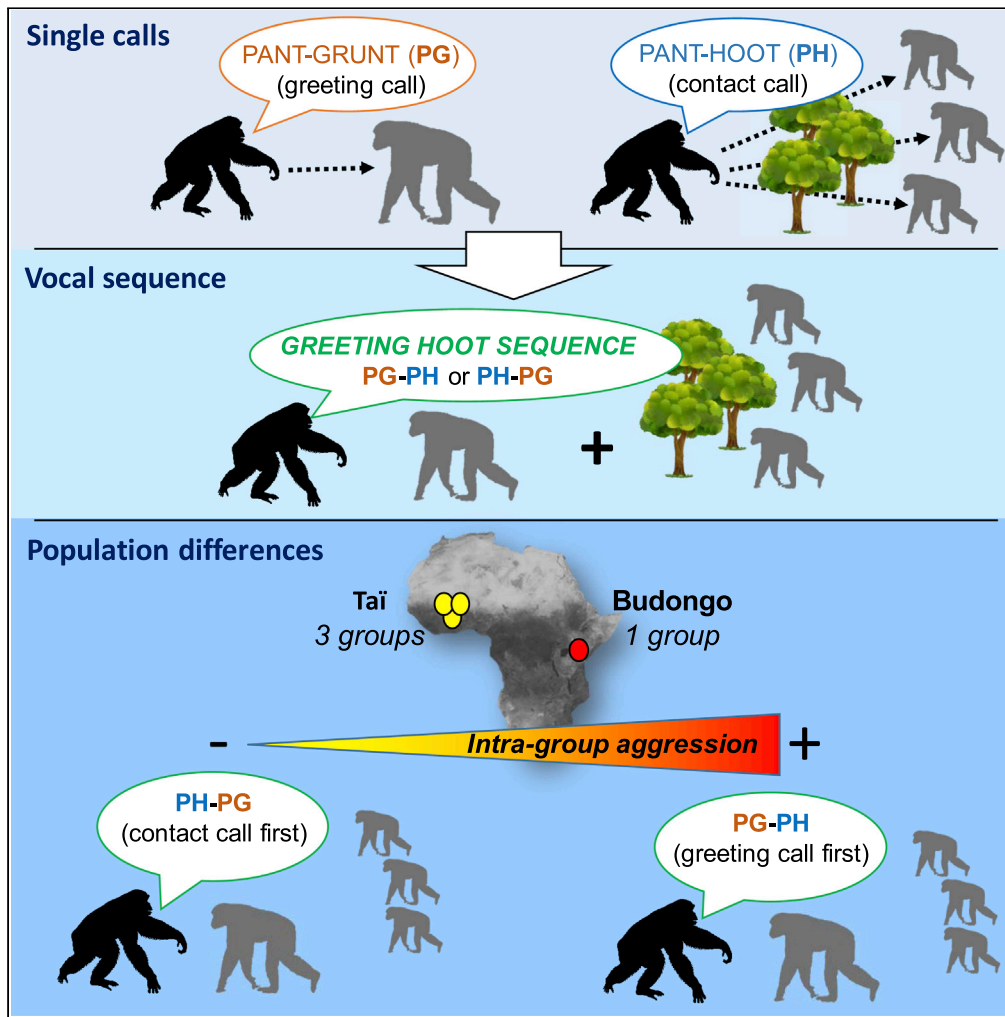


Article

Population-specific call order in chimpanzee greeting vocal sequences



Cédric Girard-Buttoz, Tatiana Bortolato, Marion Laporte, Mathilde Grampp, Klaus Zuberbühler, Roman M. Wittig, Catherine Crockford

cedric.girardbuttoz@isc.cnrs.fr (C.G.-B.)
crockford@isc.cnrs.fr (C.C.)

Highlights

Chimpanzees combine pant-grunt and pant-hoot calls into a greeting hoot sequence

Call-order of these greeting and contact calls is population specific

Pant-grunt is uttered first in the population with higher in-group aggressions

Vocal usage learning may lead to these population differences in sequence structure

Girard-Buttoz et al., iScience 25, 104851 September 16, 2022 © 2022 The Author(s). <https://doi.org/10.1016/j.isci.2022.104851>



Article

Population-specific call order in chimpanzee greeting vocal sequences

Cédric Girard-Buttoz,^{1,2,3,9,*} Tatiana Bortolato,^{1,2,3} Marion Laporte,^{4,5} Mathilde Grampp,^{1,2,3} Klaus Zuberbühler,^{6,7,8} Roman M. Wittig,^{1,2,3} and Catherine Crockford^{1,2,3,*}

SUMMARY

Primates rarely learn new vocalizations, but they can learn to use their vocalizations in different contexts. Such “vocal usage learning,” particularly in vocal sequences, is a hallmark of human language, but remains understudied in non-human primates. We assess usage learning in four wild chimpanzee communities of Taï and Budongo Forests by investigating population differences in call ordering of a greeting vocal sequence. Whilst in all groups, these sequences consisted of pant-hoots (long-distance contact call) and pant-grunts (short-distance submissive call), the order of the two calls differed across populations. Taï chimpanzees consistently commenced greetings with pant-hoots, whereas Budongo chimpanzees started with pant-grunts. We discuss different hypotheses to explain this pattern and conclude that higher intra-group aggression in Budongo may have led to a local pattern of individuals signaling submission first. This highlights how within-species variation in social dynamics may lead to flexibility in call order production, possibly acquired via usage learning.

INTRODUCTION

Despite prolific research efforts, the evolutionary origins of human language capacities remain a largely unsolved riddle. Determining the origins of human vocal learning capacities is a crucial component in resolving this problem (Jackendoff, 1999; Jarvis, 2019; Wilbrecht and Nottebohm, 2003), with research on non-human animals (hereafter “animals”) playing a key role (reviewed in Egnor and Hauser, 2004; Nowicki and Searcy, 2014; Wilbrecht and Nottebohm, 2003). Vocal learning consists of at least three different mechanisms (production, usage, comprehension), and overall refers to the ability to modify the vocal output in response to social or individual experience, typically from interacting with other group members (Janik and Slater, 2000; Sewall et al., 2016).

In terms of vocal output, two processes have been distinguished (Janik and Slater, 2000; Sewall et al., 2016). First, *production learning* refers to the ability to produce novel signals or signal variants in response to social experience and auditory feedback. Second, *contextual or usage learning* refers to the ability to use existing signals flexibly in different situations. Usage learning can operate via two distinct processes. First, individuals can learn socially by observing conspecifics and imitating their pattern of vocal production in certain contexts (Hollén and Radford, 2009). Alternatively, individuals can learn individually via trial and error by adjusting their vocal production over time to optimize the outcome (Lattenkamp et al., 2018).

