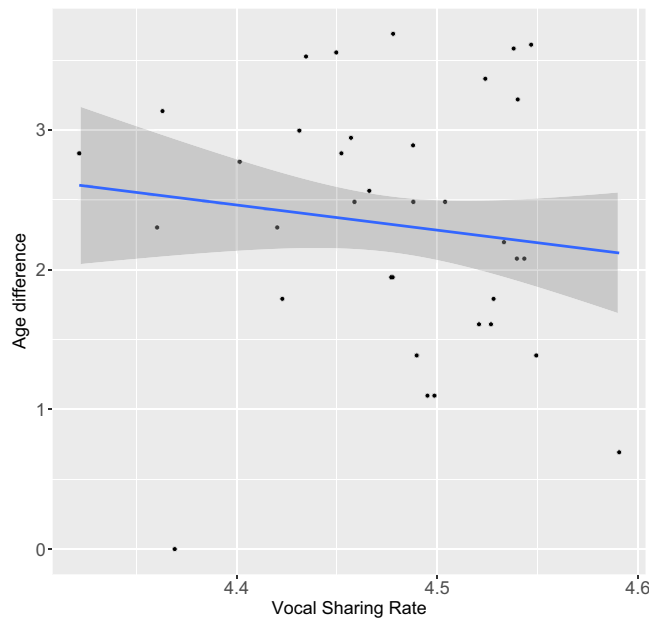


Figure 2. Effect of age difference on vocal sharing rate. ‘Vocal sharing’ was calculated for each dyad as the average of individual similarities in ‘variant’ rates of each acoustic pattern. Predictors were log-transformed. Solid line: linear regression of the estimated marginal means, grey shaded area: confidence interval.

(60.3% of positive inter-call durations $n = 232/385$), indicating that individuals typically leave a gap of a bit more than an average call duration ($247 \text{ ms} \pm 179 \text{ ms}$, this study) before responding, thus preventing overlap.

Group members did not equally participate in vocal exchanges ($n = 310$ calls extracted from a sub-sample composed of 192 vocal exchanges (see Methods and Table S1); $\chi^2 = 146.9$, $df = 8$, $P < 0.001$).

Lastly, when we examined all vocal sequences (i.e. not only dyadic vocal exchanges but all vocal sequences composed of several calls spaced by less than 2500 ms each; $n = 118$), we observed that vocal sequences typically involved two or three callers (78.3% and 18.1% of the vocal sequences respectively; mean 2.2 ± 0.5 individuals, maximum number of individuals: 4), with an average of 3.7 ± 2.1 calls in vocal interactions.

Classification of calls. The 310 exchanged calls were pre-classified into six acoustically distinct ‘variants’, based on the shape of fundamental frequency modulation pattern through visual inspections of spectrograms (see Methods and Fig. S1). Most of the exchanged calls (83%) corresponded to “peep yelps” and “peeps” according to bonobo repertoire in literature (see Methods and Fig. S1 for more details and comparisons with call types classically described in bonobo literature). A PCA followed by a DFA analysis, based on six acoustic parameters measured in the temporal and spectral domains (call duration, start frequency, maximum frequency, end frequency, ascending slope, and descending slope; see details in Methods) validated this classification. The averaged correct classification of the different ‘variants’ was 91.6% (min-max: 60–100%), largely higher than a correct classification by chance (one chance out of six: 16.7%). All the acoustic parameters significantly participated in discriminating the six ‘variants’ (*Wilks’ lambda* = 0.034, $F_{(\text{approx.}30, 1198)} = 53.019$, $P < 0.001$, see Fig. S1 for more details).

Production rate of vocal ‘variants’. During vocal exchanges, the six identified ‘variants’ were not produced in similar proportions ($\chi^2 = 171.4$, $df = 5$, $P < 0.001$) with the variants A, B and C being the most frequent (34.8%; 25.5% and 24.5% respectively, Table S1).

The production rate of the different ‘variants’ differed between individuals ($\chi^2 = 67.86$, $df = 40$, $P = 0.004$). The most often recorded ‘variant A’ (108/310 calls) represented 12.5% to 66.7% of the calls in individual repertoires. Of the six defined ‘variants’, variants A, B and D were produced by all the studied individuals ($n = 9$), variant C by eight of the individuals and variants E and F by less than half of the individuals. In order to identify the factors explaining these individual differences, we compared hereunder the vocal sharing rate, which is the production rate of vocal ‘variants’ between individuals calculating a vocal similarity score per pair of individuals (see Methods for a complete definition and calculation detail).

Socio-demographic determinants of vocal sharing. We assessed the influence of individual characteristics (age, sex, kinship) and social (affinity) factors of interacting partners on vocal sharing rate. Social affinity between individuals was defined by dyadic spatial proximities (see Methods). We found that social affinity differed among bonobos. Indeed, most of bonobos did not spend their time randomly in proximity to each group member (for seven ind.: $30.5 < \chi^2 < 187.9$, $df = 7$, $P < 0.001$; for one ind.: $\chi^2 = 21.1$, $df = 7$, $P < 0.01$; for one ind.: $\chi^2 = 1.6$, $df = 7$, $P = 0.97$). We found that only age difference between callers significantly affected vocal sharing rate (i.e. differences in call rates for each given dyad, see Methods; $n = 72$; GLM: $\chi^2 = 8.19$, $df = 1$, $P = 0.004$; Fig. 2, Table 1). The greater the age difference between two individuals, the less likely they are to share acoustic

| Fixed effects | χ^2 | d.f | P-value |
|--------------------------------|----------|-----|--------------|
| Social Affinity | 0.132 | 1 | 0.717 |
| Vocal Affinity | 1.089 | 1 | 0.765 |
| Sex composition of vocal dyads | 0.031 | 1 | 0.86 |
| Age difference | 8.19 | 1 | 0.004 |
| Kinship | 1.313 | 1 | 0.252 |