Production learning is common in songbirds (reviewed in Sewall et al., 2016; Wilbrecht and Nottebohm, 2003) but also found in non-singing birds (Wright, 1996), bats (reviewed in Vernes and Wilkinson, 2020), and cetaceans (reviewed in Sewall et al., 2016). For primates, there is a puzzling contrast between the advanced vocal production learning capacities of humans, and the very limited capacities of non-human primates (hereafter primates, Egnor and Hauser, 2004; Fedurek and Slocombe, 2011). In primates, vocal repertoires consist of structurally mostly fixed and genetically predetermined call types (reviewed in Fedurek and Slocombe, 2011; Fischer and Hammerschmidt, 2020), with individuals possessing only limited flexibility and control over the acoustic structure of the call types within the repertoires (e.g., macaques, Sugiura, 1998; marmosets, Elowson and Snowdon, 1994; Gultekin et al., 2021; Snowdon and Elowson, 1999; baboons, Fischer et al., 2020; gibbons, Geissmann, 2002, 1986; orangutans, Lameira et al., 2022,

¹The Ape Social Mind Lab, Institut des Sciences Cognitives Marc Jeannerod, UMR 5229, CNRS, 67 Boulevard Pinel, Bron, Lyon 69675 France

²Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany

³Taï Chimpanzee Project, Centre Suisse de Recherches Scientifiques, Abidjan, Ivory Coast

⁴Histoire naturelle de l’Homme préhistorique, UMR 7194, PaleoFED, Muséum National d’Histoire Naturelle, 17 place du Trocadéro et du 11 Novembre, 75116 Paris, France

⁵Institut des Sciences du Calcul et des Données, Sorbonne Université, Paris, France

⁶Université de Neuchâtel, Institut de Biologie, Cognition Compare, Neuchâtel, Switzerland

⁷School of Psychology and Neuroscience, University of St Andrews, St Andrews, Scotland

⁸Budongo Conservation Field Station, Masindi, Uganda

⁹Lead contact

*Correspondence: cedric.girardbuttoz@isc.cnrs.fr (C.G.-B.), crockford@isc.cnrs.fr (C.C.)
<https://doi.org/10.1016/j.isci.2022.104851>



and chimpanzees, [Crockford et al., 2004](#); [Marshall et al., 1999](#); [Mitani et al., 1999, 1992](#); [Watson et al., 2015](#); reviewed in [Egnor and Hauser, 2004](#); [Fischer and Hammerschmidt, 2020](#)).

Although humans (unlike other primates) are undoubtedly capable of imitating almost any arbitrary sound, the key capacity of speech arguably concerns the ability **to recombine** components of a finite phonetic repertoire into longer sequences. The human phonetic repertoire consists of around 100 sounds, although the phonemic repertoire, i.e., how people use and interpret the sounds used in their own language is often considerably smaller (15-80 phonemes, [International Phonetic Association, 2022](#); [Moran et al., 2012](#)). In primates, the number of calls per repertoire also varies largely, albeit across species, but the number of call types reported per repertoire overlaps with the number of sounds used in at least certain human languages (i.e., 2-38 call-types, [McComb and Semple, 2005](#)). However, it is important to note that call types in primates are often graded which makes reliable estimates imprecise or just plainly impossible. Highly relevant is that humans use their phoneme repertoire to generate large (and in principle infinite) numbers of sound sequences by recombining syllables, the smallest sound units of speech, into morphemes, words, phrases, sentences, and dialogues.

In contrast, most non-human primates are thought to produce few, if any call sequences, although with some notable exceptions (reviewed in [Girard-Buttoz et al., 2022](#)). For instance, studies in marmosets highlighted the use of several vocal sequences in this species ([Bezerra and Souto, 2008](#)) and suggested that these sequences may be subject to usage learning ([Gultekin et al., 2021](#)). Also, chimpanzees have been shown to produce a large range of call sequences ([Crockford et al. 2005](#); [Girard-Buttoz et al., 2022](#)), although little is yet known about their semantic content. Establishing whether chimpanzees and other primates have an ability to learn the use of vocal sequences is thus of direct relevance to the study of the origins of language. More specifically, it is important to establish if the use of vocal sequences within a non-human primate species is innately fixed or can be flexibly adjusted, which would indicate that vocal sequences can be learned either socially, or individually via trial and error. Chimpanzees are an interesting species to examine as vocal sequences increase in structural complexity across ontogeny ([Bortolato et al., in revision](#)). Also, chimpanzees show some flexibility with regard to the production of vocal sequences and demonstrate the ability to learn vocal sequence variants. In a group of captive chimpanzees, male added voiceless raspberries to the end of his pant-hoot sequences ([Marshall et al., 1999](#)), which generated a new variant subsequently adopted by other males ([Marshall et al., 1999](#)).

Learning to recombine existing calls into sequences has been described as a form of *usage learning* as no acoustic modification of the existing sound set is required ([Janik and Slater, 2000](#); [Vernes et al., 2021](#)). There is, however, little knowledge about the specific mechanisms through which animals learn to use calls in sequences, and how widespread this form of learning is in the animal kingdom ([Vernes et al., 2021](#)). Usage learning of vocal sequences has been reported in songbirds, which can learn to modify the temporal arrangement of signals ([Hultsch and Todt, 1989](#); [Podos et al., 1999](#)). However, most studies on usage learning in non-singing animals, including primates, focus on the ability of individuals to learn, either socially or individually, the use of a specific call type in a certain context (e.g., [Hollén et al., 2008](#); [Koda et al., 2007](#); [Seyfarth and Cheney, 1980](#); [Shapiro et al., 2004, 2004](#), reviewed in [Hollén and Radford, 2009](#); [Janik and Slater, 1997](#); [Seyfarth and Cheney, 2010](#)), with only few studies have examined usage learning at the vocal sequence level ([Janik and Slater, 2000](#); [Sewall et al., 2016](#); [Vernes et al., 2021](#)). Furthermore, the socioecological drivers leading to variation in the ordering of calls within a sequence are understudied ([Vernes et al., 2021](#)).

Animal communication functions to convey information to others to modify their current and future behavior so that it benefits the signaler and often also the recipient ([Laidre and Johnstone, 2013](#); [Smit, 1977](#)). One established benefit is to prevent escalated aggressions between signalers and recipients ([Laidre and Johnstone, 2013](#)), such as by revealing their own or group identity to conspecifics ([Crockford et al., 2004](#)), or by signaling dominance ([Neumann et al., 2010](#); [Santos et al., 2011](#)) or subordination ([Colmenares et al., 2000](#); [East et al., 1993](#); [Fedurek et al., 2021](#)). Individual identity is often encoded in the acoustic features of calls (e.g., bats, [Boughman, 1997](#); chimpanzees, [Fedurek et al., 2016](#); dolphins [Janik et al., 2006](#); and meerkats, [Townsend et al., 2010](#)), which provides important additional social information in species that live in individualized groups with social bonds, kin or dominance relations ([Fedurek et al., 2019](#); [Setchell and Wickings, 2005](#)). However, conveying identity is not equally important in all social situations (e.g. [Coye et al., 2022](#)).

As vocal learning has mainly been observed in relation to animal song, it has been interpreted as being driven by sexual selection. However, it has been proposed that within-species variation in social dynamics

may also favor vocal flexibility and promote the evolution of vocal learning, in particular of social calls (Sewall et al., 2016; Vernes and Wilkinson, 2020). We name this idea the “social pressure hypothesis” whereby as social dynamics shift within a species, individuals may gain benefits from using calls in specific contexts or from combining call types in a certain order, hence causing selective pressure for the evolution of vocal learning. This hypothesis predicts that species or social groups, where social flexibility in the social system requires flexible adjustment of the vocal output beyond the species-specific innate vocal system, will be more prone to modifying their vocal production through learning in order to optimize their fitness outcome. Such a flexible system allows to adapt vocalizations to accommodate social threats that may change over time depending on temporal variation in social group dynamics. This hypothesis encompasses a large range of social pressures not limited to aggression but could equally apply to other social challenges such as coordination dynamics. For our study, we tested this hypothesis by using naturally occurring vocal behavior during situations of variable social pressure, specifically dyadic encounters in forest chimpanzees, which have considerable potential for escalated aggression. Depending on the risk of receiving aggression during an encounter, individuals may signal submission first before pursuing other communication needs, such as signaling identity.

Here, we addressed the “social pressure hypothesis” by studying which signaling function callers prioritize during social encounters, socially directed signals of submission (i.e., pant grunts), or often socially undirected but individually distinct long-distance contact calls (i.e. pant hoots). To this end, we compared two populations of chimpanzees with documented differences in intra-group and inter-group aggression (Taï in Ivory Coast and Budongo in Uganda, Townsend et al., 2007; Wilson et al., 2014, see below).

Pant-hoots can themselves be sequences of one to four distinct vocal units, which carry individual and group identity markers (Crockford et al., 2004; Desai et al., 2021; Marshall et al., 1999; Mitani et al., 1996, 1999). Pant-hoot sequences are produced in a variety of contexts but mainly serve in inter-party vocal exchanges, to establish and maintain contact (Eckhardt et al., 2015) and recruit individuals in other parties (Crockford, 2019; Kalan et al., 2015; Notman and Rendall, 2005). Chimpanzees also produce a greeting vocalization, the pant-grunt (PG), during approaches to dominant individuals (Goodall, 1986; Laporte and Zuberbühler, 2010; Wittig and Boesch, 2003). Emitting PG may reduce the chances of signallers receiving aggression from dominants (Fedurek et al., 2021). Chimpanzees combine part of the pant-hoot sequence, the build-up phase, consisting of panted hoots (PH, Notman and Rendall, 2005) and PG vocalization into a sequence, which we call the greeting hoot sequence and which may occur predominantly in fusion contexts (Crockford and Boesch, 2005).

In the current study, we assessed whether the order in which PH and PG appeared in the greeting hoot sequence (either PH + PG, Figure 1A, or PG + PH, Figure 1B) differed between individuals of three neighboring social groups within the same population and across two populations (Taï in Ivory Coast and Budongo in Uganda). We chose these communities since, as in all chimpanzee populations studied so far (Wilson et al., 2014), they all have a high degree of inter-group competition such that chimpanzees sometimes kill individuals of other communities. However, the degree of intra-group competition varies across populations in this study. Intra-group killings are relatively frequent in the Sonso community of Budongo Forest (13 intra-group killings observed in 24 years of observation) but were never reported in Taï over 15-40 years of observation per community, across all four communities (Wilson et al., 2014). We hypothesized that the production of PH or PG first in the greeting hoot sequence will be influenced by the degree of intra- and inter-group competition. Specifically, for a species such as chimpanzees, which face high levels of intra- and inter-group competition, and which live in low-visibility environments, vocal signals emitted at a safe physical distance from conspecifics may be crucial in limiting aggression when encountering community members not seen for some hours or days. Within species, populations experiencing different levels of resource competition may experience different exposure to aggression. In communities facing high levels of inter-group competition, producing PH first might be crucial to be recognized as an in-group member by community members to prevent receiving injurious aggressions (Prediction P1). If the degree of intra-group competition is very high, however, producing PG first to show submission might reduce injurious aggressions from higher ranking individuals (Prediction P2).

As a first step, we investigated the contexts of production of greeting hoot sequences, compared to the component parts, PH and PG. We also tested whether the contexts of production were similar for greeting hoot sequences with PH first compared to those with PG first. We conducted the latter analysis to rule out that potential

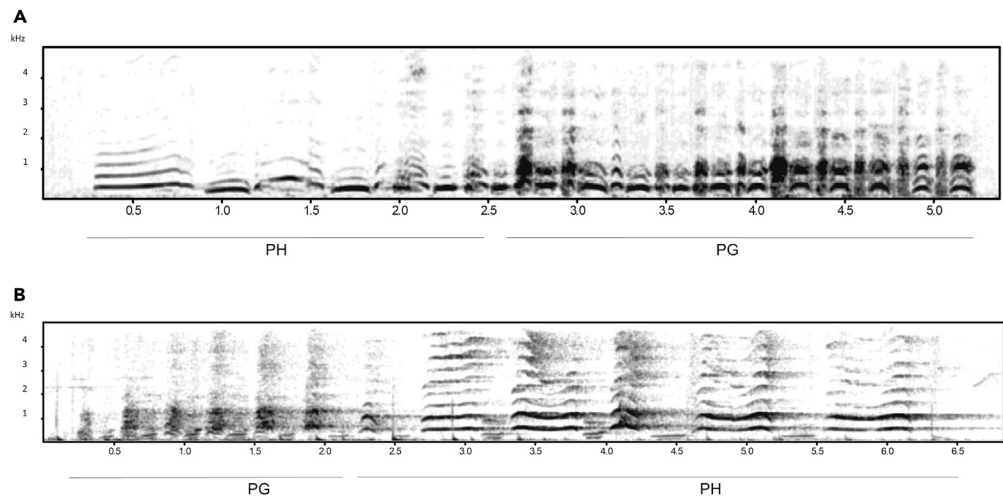


Figure 1. Greeting hoot sequence variants: A) pant hoot first, B) pant grunt first

Pant grunts (PG) are sequences of repeated grunts with a voiced inhalation between each grunt. Pant hoots (PH) are sequences of repeated hoots with a voiced inhalation between each hoot and correspond to the build-up phase of the pant-hoot sequence without the introduction or the climax. PH and PG can be combined to form the greeting hoot sequence (GH). Different call types are defined as being combined into a sequence when there is <1 s gap between calls. In the vast majority of cases, this gap is considerably shorter, as shown here. Note that only panting grunts were included in our analyses as they are only emitted in greeting contexts, in contrast to un-panted grunts that can be given in both greeting and food contexts.

differences in call ordering arise from differences in the context of call production between the two study populations. Secondly, we tested if, as predicted, chimpanzees in Budongo, likely exposed to higher degree of intra-group competition, utter PG first in the greeting hoot, whereas chimpanzees from Tai utter PH first. Finally, we assessed if potential differences in the ordering of PH and PG in the greeting hoot sequence were driven by long-term population-wide levels of intra- and inter-group threat or by the threat of receiving aggression experienced by each community member (i.e., the rates of contact aggression received).

RESULTS

For this study, we used vocal and behavioral data recorded from four wild chimpanzee communities. Three communities of western chimpanzees in Tai (East, North, and South communities) in the Tai National Park, Ivory Coast, within the Tai Chimpanzee Project (Wittig and Boesch, 2019) and one community of eastern chimpanzees in the Budongo forest (Reynolds, 2005), Uganda (the Sonso community). We sampled 78 chimpanzees in total (15 females and 11 males in Budongo and 33 females and 19 males in Tai, see details in the STAR Methods).

Comparing the context of production of single calls (pant-grunt, pant hoots) and call sequence (greeting hoot)

We used 232 PG, 178 PH, and 83 GH recorded across the four study communities for which we had all available information to test for differences in contexts of production. We used two generalized linear mixed models (GLMM) to test for differences in contexts of production between PG and GH (Model 1a) and between PH and GH (Model 1b). In both models, four contexts were considered as test predictors to determine if each call was produced during an approach, during a fusion, during inter-party calling, or addressed to a higher ranking individual. The response variable was whether the call was PG or GH (Model 1a) and PH or GH (Model 1b). In addition, we included in each model individual ID and community ID as random factors to account for repeated sampling on the same individuals and within the same communities.