Table 1. Statistical analysis of the effect of socio-demographic determinants on vocal sharing rate between two consecutive callers. P-values were obtained independently for each predictor with likelihood-ratio tests comparing the fit of the full model with a reduced model lacking the fixed effect. Significant P-values are given in bold. Generalized Linear mixed model (GLMM) with a Gaussian error structure. Normality of the residuals was attained after applying a log-transformation.

| Voc 1 | Voc 2 | | | | | |
|--------|-----------|-----------|---------|----------|--------|--------|
| | A | B | C | D | E | F |
| A (62) | 48.4 (30) | 17.7 | 22.6 | 9.7 | 1.6 | 0 |
| B (51) | 29.4 | 37.3 (19) | 19.6 | 9.8 | 3.9 | 0 |
| C (50) | 22 | 26 | 30 (15) | 14 | 4 | 4 |
| D (19) | 42.1 | 15.8 | 21.1 | 21.1 (4) | 0 | 0 |
| E (5) | 0 | 20 | 20 | 40 | 20 (1) | 0 |
| F (5) | 40 | 0 | 20 | 20 | 0 | 20 (1) |

Table 2. Matrix representing the percentage of the different ‘variant’ combinations in vocal exchanges ($n = 192$ pairs of consecutive calls). On the first left column, ‘Voc 1’ is the ‘variant’ starting a vocal exchange and in bracket is the number of observations where each variant was produced in first position. ‘Voc 2’ is the ‘variant’ given in ‘response’ namely produced within a maximum delay of 2.5 sec (see the body text for the definition of this threshold). ‘Variant’ matching is on the diagonal line and in brackets is the number of vocal exchanges where Voc 2 was the same that Voc 1.

‘variants’. The other tested fixed factors (social affinity rate, vocal affinity rate (i.e. dyadic vocal response rate, see Methods for detailed definition), sex composition of vocal dyads and kinship) were not significant predictors of the vocal sharing rate (see Table 1). Interestingly, all the variants could induce a vocal ‘response’ (Fisher’s Exact Test $P = 0.95$, Table S1) and the matching of ‘variants’ within vocal exchanges was not systematic. Indeed, we found that only 36.5% of vocal exchanges ($n = 192$) were composed of similar ‘variants’ (Table 2). It is worth noting that this rate still remains higher than ‘variant’ matching by chance (16.7%). A similar result was found when only taking into account the first two callers of all the vocal sequences recorded (composed of several calls spaced with a short response delay, see above) (39% of ‘variant’ matching, $n = 118$ of vocal exchanges between the two first contributors of vocal sequences; see Methods).

Socio-demographic determinants of vocal affinity. We assessed the influence of individual characteristics (age, sex, kinship) and social (affinity) factors of interacting partners on their amount of vocal interactions. Vocal affinity (the frequency of vocal interactions for each given dyad, see Methods for detailed definition) was only explained by the ‘social affinity’ of individuals ($n = 72$; GLM: $\chi^2 = 8.4$, $df = 1$, $P = 0.003$; Fig. 3 Table 3) while the sex of callers, age differences and kinships were not significant predictors. Thus the stronger the ‘social affinity’ between two individuals, the more they are preferred vocal partners.

Lastly, we identified preferred exchanging partners amongst individuals, as some bonobos induced more responses from specific group members than others ($\chi^2 = 89.77$, $df = 8$, $P < 0.001$, Figs 4 and 5). The dominant female (ind. 1) and her adult son (ind. 4) induced the highest numbers of vocal responses.

Discussion

Evidence of human-like conversations is scarce in great apes. Vocal exchanges have been reported in the context of long-distance communication in chimpanzees⁶¹ and bonobos⁶², as a means to coordinate their movement between parties. Here we demonstrate for the first time that orderly vocal exchanges in short distance communication exist in bonobos. We show that their vocal exchanges respect basic temporal (i.e. turn-taking, overlap avoidance) and social (i.e. interlocutor selectivity) rules, similar to the rules guiding human conversations. Second, vocal sharing (i.e. typically resulting from a progressive vocal convergence between at least two individuals) was determined by the age of individuals, independently of their social affinity, sex and kinship. Third, the frequency of vocal interactions between interlocutors was only influenced by social bonds.

In the context of spontaneous vocal interactions at short distances and in a peaceful context, we found that bonobos displayed two vocal interaction patterns (overlapped and successive calling). Overlapped calling may function as the well-known chorus in chimpanzees to maintain spatial contact over long distances⁶³ (notably to locate and facilitate reunions with other community members, especially males) while successive calling (vocal exchanges hereafter), which until now has only been described in gorillas³¹ among great apes, may

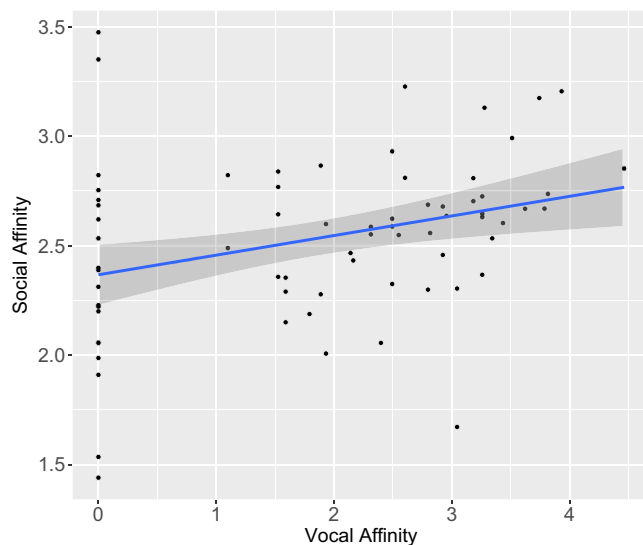


Figure 3. Effect of social affinity on vocal affinity. ‘Social affinity’ was calculated for each dyad as the frequency of occurrences of peaceful spatial proximities and ‘vocal affinity’ as the frequency of their vocal interactions. Predictors were log-transformed. Solid line: linear regression of the estimated marginal means, grey shaded area: confidence interval.