In Model 1a, the full model was significantly different from the null model ($N = 261$ calls from 65 individuals, $LRT, df = 5, \chi^2 = 14.95, p = 0.011$), indicating that the greetings consisting of PG only and greetings consisting of GH sequences (PG + PH or PH + PG) differed in their contexts of production. Specifically, GH sequences were significantly more likely to be produced during inter-party communication ($p = 0.029$,

Table 1. Results of models comparing contexts of production of chimpanzee pant grunts, pant hoots, and the combination of both, the greeting hoots (Models 1a and 1b)

Model	Response	Predictor	Estimate	SE	CI _{low}	CI _{high}	χ^2	P
1a	Binomial: PG (0) vs. GH (1)	(Intercept)	-1.14	0.84	-1.68	-0.74		
		Inter-party calling (Yes)	1.44	0.68	0.07	2.88	4.789	0.029
		Fusion (Yes)	1.03	0.44	0.24	1.92	6.087	0.014
		Approach (Yes)	-0.12	0.49	-0.61	0.16	0.000	1.000
		Caller subordinate to receiver (Yes)	-0.84	0.90	-1.76	0.00	1.322	0.516
		Caller subordinate to receiver (No receiver)	-0.12	0.87	-0.40	0.22		
1b	Binomial: PH (0) vs. GH (1)	(Intercept)	-2.77	0.74	-5.61	-1.73		
		Inter-party calling (Yes)	-0.64	0.59	-2.03	0.52	1.096	0.295
		Fusion (Yes)	0.72	0.52	-0.35	1.80	1.812	0.178
		Approach (Yes)	2.69	0.52	1.74	4.30	29.594	<0.001
		Caller subordinate to receiver (Yes)	-0.20	0.78	-1.63	2.72	18.688	<0.001
		Caller subordinate to receiver (No receiver)	2.07	0.73	0.87	4.66		

PG, PH, and GH indicate pant grunts, pant hoots, and greeting hoots, respectively. SE indicates the SE of the estimate for each predictor. The coded level for each categorical predictor is indicated in parentheses. Significant p value ($p < 0.05$) is indicated in bold. CI_{low} and CI_{high} indicate the lower and upper limits of the 95% confidence interval for the estimates of each predictor.

Table 1, Figure 2) and party fusion ($p = 0.014$, Table 1, Figure 2). However, there were no significant differences in the likelihood of GH and PG being produced during approaches or to higher ranking recipients (all $p > 0.5$, Table 1, Figure 2). The marginal R^2 and the conditional R^2 for Model 1a were 0.061 and 0.388, respectively.

In Model 1b, the full model was significantly different from the null model ($N = 315$ calls and 70 individuals, LRT, $df = 5$, $\chi^2 = 154.11$, $p < 0.001$), indicating that PH and GH differed in their context of production. GH were significantly more likely to be produced during approaches ($p < 0.001$, Table 1, Figure 2) and to be directed toward higher ranking individuals ($p < 0.001$, Table 1, Figure 2) than PH. There was no significant difference in the likelihood of GH and PH to be produced during inter-party communication ($p = 0.295$) or during fusion ($p = 0.178$, Table 1). The marginal R^2 and the conditional R^2 for Model 1b were 0.469 and 0.490, respectively.

A visual inspection of context plotted by population (Figure S1) revealed that the differences in context between GH, PH, and PG (Figure S1) were mostly consistent with the differences described above (the differences over the entire dataset). In both Budongo and Tai chimpanzees, GH was produced more frequently during approaches and was more often produced toward higher ranking individuals than PH. Furthermore, GH was more often produced during party fusion than PG alone in both populations, even if the differences in Budongo were less clear than in Tai. Finally, while GH was more often produced during inter-party communication in Tai as compared to PG, this difference between GH and PG in the frequency of production during inter-party communication was not present in Budongo.

In a third model (Model 1c), we established if GH sequences starting with a PH were produced in similar contexts to GH sequences starting with a PG to rule out that contexts of production were the factor driving potential population differences in call ordering. We used the same predictor and random factors as for models 1a and 1b and used as binomial response whether the GH sequence started with a PH or a PG. In this model, the full model was not significantly different from the full model ($N = 75$ calls and 32 individuals, LRT, $df = 4$, $\chi^2 = 3.54$, $p = 0.473$, Figure S3), indicating that greeting hoot sequences starting with a PH were produced in similar contexts than greeting hoot sequences starting with PG. The marginal R^2 and the conditional R^2 for Model 1c were 0.058 and 0.445, respectively.

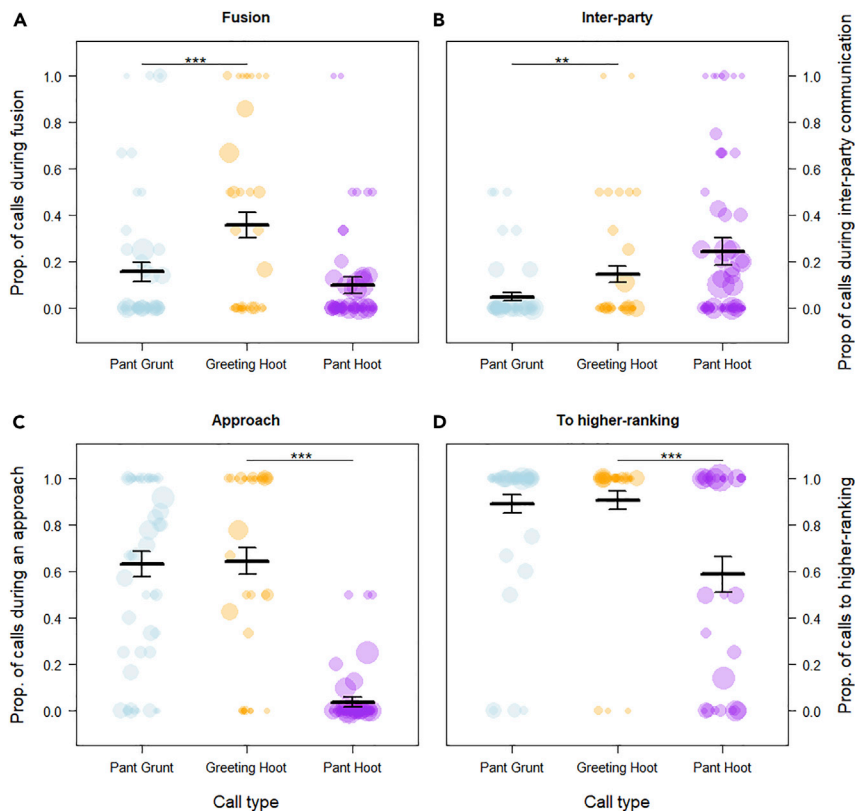


Figure 2. The contexts of production of the greeting hoot sequence compared with its component parts produced singularly, pant grunt and pant hoot, across four chimpanzee communities

The four contexts of interest are depicted across the four panels from top left to bottom right: A) fusion, B) Inter-party communication, C) Approach, D) Caller directed the call to a higher ranking recipient. Pant-grunts (PG) are depicted in blue, greeting hoots (GH) in yellow, and panted-hoots (PH) in purple. Each dot represents one individual chimpanzee and the size of the dot is proportional to the number of calls recorded for this individual. The thick line depicts the mean proportion of each context for each call and the upper and lower bars of the SE. The “***” indicate the differences which were significant (* $p < 0.1$, ** $p < 0.05$).