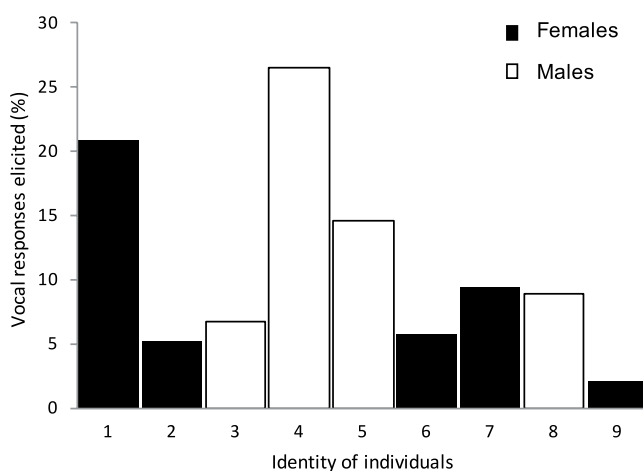


Figure 4. Vocal responses elicited by the different group members. Percentage of times that a call of a given bonobo elicited a vocal response from a conspecific. Individuals are ordered by decreasing age. Individuals 1 and 4 correspond to a mother-son pair.

| Fixed effects | χ^2 | d.f | P-value |
|--------------------------------|----------|-----|----------------|
| Social Affinity | 8.4 | 1 | 0.00375 |
| Sex composition of vocal dyads | 0.009 | 1 | 0.923 |
| Age difference | 1.795 | 1 | 0.180 |
| Kinship | 2.764 | 1 | 0.096 |

Table 3. Statistical analysis of the effect of socio-demographic determinants on vocal affinity between individuals. P-values were obtained independently for each predictor with likelihood-ratio tests comparing the fit of the full model with a reduced model lacking the fixed effect. Significant P-values are given in bold. Generalized Linear mixed model (GLMM) with a Gaussian error structure. Normality of the residuals was attained after applying a log-transformation.

have a significant social function similar to the vocal exchanges in monkeys and human conversations. Among nonhuman primates, spontaneous orderly vocal exchanges had only been described in monkey species (e.g. vervets, macaques, marmoset *sp.*, squirrel monkeys, Campbell’s monkeys) until we recently identified them in

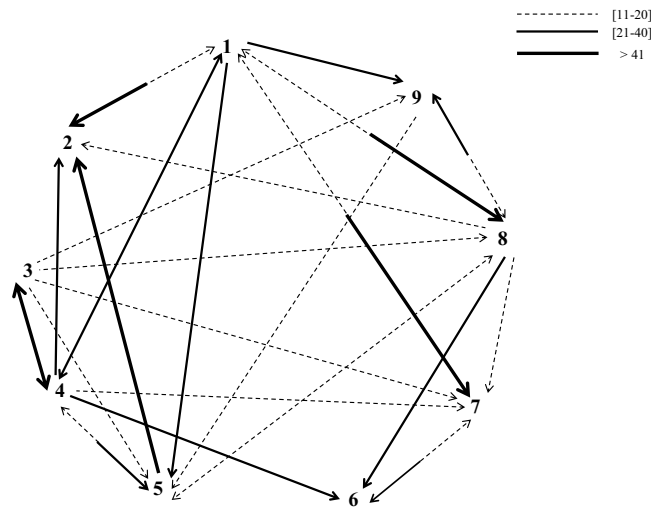


Figure 5. Sociogram showing vocal affinities. The arrows indicate who responds frequently to whom. Only dyadic interactions occurring more often than a threshold score of 10% are drawn here. The thickness of the arrows is related to the frequency of the interaction.

gorillas³¹. In orangutans, an experimental turn-taking setup showed that this ape has the cognitive control of vocal turn-taking^{64,65}. Overall, this suggests that spontaneous vocal exchanges may be spread among great ape species and that the lack of evidence is likely as a result of the lack of investigations. The only study conducted in chimpanzees concluded that they produce non-temporally organized call utterances or synchronous choruses³⁵. Their social behaviour might be less favourable to vocal exchanges than in bonobos. While bonobos are known for displaying pacifist behaviour and for their use of sexual behaviour in conflict appeasement, chimpanzees were found to resort more to force to deal with social conflicts⁵⁸. The absence of vocal turn-taking in chimpanzees, however, remains puzzling considering their complex social interactions and cooperative abilities. This finding in chimpanzees may be explained by the study conditions, which greatly differed from the bonobo and gorilla studies, and should perhaps be considered a preliminary result. First, the chimpanzees' study did not focus on peaceful contexts, which are more susceptible to induce organized vocal exchanges. Second, contrary to other ape studies, it took place in the wild where soft calls may have been undetected, as acknowledged by the author. Lastly callers were all adult male chimpanzees. Further research is thus necessary to investigate potential spontaneous cooperative vocal communication with vocal turn-taking behavior in captive chimpanzees. More generally, tracking the evolutionary pathway of proto-conversation within the primate lineage will require more comparative studies between socially-diverse species to discuss a potential convergent evolution or an ancestral origin of vocal cooperation.