Group differences in ordering of pant-grunt and pant hoots within greeting hoot

We ran a GLMM (Model 2) to test for the influence of community ID on the likelihood for the GH sequence to start with a PH. We included party size at the time of call production and the rate of aggression received by the caller during the study period as control predictors. We also included individual ID as a random factor to account for repeated sampling of the same individuals. We used $N = 118$ GH sequences from 38 individuals (31 GH sequences from six individuals in Sonso and 23, 39, and 25 GH sequences from 8, 11, and 13 individuals in Tai East, North, and South, respectively) for which we had all the information on the test and control variables. In Model 2, the full model was significantly different from the null model (LRT, $df = 5$, $\chi^2 = 153.38$, $p < 0.001$), indicating a difference between the communities in the likelihood to produce PH first in the GH sequence. We found a significant effect of community on the likelihood to produce PH first in the GH sequence (LRT, $df = 3$, $p = 0.001$, Table 2) with individuals of all of the three Tai communities being more likely to produce PH first in the GH as compared to Sonso (Figure 3). PH was 12-14 times (78-87% of cases, 18/23, 34/39, and 21/25 in Tai East, North, and South, respectively) more likely to come first in the GH sequence in all the three Tai communities as compared to in Sonso (6% of cases, 2/31). The probability for PH to come first in the GH sequence was: 0.78, 0.87, and 0.84 in Tai East, Tai North and Tai South, respectively, versus 0.06 in Sonso. The three communities in Tai did not differ significantly from each other and were all similarly likely to utter more often PH first than PG first within the GH sequence. The difference between the two study populations was consistent across the years. The percentage of PH first in the GH sequence varied in the three Tai communities and across years between 67 and 100% and between 0 and 10% across years in Sonso (Table S2). In Model 2, complete data could be entered into the model

Table 2. Results of models assessing community differences in ordering of call types in the chimpanzee bigram vocalization, the greeting hoot (Model 2)

Model	Response	Predictor	Estimate	SE	CI _{low}	CI _{high}	χ^2	P
2	Binomial: PH before PG in the GH sequence (Y/N)	(Intercept)	-3.16	1.32	-22.36	-1.04		
		Average rate of aggression received	-0.77	0.57	-3.87	0.74	1.779	0.182
		Party size	0.60	0.38	-0.16	2.62	2.469	0.116
		Group (Tai East)	5.10	1.64	2.20	32.99	16.158	0.001
		Group (Tai North)	6.68	1.75	3.85	36.42		
		Group (Tai South)	5.02	1.58	2.04	31.63		

PG, PH, and GH indicate pant grunts, pant hoots, and greeting hoot sequences, respectively. SE indicates the SE of the estimate for each predictor. The coded level for each categorical predictor is indicated in parentheses. Significant p value ($p < 0.05$) are indicated in bold. CI_{low} and CI_{high} indicate the lower and upper limits of the 95% confidence interval for the estimates of each predictor.

from six individuals from Sonso. To determine if the PG-first pattern found in Sonso was limited to these six individuals or was also representative of a more community-wide Sonso greeting hoo pattern, we computed the call ordering in the GH sequences of five additional individuals from Sonso which could not be entered into the model either because we did not have information on the rate of aggressions they received ($N = 2$), or because they produced GH sequences starting with another call (e.g. grunt + PG_PH). We analyzed the 12 GH sequences available from these individuals. Individuals uttered PG first in 11 of these 12 sequences (i.e., 92%) highlighting that the pattern found in Budongo is not limited to the six individuals which entered the statistical model.

We found no significant effect of party size or of the rate of aggression received by the caller on the likelihood to produce PH first in GH (both $p > 0.11$, Table 2). The median rate of aggression received was higher in Sonso than in all the three Tai communities but individuals receiving the highest rates of aggression were found in Tai North (Figure S4). The marginal R^2 and the conditional R^2 for Model 2 were 0.547 and 0.755, respectively. To explore further if the call ordering difference between PH and PG in the greeting hoot sequence was triggered by proximate social factors, we used the detailed behavioral data accompanying each recording to assess whether greeting hoots were uttered following aggression by the recipient of the call in each population. Out of 118 GH sequences in Model 2, we had this detailed information for 112 GH sequences (28 in Sonso and 84 in Tai). The rates of aggression received by callers before uttering the GH sequence were very similar in the two populations (7.1% - 2/28 - in Sonso and 9.5% - 8/84 - in Tai).

Population differences in aggressiveness

To examine some of the potential drivers of differences in PH and PG ordering within the GH sequence, we calculated the rate of aggression given by the most aggressive male in each community (see STAR Methods). The most aggressive individual was a male in each of the four study communities. However, the aggressiveness of the most aggressive individual was two to seven times higher in Sonso (average of 1.43 aggressions given per hours) as compared to the three Tai communities (0.41, 0.67, and 0.19 aggressions given per hour in Tai East, North, and South, respectively). The most aggressive individual was present in the party when GH was emitted most of the time in Sonso (56%) and in Tai South (58%). In contrast, the most aggressive individual was present only during 26 and 36% of GH emissions in Tai East and in Tai North, respectively.

Population differences in intra- and inter-group killing

Similar to aggressiveness above, we also compiled differences in intra- and inter-group killing rates between Budongo and Tai populations. First, we compiled the report of killings which took place during the four years preceding each of the data collection periods as well as during the years of data collection in each study community (i.e., 2003-2012 in Budongo (10 years), 2015-2020 in Tai east (6 years), 1995-2000 and 2015-2020 in Tai North (13 years) and 2015-2020 in Tai South (6 years). During these periods, eight intra-group killings were reported in Sonso (i.e., 0.8 killing per year on average), whereas no intra-group killings were reported in any of the three Tai communities. During the same periods, 4 inter-group killings

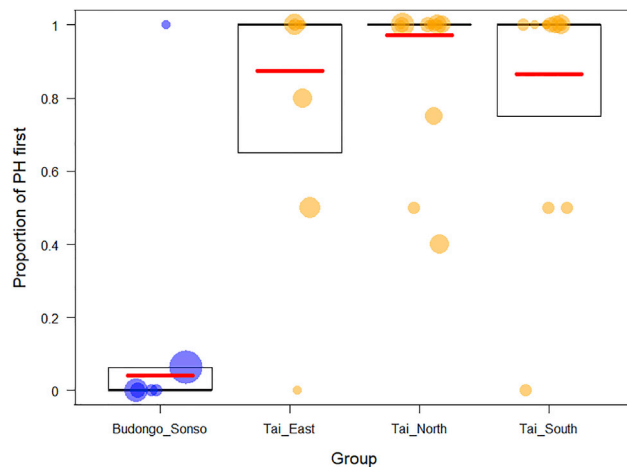


Figure 3. Variation in the order of production of single calls (pant hoots (PH) and pant grunt (PG)) within greeting hoot sequence across four chimpanzee communities

The y axis depicts the likelihood for the PH to be emitted first in the sequence (i.e. the likelihood for the greeting hoot sequence to be PH + PG as opposed to the alternative PG + PH sequence). Eastern Budongo chimpanzees are depicted in blue and Western Tai chimpanzees are depicted in orange. Each dot represents one individual chimpanzee and the size of the dot is proportional to the number of GH recorded for this individual. The boxplot depicts the median (thick line) and the 25% and 75% quartiles. The red line depicts the model line extracted predicted by Model 2.

were reported in Sonso (i.e., 0.4 inter-group killings on average), one inter-group killing was reported in Tai East (i.e., 0.17 inter-group killing per year on average), and none in Tai South and Tai North. These patterns are in line with what has been reported over longer periods in both field sites (reviewed in [Wilson et al., 2014](#)). No intra-group killing has ever been reported in any Tai community over 119 observation group_years ([Wilson et al., 2014](#) and unpublished long-term data from the Tai Chimpanzee project: [Wittig and Boesch, 2019](#)), whereas 12 intra-group killings were reported over 23 years in the Sonso community in Budongo ([Wilson et al., 2014](#)). Inter-group killings appear considerably rarer in Tai than in Budongo but have been reported on occasion (3 reports in total across several communities since 1979, [Wilson et al., 2014](#), own unpublished data).