Vocal sharing in bonobos had not been previously investigated. Among their graded vocal system^{59,60}, we identified distinctive acoustic 'variants' and found differences in their production rate among individuals. Focusing on these 'variants' in the context of vocal exchanges, only age difference explained the vocal sharing rate in bonobos. Vocal sharing was not influenced by their social affinities or by other individual characteristics (sex and kinship). This finding differs from the studies on Campbell's monkeys^{26,48,66}, chimpanzees^{37,54} and humans⁶⁷ for which social affinity predicts vocal sharing. It also differs from Japanese macaques for which vocal sharing is predicted by dominance ranks⁶⁸. The question of the functions of vocal sharing in vocal exchanges is still open. Sharing acoustic 'variants' may function as a 'vocal social badge' to advertise bonding within a group. This is the case with most monkey species e.g.^{26,53,66} and also with chimpanzees³⁷. Social systems may greatly influence the importance of signaling specific social bonds, such as hierarchical bonds in despotic societies⁵⁸, male-female affinities in monogamous groups⁵³ and female-female affinities in uni-male/multi-female groups⁴⁸. To the best of our knowledge, no studies had previously reported that age influences vocal sharing. We found in bonobos that the greater the age difference between two callers, the less likely that they shared acoustic 'variants'. We can thus hypothesize that juvenile bonobos display a 'youngster vocal badge' since they often play with peers (but not exclusively see⁶⁹), which may favor the development of social bonds during their long immaturity period. It is also interesting to note that an effect of age has been found on call exchange rate in gorillas³¹. The closer two gorillas are in age, the more likely they are to exchange grunts. These observations might thus suggest that age proximity plays an important role in great apes that are characterized by a long period of development and their longevity. At this stage, we cannot conclusively rule out alternative hypotheses, although these appear less plausible: (i) Vocal sharing might indeed result due to the phenomenon of maturation. Although a link between age and morphological changes in vocal apparatus has been shown in numerous species⁷⁰, it seems unlikely in bonobos since all age classes produced the recorded 'variants' in our focused group; (ii) Vocal sharing could be context-specific but in the present study all vocal exchanges were recorded in a relative standardized context (peaceful context); (iii) Vocal features could depend on the internal state of callers (review in mammals including primates⁷¹); (iv) Young individuals might share more similar emotions with their peers than with other age classes, which might explain the correlation found between age and vocal sharing rate in bonobos.

Interestingly, bonobos do not match the ‘variants’ during their vocal exchanges contrary to what is suggested for some monkey species (Diana monkeys⁷²; Japanese macaques⁷³) and in most of birds who display song matching⁴⁶. This is however similar to what was described in Campbell’s monkeys where individuals who share the same acoustic variants do not systematically respond to each other by matching their calls²⁶.

Lastly, we found that vocal interactions between individuals were affected by their social affinities. Indeed, bonobos did not randomly respond to any group members but preferentially to some specific conspecifics with whom they maintain close bonds. This finding is consistent with the evolutionary function of call exchanges or chorusing which is often linked to social bonding (e.g. many pair-living bird species⁷⁴, primates species: ring-tailed lemurs⁷⁵, squirrel monkeys^{14,76,77}, spider monkeys⁷⁸, rhesus macaques⁷⁹, siamangs¹⁷, gorillas⁸⁰ and chimpanzees³⁷). Interestingly, we observed an unbalanced vocal response rate between individuals (individual B systematically calls after individual A but the contrary is not true) that suggests that vocal exchanges may reflect an active search of building bonds. Hierarchical ranks may also influence the patterns of vocal exchanges, but more groups are however needed in order to test this effect. It is widely considered that socially-ruled communicative behaviour shared by humans and nonhuman primates may have been a crucial step in the coevolution of language and social life e.g.^{1,2,20}. In line with this, Dunbar²⁵ described nonhuman primate call exchanges as ‘vocal grooming’. When social groups became too large and too complex, limiting the possibility to interact physically with a lot of members, vocal exchanges, and then later language, could have become a way to maintain close bonds at distance.

In conclusion, this study conducted on a single captive group of bonobos shows that great apes, like monkeys, apply some simple conversational turn-taking rules and only social bonds determine the frequency of these vocal exchanges. Moreover, the sharing of vocal patterns appears to be essentially influenced by the age of bonobos, independent of their social bonds. These findings fill an important gap regarding the knowledge of vocal patterns along the primate phylogeny. Whether the main role of vocal exchanges is to reinforce social bonds between vocal interlocutors, advertise the existing ones or to convey more informative signals remains to be determined by conducting further playback experiments. We suggest that further comparative work, and the consideration of multiple sensory modalities in investigations of turn-taking behavior⁸¹ are now required to identify the key factors that have driven the emergence of the different components of human language.

Methods

Ethics Statement. All methods were carried out in accordance with the relevant French national guidelines and regulations, under the authorization no. C42-218-0901-38 SV 09 delivered by the «Direction Départementale de la Protection des Populations» committee, Préfecture du Rhône. This study only involves behavioural observations and spontaneous vocal recordings of animals in their natural social group and in their usual environment. No experimental protocols have been conducted in this study.

Study site and animals. A captive group of nine bonobos composed of four males (8 to 20 years old) and five females (5 to 44 years old) was observed at the zoological Park of La Vallée des Singes (Romagne, France). Wild bonobos live in fission-fusion social system^{82,83} and their community size greatly varies according to the studied field sites^{82,84,85}. The studied group composition reflects natural mixed subgroups observed in the wild. No studied individuals have been rescued or experienced bad treatment. The group was initially composed of five individuals (ind. #1, 3, 4, 5, and 9; Figs 4 and 5) who have lived together for at least four years and one year later, joined by four additional individuals (April-May 2011) before our data collection started. The study group was composed of the following pairs: three mother/offspring, one father/offspring, one half-brother, one nephew and one second-cousin. They lived in a large enriched building of 900 m² with access to a one-hectare wooded island depending on the meteorological conditions. Bonobos were fed eight times per day with fresh fruits and vegetables. Water was available *ad libitum*.

Data collection. Bonobos were monitored about six hours per day during 31 days from March to May 2012. Observations consisted in alternating focal group and scan sampling. Focal groups lasted 10 minutes each during which all vocal utterances were recorded (Sennheiser MKH70 ultra-directional microphone, Marantz PMD670 recorder 44.1 kHz sample rate/16 bit resolution, Sennheiser HD 25-1 II headphones) with the identity of the callers given in a handheld digital voice recorder (Yamaha Pocketrak2G). Between two given focal group samples, scan sampling was scored every 15 minutes to identify the closest neighbour of each group member and record their respective activity using the aforementioned voice recorder. We focused on vocal interactions produced within peaceful contexts such as resting and foraging periods, and peaceful social interactions like grooming sessions. Calls recorded in agonistic interactions (e.g. conflicts characterized by chasing, hitting and biting behaviours) or in response of external disturbance (e.g. arrival of caregivers) were thus excluded from the subsequent analyses. Indeed, turn-taking and vocalization spacing require calmness, control and attention toward the others while excitation or conflicts may lead to more overlap (e.g. in humans^{7,86}). The distance between individuals involved in these vocal exchanges were between less than 1 meter and about 8 meters.