DISCUSSION

The limited capacity of primates to modify the acoustic structure of their species-specific calls ([Egnor and Hauser, 2004](#); [Fedurek and Slocombe, 2011](#); [Fischer et al., 2015](#); [Fischer and Hammerschmidt, 2020](#)) is in stark contrast with the ability of vocal production learners, such as humans, birds, and cetaceans, to produce completely new songs or calls ([Sewall et al., 2016](#); [Vernes and Wilkinson, 2020](#); [Wilbrecht and Nottebohm, 2003](#)). In this study, we highlight the flexibility in vocal sequence production across two populations of the same species. We showed that chimpanzees from different populations produced greeting hoot (GH) sequences with reversed order of PH and PG in the sequence. A vast majority of GH sequences uttered by chimpanzees in Budongo started with a PG, whereas the opposite was found in chimpanzees from all three communities in Tai. We found no effect of the rate of aggression received by the caller on the likelihood of uttering PH or PG first in the greeting hoot. However, when comparing the general aggressiveness of the most aggressive individual in each community, we found that this individual was much more aggressive in Budongo than in any of the three Tai study populations. The call ordering in the GH sequences is, however, unlikely to be driven by a proximate reaction, such as fear. First, PG-first sequences are not necessarily emitted by the chimpanzees that receive the most aggression, nor by those greeting the most aggressive male (for the latter, 55% of GH sequences in Budongo were not uttered toward the most aggressive individual but toward nine other individuals, including both males and females, and toward eight, seven, and seven different receivers, also including males and females, in Tai North, Tai South, and Tai East, respectively). Furthermore, recipients of the calls were not more aggressive toward the caller in Budongo than in Tai. Our results rather support the hypothesis that uttering PH or PG first in the GH sequence is influenced by consistent long-term population differences in the general levels of intra-group aggression and lethal risk rather than short-term variation in the level of threat experienced by each caller.

In fact, several intra-group killings were documented during the 4 years leading up to and including the study period in Budongo but not in Tai.

When examining contexts of production, the single calls PG and PH demonstrate differences in the context of production, such that PG is more likely to be emitted during approaches to dominants, indicating a function of submissive greeting (Fedurek et al., 2021), whereas PH is more likely to be emitted during inter-party calling (also described in Arcadi, 1996; Goodall, 1986). We also found that when PG and PH are combined into the GH sequence, the greeting component of the PG (emission during approaches to dominants), is retained, as is the inter-party calling component of the PH. The GH sequence is a form of greeting more frequently used during fusion (36% of GH sequences are emitted during fusions) than PG alone (emitted only 16% of the cases during fusions).

Fusions can be tense, highly vocal events characterized by aggressions or affiliations, as dominance and bonding relationships are re-established after a period of absence. Rapid juxtaposition of vocalizing own identity and submission may reduce received aggression, such that the greeting hoot sequence may have emerged as a beneficial vocal trait. This suggests that as recently shown for another chimpanzee vocal sequence, the combination pant-hoot + food call (Leroux et al., 2021), greeting hoot sequences may also follow the “principle of compositionality” in which the “meaning” of a sequence retains the integrity of the meaning of its parts (Hurford, 2012; Leroux and Townsend, 2020; Suzuki et al., 2018). Also, the single calls, pant hoot, and food grunt were each found to be acoustically similar to their respective calls when combined into a sequence (pant hoot + food grunt; Leroux et al., 2021). Likewise, being easily recognizable by the trained human ear, it is highly likely that the acoustic structure of the PH and PG elements is retained when these elements are combined into the GH sequence, although this would need to be confirmed with acoustic analyses. Our study, combined with the findings from Leroux et al., 2021, show that PH is often combined with other more context-specific calls. Such a combination of an identity marker contact call (such as panted hoots) with another more context-specific (submissive greeting) call is not unique to chimpanzees but has been found in other primate and non-primate mammalian species (Coye et al., 2016; Jansen et al., 2012). Documenting such simple forms of compositionality in non-human animals is important in retracing the evolutionary steps toward simple syntactic constructs in human speech (Leroux and Townsend, 2020; Townsend et al., 2018). Playback experiments (e.g. as in Japanese tits, Suzuki et al., 2016) will be important for confirming the function and meaning of greeting hoots and whether receivers are sensitive to call order within greeting hoot sequences.

Greeting hoots may show compositionality but what mechanisms explain the difference in PH and PG call order between the two study populations? Geographical distance between populations is well known to influence the acoustic properties of acoustic signals, often single call types, across mammals, even in non-singing species (reviewed in Lameira et al., 2010). However, in non-singing species, geographical variation in the construction of vocal sequences is rarely examined and there are few theoretical frameworks to propose selection pressures that promote such variation. Thus, we will discuss the most common mechanisms proposed to explain geographical variation in the acoustic properties of vocal signals in general (Ey and Fischer, 2009; Fischer and Hammerschmidt, 2020; Lameira et al., 2010), and those which have been proposed to explain the differences between chimpanzee populations in the acoustic structure of pant-hoots (Mitani et al., 1992, 1999), can explain why Budongo chimpanzees utter PG first and Tai chimpanzees utter PH first within the greeting hoot sequence. We considered the five following mechanisms: 1) genetic differences between populations, 2) differences in habitat structure (visibility and acoustic properties), 3) differences in the context of production, 4) temporal variation, and 5) vocal usage learning.

Genetic differences between populations

In chimpanzees (Menzel, 1964), as in other primates (Egnor and Hauser, 2004; Fedurek and Slocombe, 2011; Fischer et al., 2015; Fischer and Hammerschmidt, 2020; Menzel, 1964), the acoustic structure of single call types is mostly genetically programmed. The same call types are found across populations (reviewed in Crockford, 2019). However, could the population-specific ordering of such call types into sequences be transmitted genetically? Each study population is a different subspecies, eastern chimpanzees in Budongo (*Pan troglodytes schweinfurthii*) and western chimpanzees in Tai (*P. troglodytes verus*) which separated c.a. 0.84 million years ago (Becquet et al., 2007). Could they have inherited genetically different greeting hoot sequences? To our knowledge, there is no report of genetic transmission of call order in vocal sequences. Furthermore, while a large majority of greeting hoot sequences start with PG in Budongo and with PH in

Tai, some individuals in each population (1 in Budongo and seven in Tai) uttered both variants of greeting hoot sequences with either PH or PG first. This rules out the possibility of a fixed genetically programmed sequence ordering in each population.

Differences in habitat structure and visibility

Animals may adjust their acoustic signal to optimize sound propagation in diverse environments, such as by lowering fundamental frequency in more closed habitats (reviewed in [Ey and Fischer, 2009](#); [Lameira et al., 2010](#)). However, it is hard to conceive how this mechanism would select individuals producing PH or PG first within GH sequences, given that both populations of chimpanzees are forest dwelling, living in closed habitats. Furthermore, PH is loud calls that can travel up to 500 m ([Kalan et al., 2016](#)) and it is unlikely that PH lose their *far reaching properties* if they are preceded or succeeded by a PG. Call order may rather be impacted by a combination of the social environment and visibility and in particular, by group cohesion. For example, if visibility impacts call order, we might expect that in forest areas where individuals can only see each other over short distances, pressure to reveal their identity should favor the emission of PH first. Budongo chimpanzees live in a secondary forest ([Reynolds, 2005](#)) with generally lower visibility (<20 m) than the more open primary rain forest in Tai ([Wittig and Boesch, 2019](#)), with visibility >40 m. However, visibility varies considerably within all communities. Budongo chimpanzees often use the dense trail system to travel in the forest where the visibility is higher (>40 m) and Tai Forest also comprises denser areas owing to tree falls or rocky outcrops with visibility <20 m. If visibility played a role, given that GH is emitted in variable habitat, much less consistency in GH call order within a given population would be expected than we observe, given the variability in visibility across their home ranges. Finally, we did not find a significant effect of party size on the likelihood to produce PH or PG first in the GH sequence which indicates that group cohesion does not affect call ordering.

Differences in the context of production

We found no significant differences in the context of production between greeting hoot sequences starting with PH and the ones starting with PG. Furthermore, greeting hoot sequences were produced in very similar contexts in Tai and Budongo ([Figure S3](#)) and the slight differences in the context of production (e.g. GH is more often produced during inter-party calling than PG in Tai, but not in Budongo) may be related to the limited sample size in Budongo. We can thus rule out that the difference found is driven by different contexts of production.