Occurrence of ‘vocal responses’. The group vocalized in 55% of the focal group samples performed (70/127 ten-minute sequences). To describe the temporal organization of call production, we studied the distribution of all the inter-call durations from distinct identified consecutive callers²⁹. When there were more than two exchanged calls we focused on each of the consecutive calls: for example, in calls exchanged between the individuals ABC, the vocal dyads AB and BC were studied; between ABA, the vocal dyads AB and BA were studied (as in³¹). We are aware that such divisions in dyads are a debated question, so for the study of the matching of ‘variants’ we also report the analysis with only the first two callers (i.e. in a sequence between ABC, we only considered AB). We measured on spectrograms the time intervals between two consecutive calls from the offset of a call to

the onset of the next one (using PRAAT software version 5.3.56). A negative time interval meant that the two calls overlapped. In total, we computed 779 inter-call durations, which were used to assess the time window referring to, what could later be named a ‘vocal response’, defining the consecutive pairs of exchanged calls without overlap. To compare the proportion of calls eliciting a vocal response according to the identity of the emitters, we calculated the percentage for each individual of his/her total number of calls produced during the focal group samples that were followed by a consecutive caller within 2500 ms (i.e. response delay threshold, see Fig. 1).

Classification of calls. 310 exchanged calls were isolated from 192 vocal exchanges, a sub-sample of our data taken in chronological order of the recordings. These 310 calls were pre-classified into acoustic ‘variants’ based on the shape of fundamental frequency modulation pattern through visual inspections of spectrograms (see Fig. S1), as is classically done with birds⁸⁷, monkeys⁶⁶ and cetaceans^{88,89}. The bonobo vocal repertoire is described as highly graded, namely that there are no clear boundaries in the acoustic structure of the different call types (see the spectrograms in^{59,60}). In the present study, most of exchanged calls were soft and short as expected in peaceful context and mostly corresponded to “peep yelps” and “peeps” according to the bonobo repertoire in the literature. Peep yelps and peeps are both described in literature as presenting “very variable shapes”^{59,90}. Peep yelps and peeps can have a flat, ascending, descending or double frequency modulation. Peeps are also characterized by their short duration^{60,91} (<100 ms). Since we focused on call production within only one general peaceful context, we looked for more subtle call classification based on their frequency modulation patterns. We thus obtained 6 ‘variants’ (Fig. S2) from the visual inspection of spectrograms (with PRAAT software version 5.3.56, available from www.praat.org): **Variant A:** no frequency modulation, varied duration (n = 108 calls classified as variant A). **Variant B:** ascending frequency with an optional plateau at the end (start frequency < end frequency), varied duration (n = 79). **Variant C:** frequency modulation with an ascending and a descending slope (start frequency = end frequency), varied duration (n = 76). **Variant D** (opposite to variant 2): descending frequency modulation with an optional plateau at the beginning (start frequency > end frequency), varied duration (n = 33). **Variant E:** frequency modulation with an ascending and a descending slope (start frequency > end frequency), short duration (n = 9). **Variant F:** strong frequency modulation with two ascending and two descending slopes, long duration (n = 5). Sixty one percent of calls were classified for accuracy by a second observer (FL) and tested using the Cohen’s Kappa coefficient to ensure inter-observer reliability. We obtained a “very good” level of agreement ($\kappa = 0.85$).

For general information, we could classify 61.3% of the calls produced during vocal exchanges (N = 310 calls) in call types according to the bonobo repertoire⁵⁹ and found 51% of peep yelps, 32% of peeps, 6% of soft barks and less than 3% of barks, bark screams, whistles, whine whistles, pout moans, yelps, whistle barks and grunts). Bermejo & Omedes⁵⁹ described various call types (e.g. peep yelp, peep, soft bark, whistle) produced in peaceful contexts (foraging, grooming, resting, play incitation, play sequences). Our classification in ‘variants’ from A to F thus distinguishes the acoustic variations within call types. A call type can be found in different ‘variants’ (each of our ‘variants’ includes at least two distinct call types).

Acoustic measurements. Six classical acoustic parameters had been manually measured on the fundamental frequency with PRAAT software using the following settings: analysis window length 0.05 s, dynamic range 70 dB; spectrogram view range 0–8 kHz. In temporal and frequency domains we measured: call duration (d), start frequency (Fstart), maximum frequency (i.e. highest pitched frequency, Fmax), end frequency (Fend), ascending slope ((Fmax-Fstart)/duration) and descending slope ((Fmax-Fend)/duration). A Principal Component Analysis (PCA) followed by a Discriminant Function Analysis (DFA) was performed to test the relevance of our variants pre-classification¹⁶.

Socio-demographic determinants. For each possible dyad of bonobos (bidirectional data: N = 36) we calculated the ‘age difference’ in years and the ‘coefficient of relatedness’. We then calculated for each ‘variant’ a dyadic ‘vocal sharing’ score (n = 36) which was the differences in call rates of the two individuals forming a given dyad. Individual call rates were the number of times that each variant was uttered divided by the total number of calls produced by one individual, then multiplied by 100. For each dyad, the differences of call rate between all ‘variants’ were then averaged; we thus obtained a score of vocal dissimilarity that we subtracted from 100 to get a more intuitive score (i.e. the higher the score, the higher the acoustic similarity between two individuals). We also calculated, based on the identification of the closest neighbour of each individual from 487 scan samples, a dyadic ‘social bonding’ score quoted thereafter ‘social affinity’ (unidirectional data: N = 72) relying on the frequency of occurrence of peaceful spatial proximities. Individuals were considered in peaceful spatial proximity when they interacted positively (e.g. grooming) or rested or foraged at 5 meters or less from each other (see also⁹²). The score for dyad AB was calculated according to the formula: number of scans where B is the closest neighbour (and is in peaceful spatial proximity) of A * 100/total number of scans where A had been observed. Last the dyadic vocal response rate, quoted ‘vocal affinity’ thereafter (n = 72) was calculated for each dyad. For the dyad AB for instance, we recorded the percentage of times A got a vocal response from B out of the total A calls and *vice versa* (i.e. the higher the score, the higher the vocal affinity between two individuals). Unilateral dyads were considered here since reciprocity for vocal affinity was not observed within dyads, namely individual A can elicit more vocal response rate from individual B than B from A.

Statistical analyses. *Studied group.* Because the studied group resulted of the fusion of two former groups (composed of five and four individuals respectively) one year before our study (see above in Methods), we checked that the nine individuals truly formed a united group in terms of social and vocal behaviour. When necessary we performed generalized linear mixed models (GLM), which account for the repeated variables of identity (i.e. the same individual was involved in several vocal or social dyads). First no differences were found in the production

rate of the six ‘variants’ regarding the past membership of individuals (for each variant: Mann–Whitney tests: z ranged from 0 to 1.715 and P from 0.09 to 1). Second no social preferences, defined by dyadic spatial proximities (see definition above in Methods) have been highlighted between the individuals who initially belonged to the same subgroup ($n = 72$; GLM: $\chi^2 = 0.55$, $df = 1$, $P = 0.456$). Lastly, vocal affinity (i.e. dyadic vocal response rate, see above in Methods) between exchanging callers was not influenced by the past membership of individuals ($n = 72$; GLM: $\chi^2 = 1.3$, $df = 1$, $P = 0.254$; see also Fig. 5). These results showed that the nine individuals indeed form a cohesive group whatever their past membership and can be considered as belonging to a united social group at the time of the study.

Socio-demographic determinants of vocal sharing and vocal affinity. We performed GLM (accounting for the repeated variables of identity, namely the same individual was involved in several vocal exchanges) to investigate the socio-demographic factors that influence vocal sharing on the one hand, and on the other hand the vocal affinity (function lmer in lme4 package of R software).

First we tested the effects of social affinity, vocal affinity and individual characteristics (age, sex and kinship) on vocal sharing rate (fixed factors: social affinity, vocal affinity, sex composition of vocal dyads, age difference and kinship; random effects: the identity of the first caller and the identity of the second caller). Second we tested the effects of social affinity and individual characteristics (age, sex and kinship) on vocal affinity (fixed factors: social affinity, vocal sharing, sex composition within vocal dyads, age difference and kinship; random effects: the identity of the first caller and the identity of the second caller). In this model, vocal sharing rate and social affinity*vocal sharing have been excluded from the tested fixed factors since the previous model (see above) revealed that vocal affinity was not a significant predictor of the vocal sharing.

For all analyses, the variables of interests were centered and normalized (i.e. transformed into z-scores) to insure correct weighting since our parameters had different units and then log-transformed to meet the homoscedasticity assumption. We used a GLM with a Gaussian error structure since the normality of the residuals was attained after data transformation. Collinearity was checked for all fixed factors using variance of inflation factor (function vif in car R package). For each predictor we obtained a very low vif (< 1.45 for all predictor for vocal sharing analysis, and < 1.24 for vocal analysis) indicating that the used predictors were not correlated. The P-values were obtained with likelihood-ratio tests comparing the fit of the full model with a reduced model lacking the fixed effect.

All the statistical tests were performed using R (version 3.4.0) software with a significance level set at $\alpha = 0.05$.

References

- Levinson, S. C. Turn-taking in human communication – Origins and implications for language processing. *TICS* **20** (2016).
- Snowdon, C. T. Learning from monkey “talk”. *Science* **355**, 1120 (2017).
- Coye, C., Townsend, S. & Lemasson, A. From animal communication to linguistics and back: insight from compositional abilities in monkeys and birds in *Origins of human language: continuities and splits with nonhuman primates* (eds L. J. Boë, J. Fagot, P. Perrier, & J. L. Schwartz) 187–232 (Peter Lang Publisher, 2017).
- 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**, 20180598 (2018).
- Grice, H. P. Logic and conversation in *Syntax and semantics* Vol. 3 (eds P. Cole & J. Morgan) 43–58 (New York: Academic Press, 1975).
- Levinson, S. C. *Pragmatics* (Cambridge University Press, 1983).
- Sacks, H., Schegloff, E. A. & Jefferson, G. Simplest systematics for organization of turn-taking for conversation. *Language* **50**, 696–735 (1974).
- Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc Natl Acad Sci USA* **106**, 10587–10592 (2009).
- Yoshida, S. & Okanoya, K. Evolution of turn-taking: A bio-cognitive perspective. **12**, 153–165 (2005).
- Méndez-Cárdenas, M. G. & Zimmermann, E. Duetting—A mechanism to strengthen pair bonds in a dispersed pair-living primate (*Lepilemur edwardsi*)? *Am. J. Phys. Anthropol.* **139**, 523–532 (2009).
- Chow, C. P., Mitchell, J. F. & Miller, C. T. Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. R. Soc. B* **282**, 20150069 (2015).
- Snowdon, C. T. & Cleveland, J. “Conversations” among pygmy marmosets. *Am J Primatol* **7**, 15–20 (1984).
- Takahashi, D. Y., Narayanan, D. Z. & Ghazanfar, A. A. Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr. Biol.* **23**, 2162–2168 (2013).
- Biben, M., Symmes, D. & Masataka, N. Temporal and structural analysis of affiliative vocal exchanges in squirrel monkeys (*Saimiri sciureus*). *Behaviour* **98**, 259–273 (1986).
- Müller, A. E. & Anzenberger, G. Duetting in the Titi monkey *Callicebus cupreus*: structure, pair specificity and development of duets. *Folia Primatol.*, 104–115 (2002).
- Lemasson, A. *et al.* Youngsters do not pay attention to conversational rules: is this so for nonhuman primates? *Sci. Rep.* **1**, 22 (2011).
- Geissmann, T. & Orgeldinger, M. The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim. Behav.* **60**, 805–809 (2000).
- Haimoff, E. H. Video analysis of siamang (*Hylobates syndactylus*) songs. *Behaviour* **76**, 128–151 (1981).
- Levinson, S. C. On the human “interaction engine” In *Roots of human sociality: Culture, cognition and interaction* (eds N. J. Enfield & S. C. Levinson) 39–69 (Berg, 2006).
- Tomasello, M. *Origins of human communication* (MIT Press, 2008).
- Clark, H. H. *Using Language* (Cambridge University Press, 1996).
- Arlet, M., Jubin, R., Masataka, N. & Lemasson, A. Grooming-at-a-distance by exchanging calls in non-human primates. *Biol. Lett.* **11** (2015).
- Miller, C. T., Mandel, K. & Wang, X. The communicative content of the common marmoset phee call during antiphonal calling. *Am. J. Primatol.* **72**, 974–980 (2010).
- Schegloff, E. A. Overlapping talk and the organization of turn-taking for conversation. *Lang. Soc.* **29**, 1–63 (2000).
- Dunbar, R. I. M. *Grooming, gossip and the evolution of language* (Faber and Faber, 1998).
- Lemasson, A. & Hausberger, M. Patterns of vocal sharing and social dynamics in a captive group of Campbell’s monkeys (*Cercopithecus campbelli campbelli*). *J. Comp. Psychol.* **118**, 347–359 (2004).
- Lemasson, A., Gandon, E. & Hausberger, M. Attention to elders’ voice in non-human primates. *Biol. Lett.* **6**, 325–328 (2010).