Temporal variation

If the within-population tendency to produce PH or PG first in the greeting hoot sequence varied across time in response to proximate situational changes (e.g., a period with a particularly high frequency of inter-group encounters which could promote the production of PH first, to aid rapid identification as an in-group member by other group members), our result could be attributed to different sampling periods rather than to consistent long-lasting population differences. However, sampling the Tai chimpanzees across a discontinuous 22 years period (1998-2020) and the Budongo chimpanzees across a 6 year period (2007-2012) revealed within-population consistency. In fact, across each of the study years, chimpanzees in Tai consistently uttered a vast majority of GH sequences starting with PH, whereas chimpanzees in Budongo started a vast majority of their sequences with a PG ([Table S2](#)). Thus, it seems unlikely that the population difference in call ordering is owing to temporal sampling bias.

Vocal usage learning

After discarding alternative hypotheses, we argue that consistent population differences in call ordering in the greeting hoot sequence are most likely explained by usage learning ([Janik and Slater, 2000](#); [Sewall et al., 2016](#)) of vocal sequences. We cannot prove with our results that chimpanzees learn vocal sequences but this hypothesis appears to be the most parsimonious explanation. In fact, if the order of PG and PH in the GH sequence was determined by a proximate reaction to the aggressive or threatening behavior of the call receiver, we would find clear population differences in receiver's behavior, which we did not find. Another possible explanation for stable patterns within populations could be related to our small sample size, which could underrepresent variation. However, given that the greeting hoots were emitted from a number of emitters to a number of receivers over several years of data collection in each population makes this explanation unlikely. Replication of these results in future studies remains nevertheless important to confirm our conclusion.

The capacity for usage learning of vocal sequences is not beyond the reach of primates. Vocal usage learning can operate via social learning as demonstrated by a recent study showing that social exposure during ontogeny is crucial to the acquisition of vocal sequences in marmoset monkeys (Gultekin et al., 2021). Studies on the acoustic structure of the pant-hoot sequence suggest that chimpanzees may also learn the use of vocal sequences socially. The pant-hoot sequences comprise hoos, panted-hoos (the PH in our study), screams or panted-scream, and a let-down phase. In one study, most of the pant-hoot sequences which deviated from the stereotypical structure were uttered by subadult males, and younger males often joined older males in producing this vocalization but rarely initiated it (Arcadi, 1996). If call order in the greeting hoot sequence is socially transmitted, this sequence could thus be a cultural trait transmitted through vocal usage learning as proposed for other population-specific vocal traits in orangutans (Lameira et al., 2010; van Noordwijk et al., 2006; van Schaik et al., 2003). In terms of non-vocal behaviors, great apes in general, and chimpanzees and orangutans in particular, are able to learn from others and innovate at a higher rate than other primate taxa (Reader and Laland, 2002), and possess a wider range of cultural behavioral traits (Schaik et al., 2003; Whiten et al., 1999), although this has rarely been reported in the vocal domain. Social transmission of vocal sequences in chimpanzees would explain population consistency in call ordering in our study. Alternatively, vocal usage learning could operate via individual learning whereby individuals would learn the most beneficial call order in the GH sequence via trial and error (e.g., by receiving aggressions more frequently when producing the sequence in the reverse order). Our results do not allow to disentangle whether individuals learn socially or individually the order in the greeting hoot sequence.

Regardless of the mechanisms, the consistency of the call ordering in each population shows that the GH sequence is a stable within-population vocal phenotype maintained through time and across individuals.

As we did not find an effect of the overall rate of aggression received by a caller during the years of data collection on their likelihood to put PH first in the GH sequence it is unlikely that social experience as an adult drives call ordering in the GH sequence. This was not owing to a lack of variation in the rate of aggression received in each study community as this rate varied greatly and more or less equally within each community (Figure S4). Whilst it is difficult to rule out all possible ecological causes that could promote such consistent population-specific call order effects, we have assessed all those that we considered could be applicable from the literature. We cannot completely exclude that call recipients consistently express subtle behaviors during approaches of the caller which could trigger different reactions in each population. However, our detailed observations of the caller's and recipient's behavior before and after the GH sequences were uttered indicate that this is unlikely to be the case. Furthermore, nuanced behaviors are more likely to promote flexible call ordering rather than lead to the population-specific patterns found here.

Sexual selection is also unlikely to have led to the observed patterns as individuals of each sex were sampled in each population and both males and females produced greeting hoot sequences with PH and with PG first.

Overall, our results indicate that call ordering in GH sequences is likely to be learned during ontogeny and before reaching adulthood (only adult individuals were considered in our study). Comparative ontogenetic studies establishing whether young chimpanzee shift from random call ordering in the GH sequence to population-specific ordering throughout development would provide important information to show that this call sequence is, indeed, learned, as suggested, but not demonstrated, by our study. Such studies should monitor in particular whether young chimpanzees are more likely to receive aggressions when uttering the GH sequence in the reverse order to the population-specific order and whether aggressions received impact the call order within the next GH sequence produced to fully investigate the possibility that GH sequences are individually learnt by trial and error. Alternatively, studies in captivity could expose young chimpanzees to GH sequences in a certain order to assess if these sequences can also be learnt "socially" without trial and error and behavioral feedback such as aggressions.

Differences in population-level risk of lethal aggression may have driven the divergence in call order across the two populations, and possibly its transmission over time either via social or individual usage learning. In fact, throughout the study period, we documented consistently higher aggressiveness of the most aggressive male in Budongo compared to Tai across all study periods, which may promote individual learning

during ontogeny, and evidence of intra-group lethal aggression in Budongo but never in Tai, which may, in turn, promote the selection of a socially transmitted trait in a population. PH_PG may have been the ancestral form of the GH sequence to respond to the *inter*-group threat which is high across all chimpanzee populations (Wilson et al., 2014). Tai chimpanzees may have retained this pattern as they are exposed to high *inter*-group threat (Lemoine et al., 2020), despite the low rate of *inter*-group killing, whereas in Budongo the high *intra*-group threat and killing risk may have led to a reversal of the sequence toward PG_PH. Our study thus supports the social pressure hypothesis, that social pressure may promote flexibility in call order and be a driver of vocal usage learning in social calls, particularly here in the ordering of non-human animal vocal sequences (Sewall et al., 2016; Vernes and Wilkinson, 2020).

In conclusion, our study provides some rare evidence for consistent population variation in call ordering within a vocal sequence combining social calls in a non-singing animal species. The contextual usage of single calls compared to that of the combined call showed some features of compositionality. However, the order of the single units in the combined call did not change contextual use. Such consistency may constitute a vocal tradition in each population transmitted temporally from one generation to the next via social learning like other cultural traits found in great ape species and in particular chimpanzees such as tool use (Whiten et al., 1999). Alternatively, chimpanzee may learn individually from trial and error the most efficient call order leading to reduced risk of receiving aggression and thereby adopt the population-specific call order. Regardless of the mechanism, our study supports the general hypothesis that within-species variation in social pressure may select for flexibility in call ordering and hence promote the evolution of learning of vocal sequence order. In this case, social pressure (rates of lethal aggression) differs at the population level (likely driven by different ecological pressures) and this social pressure may contribute to differential benefits in producing one call before another, resulting in a proclivity for emitting different sequence variants in different populations (Sewall et al., 2016; Vernes and Wilkinson, 2020). Like other primates, chimpanzees produce a limited range of single-call types. However, they may need communicative strategies that allow them to communicate a wider range of meanings than there are call types to navigate their complex social environment. Here, flexibility in call ordering within vocal sequences may be one such “strategy” that contributes to a flexible communicative system that can convey more meanings than there are call types (Girard-Buttoz et al., 2022).