28. Lemasson, A. *et al.* Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291 (2013).
29. 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).
30. Digweed, S. M., Fedigan, L. M. & Rendall, D. Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus Capucinus*). *Am. J. Primatol.* **69**, 829–835 (2007).
31. Lemasson, A., Pereira, H. & Levréro, F. Social basis of vocal interactions in western lowland gorillas (*Gorilla g. gorilla*). *J. Comp. Psychol.* (2018).
32. 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).
33. Sugiura, H. Adjustment of temporal call usage during vocal exchange of coo calls in Japanese macaques. *Ethology* **113**, 528–533 (2006).
34. Henry, L., Barbu, S., Lemasson, A. & Hausberger, M. Dialects in animals: Evidence, development and potential functions. *Anim. Behav. Cogn.* **2**, 132–155 (2015).
35. Arcadi, A. C. Vocal responsiveness in male wild chimpanzees: implications for the evolution of language. *J. Hum. Evol.* **39**, 205–223 (2000).
36. Call, J. & Tomasello, M. Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* **12**, 187–192 (2008).
37. Fedurek, P., Machanda, Z. P., Schel, A. M. & Slocombe, K. E. Pant hoot chorusing and social bonds in male chimpanzees. *Anim. Behav.* **86**, 189–196 (2013).
38. Mitani, J. C. & Brandt, K. L. Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology* **96**, 233–252 (1994).
39. Soltis, J., Leong, K. & Savage, A. African elephant vocal communication I: antiphonal calling behaviour among affiliated females. *Anim. Behav.* **70**, 579–587 (2005).
40. Janik, V. M. & Slater, P. J. B. The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11 (2000).
41. 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).
42. Vernes, S. C. What bats have to say about speech and language. *Psychon. Bull. Rev.* **24**, 111–117 (2017).
43. Yosida, S., Kobayashi, K. I., Ikebuchi, M., Ozaki, R. & Okanoya, K. Antiphonal vocalization of a subterranean rodent, the naked mole-rat (*Heterocephalus glaber*). *Ethology* **113**, 703–710 (2007).
44. Rendall, D., Seyfarth, R. M., Cheney, D. L. & Owren, M. J. The meaning and function of grunt variants in baboons. *Anim. Behav.* **57**, 583–592 (1999).
45. Hafen, T., Neveu, H., Rumpler, Y., Wilden, I. & Zimmermann, E. Acoustically dimorphic advertisement calls separate morphologically and genetically homogenous populations of the grey mouse lemur (*Microcebus murinus*). *Folia Primatol.* **69**, 342–356 (2000).
46. Snowdon, C. T. & Hausberger, M. *Social influences on vocal development* (Cambridge University Press, 1997).
47. Rukstalis, M., Fite, J. E. & French, J. A. Social change affects vocal structure in a callitrichid primate (*Callithrix kuhlii*). *Ethology* **109**, 1–14 (2003).
48. Lemasson, A., Ouattara, K., Petit, E. J. & Zuberbühler, K. Social learning of vocal structure in a nonhuman primate? *BMC Evol. Biol.* **11**, 362 (2011).
49. Geissmann, T. Mate change enhances duetting activity in the siamang gibbon (*Hylobates syndactylus*). *Behaviour* **17–27** (1986).
50. Geissmann, T. Taxonomy and evolution of gibbons. *Evol Anthropol: Issues, News, and Reviews* **11**, 28–31 (2002).
51. Watson, S. K. *et al.* Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* **25**, 495–499 (2015).
52. Elowson, A. M. & Snowdon, C. T. Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Anim. Behav.* **47**, 1267–1277 (1994).
53. Snowdon, C. T. & Elowson, A. M. Pygmy marmosets modify call structure when paired. *Ethology* **105**, 893–908 (1999).
54. Mitani, J. C. & Gros-Louis, J. Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour* **135**, 1041–1064 (1998).
55. Pardo, J. Measuring phonetic convergence in speech production. *Front. Psychol.* **4**, 559 (2013).
56. Pardo, J. S. On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* **119**, 2382–2393 (2006).
57. White, F. J. Comparative socio-ecology of *Pan paniscus* In *Great ape societies* (eds W. C. McGrew, L. F. Marchant & T. Nishida) **29–41** (Cambridge University Press, 1996).
58. Gruber, T. & Clay, Z. A comparison between bonobos and chimpanzees: A review and update. *Evol Anthropol: Issues, News, and Reviews* **25**, 239–252 (2016).
59. Bermejo, M. & Omedes, A. Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu. *Folia Primatol.* **70**, 328–357 (1999).
60. De Waal, F. B. M. The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour* **106**, 183–251 (1988).
61. Fedurek, P., Donnellan, E. & Slocombe, K. E. Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behav. Ecol. Sociobiol.* **68**, 1345–1355 (2014).
62. Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G. & Seyfarth, R. M. Call combinations, vocal exchanges and interparty movement in wild bonobos. *Anim. Behav.* **122**, 109–116 (2016).
63. Mitani, J. C. & Nishida, T. Contexts and social correlates of long-distance calling by male chimpanzees. *Anim Behav* **45**, 735–746 (1993).
64. 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).
65. Lameira, A. R. *et al.* Orangutan (*Pongo spp.*) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *J. Acoust. Soc. Am.* **134**, 2326–2335 (2013).
66. Lemasson, A., Hausberger, M. & Zuberbühler, K. Socially meaningful vocal plasticity in adult Campbell's monkeys (*Cercopithecus campbelli*). *J. Comp. Psychol.* **119**, 220–229 (2005).
67. Giles, H. & Coupland, N. *Language: Contexts and consequences, mapping social psychology* (Thomson Brooks/Cole Publishing Co, 1991).
68. Lemasson, A., Jubin, R., Masataka, N. & Arlet, M. Copying hierarchical leaders' voices? Acoustic plasticity in female Japanese macaques. *Sci. Rep.* **6** (2016).
69. Behncke, I. Play in the Peter Pan ape. *Curr. Biol.* **25**, R24–R27 (2015).
70. Ey, E., Pfefferle, D. & Fischer, J. Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. *Primates*, 253–267 (2007).
71. Briefer, E. F. Vocal expression of emotions in mammals: mechanisms of production and evidence. *J. Zool.* **288**, 1–20 (2012).
72. Candiotti, A., Zuberbühler, K. & Lemasson, A. Convergence and divergence in Diana monkey vocalizations. *Biol. Lett.* **8**, 382–385 (2012).
73. Sugiura, H. Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim. Behav.* **55**, 673–687 (1998).
74. Hall, M. L. A review of hypotheses for the functions of avian duetting. *Behav. Ecol. Sociobiol.* **55**, 415–430 (2004).

75. Oda, R. Effects of contextual and social variables on contact call production in free-ranging ringtailed lemurs (*Lemur catta*). *Int. J. Primatol.* **17**, 191–205 (1996).
76. Soltis, J., Bernhards, D., Donkin, H. & Newman, J. D. Squirrel monkey chuck call: vocal response to playback chucks based on acoustic structure and affiliative relationship with the caller. *Am. J. Primatol.* **57**, 119–130 (2002).
77. Biben, M. Recognition of order effects in squirrel monkey antiphonal call sequences. *Am. J. Primatol.* **29**, 109–124 (1993).
78. Ramos-Fernández, G. Vocal communication in a fission–fusion society: do spider monkeys stay in touch with close associates? *Int. J. Primatol.* **26**, 1077–1092 (2005).
79. Hauser, M. D. A mechanism guiding conversational turn-taking in vervet monkeys and rhesus macaques in *Topics in Primatology Vol. 1: Human origins* (eds T. Nishida *et al.*) 235–248 (University of Tokyo Press, 1992).
80. Harcourt, A. H. & Stewart, K. J. Function and meaning of wild gorilla ‘close’ calls. *Correlations with rank and relatedness. Behaviour* **133**, 827–845 (1996).
81. Fröhlich, M. Taking turns across channels: conversation-analytic tools in animal communication. *Neurosci. Biobehav. R.* **80**, 201–209 (2017).
82. White, F. J. Party composition and dynamics in *Pan paniscus*. *Int. J. Primatol.* **9**, 179–193 (1988).
83. Kano, T. The social group of pygmy chimpanzees (*Pan paniscus*) of Wamba. *Primates* **23**, 171–188 (1982).
84. Kano, T. *The last ape: pygmy chimpanzee behavior and ecology* (Stanford University Press, 1992).
85. Furuichi, T. Factors underlying party size differences between chimpanzees and bonobos: a review and hypotheses for future study. *Primates* **50**, 197–209 (2009).
86. 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).
87. Hausberger, M., Richard-Yris, M. A., Henry, L., Lepage, L. & Schmidt, I. Song sharing reflects the social organization in a captive group of European starling (*Sturnus vulgaris*). *J. Comp. Psychol.* **109**, 222–241 (1995).
88. Kremers, D., Briseño Jaramillo, M., Böye, M., Lemasson, A. & Hausberger, M. Do dolphins rehearse show-stimuli when at rest? Delayed matching of auditory memory. *Front. Psychol.* **2**, 386 (2011).
89. Reiss, D. & McCowan, B. Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *J. Comp. Psychol.* **107**, 301–312 (1993).
90. Clay, Z., Archbold, J. & Zuberbühler, K. Functional flexibility in wild bonobo vocal behaviour. *PeerJ* **3**, e1124 (2015).
91. Clay, Z. & Zuberbühler, K. Food-associated calling sequences in bonobos. *Anim. Behav.* **77**, 1387–1396 (2009).
92. Parish, A. R. Female relationships in bonobos (*Pan paniscus*). *Hu. Nat.* **7**, 61–96 (1996).

Acknowledgements

We are very grateful to the bonobo caregivers from La Vallée des Singes, Romagne (France) for their valuable assistance. We acknowledge Maxime Blanchon for his contribution and warmly thank Nicolas Mathevon for his valuable comments on the earlier versions of this manuscript, Marie Fernandez for writing the routine on R Software, David Reby for his valuable comments on the previous version and, Mark van Niekerk and David Reby for english corrections. This work was financially supported by the Centre National de la Recherche Scientifique “GDR Ethologie 2822”, the University of Lyon/Saint-Etienne and the University of Rennes 1.

Author Contributions

F.L. and A.L. conceived the project. J.F. performed the acoustic recordings with the assistance of J.P.G. B.N. analyzed the time interval measurements. S.T. performed all other acoustic measurements. F.L., S.T. and A.L. performed the statistical analyses. F.L., S.T. and A.L. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36024-9>.

Competing Interests: The authors declare no competing interests.

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) 2019