Limitations of the study

In the current study, we provide evidence for population differences in call ordering and use a process of discarding alternative hypothesis to conclude that the most likely mechanism explaining these results is through vocal usage learning of vocal sequences in each population. One drawback of this approach is that it provides evidence by exclusion and direct evidence would need to be assessed in the future. One option is to examine the ontogenetic development of the greeting utterance. If social learning shapes call order, greater initial flexibility in call order might be expected at young ages compared to adults. Future studies investigating the ontogeny of greeting hoot sequence acquisition using a combination of naturalistic vocal recordings and playback experiments on both immature and adult individuals (see work in Japanese tits Suzuki et al., 2016) would help shed light on such mechanisms. Further tests of the “social pressure hypothesis” are needed, across species, to assess whether other species demonstrate similar call order flexibility, particularly when within-species variation in social pressures would make call order flexibility beneficial, either within or between populations. Another limitation of the behavioral but not the vocal part of our study is that we did not conduct inter-observer reliability tests across all the observers on the behavioral recordings at the two sites. Behaviors included, however, such as fusions, and aggressions involving contact or locomotion are very conspicuous and thus robust to observer variation (Witig and Boesch, 2019). Future comparative studies could implement inter-observer tests at the beginning of the study or ensure that data are collected across sites by the same observers. This might, however, reduce the opportunity to collect data across long time periods. Finally, we bear in mind that our sample size is relatively small. We see this study as one to generate interest in the potential of the social pressure hypothesis to explain selection for flexibility in vocal usage. This can be further tested in future studies by collecting a larger number of sequences within each study population, and across a wider range of populations with variable socio-ecology.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- METHOD DETAILS
 - Vocalisation recordings
 - Vocal coding
 - Behavioral observations
 - Dominance hierarchy
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Context of production for PG, PH and GH
 - Group differences in ordering of PG and PH within GH

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104851>.

ACKNOWLEDGMENTS

We thank the Ministère de l'Enseignement Supérieur et de la Recherche Scientifique and the Ministère de Eaux et Forêts in Côte d'Ivoire, and the Office Ivoirien des Parcs et Réserves for permitting the study in Ivory Coast. We thank the Uganda Wildlife Authorities and the Ugandan National Council for Science and Technology for permitting the study in Uganda. We are very grateful to Christophe Boesch for his tremendous work in establishing and nurturing the Taï Chimpanzee Project for 30 years. Thanks to the Center Suisse de Recherches Scientifiques en Côte d'Ivoire and the staff members of the Taï Chimpanzee Project and Kayla Kolff for their support and collecting data in Taï. Thanks to the staff members of Budongo Conservation Field Station for their support in collecting data in Budongo. This study was funded by the Max Planck Society and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program awarded to C.C. (grant agreement no. 679787) and ERC (Prilang GA283871) and by Leverhulme Trust Research Leadership Award. Core funding for the Taï Chimpanzee Project has been provided by the Max Planck Society since 1997 and for Budongo Conservation Field station by the Royal Zoological Society of Scotland.

AUTHOR CONTRIBUTIONS

CC and CGB designed the study. CC, MG, ML, RMW, and TB collected the data. KZ and RMW directed the field sites in which the data were collected. CGB conducted the statistical analysis and wrote the first draft of the article with a substantial contribution from the co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 10, 2022

Revised: April 1, 2022

Accepted: July 25, 2022

Published: September 16, 2022

REFERENCES

- Arcadi, A.C. (1996). Phrase structure of wild chimpanzee pant hoots: patterns of production and interpopulation variability. *Am. J. Primatol.* 39, 159–178. [https://doi.org/10.1002/\(SICI\)1098-2345\(1996\)39:3<159::AID-AJP2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1098-2345(1996)39:3<159::AID-AJP2>3.0.CO;2-Y).
- Baayen, R.H., Davidson, D.J., and Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>.
- Barr, D.J., Levy, R., Scheepers, C., and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Barton, K. (2020). MuMIn: Multi-Model Inference. R package version 1.43.17. <https://CRAN.R-project.org/package=MuMIn>.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Becquet, C., Patterson, N., Stone, A.C., Przeworski, M., and Reich, D. (2007). Genetic structure of chimpanzee populations. *PLoS Genet.* 3, e66. <https://doi.org/10.1371/journal.pgen.0030066>.
- Bezerra, B., and Souto, A. (2008). Structure and usage of the vocal repertoire of *Callithrix jacchus*.

- Int. J. Primatol. 29, 671–701. <https://doi.org/10.1007/s10764-008-9250-0>.
- Bortolato, T., Mundry, R., Wittig, R.M., Girard-Buttoz, C., Crockford, C., Development of vocal sequences: structural complexity and combinatorial potential, in wild chimpanzees (*Pan troglodytes verus*), in revision.
- Boughman, J.W. (1997). Greater spear-nosed bats give group-distinctive calls. *Behav. Ecol. Sociobiol.* 40, 61–70. <https://doi.org/10.1007/s002650050316>.
- Bygott, J. (1979). Agonistic behaviour, dominance, and social structure in wild chimpanzees of the Gombe National Park. *The Great Apes (Benjamin-Cummings)*, pp. 405–428.
- Colmenares, F., Hofer, H., and East, M.L. (2000). Greeting ceremonies in baboons and hyenas. In *Natural Conflict Resolution*, F. Aureli and F.B.M. de Waal, eds. (University of California Press), pp. 94–96.
- Coye, C., Zuberbühler, K., and Lemasson, A. (2016). Morphologically structured vocalizations in female *Diana* monkeys. *Anim. Behav.* 115, 97–105. <https://doi.org/10.1016/j.anbehav.2016.03.010>.
- Crockford, C., and Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour* 142, 397–421. <https://doi.org/10.1163/1568539054012047>.
- Crockford, C., Herbinger, I., Vigilant, L., and Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* 110, 221–243. <https://doi.org/10.1111/j.1439-0310.2004.00968.x>.
- Coye, C., Zuberbühler, K., and Lemasson, A. (2022). The evolution of vocal communication: inertia and divergence in two closely related primates. *Int. J. Primatol.* 43, 712–732.
- Crockford, C. (2019). Why Does the chimpanzee vocal repertoire remain poorly understood? - and what can be done about it. In *The Chimpanzees of the Tai Forest: 40 Years of Research*, Boesch, R.M. Christophe, R.M. Wittig, C. Crockford, L. Vigilant, T. Deschner, and F.H. Leendertz, eds. (Cambridge University Press), pp. 394–409.
- de Vries, H. (1998). Finding a dominance order most consistent with a linear hierarchy: a new procedure and review. *Anim. Behav.* 55, 827–843. <https://doi.org/10.1006/anbe.1997.0708>.
- Desai, N.P., Fedurek, P., Slocombe, K.E., and Wilson, M.L. (2021). Chimpanzee pant-hoots encode information about individual but not group differences. Preprint at bioRxiv. <https://doi.org/10.1101/2021.03.09.434515>.
- Dobson, A.J. (2002). *An Introduction to the Generalized Linear Models*, 2nd edn. (Chapman and Hall/CRC Press).
- East, M., Hofer, H., and Wickler, W. (1993). The erect 'penis' is a flag of submission in a female-dominated society: greetings in Serengeti spotted hyenas. *Behav. Ecol. Sociobiol.* 33, 355–370.
- Eckhardt, N., Polansky, L., and Boesch, C. (2015). Spatial cohesion of adult male chimpanzees (*Pan troglodytes verus*) in Tai National Park, Côte d'Ivoire. *Am. J. Primatol.* 77, 125–134. <https://doi.org/10.1002/ajp.22316>.
- Egnor, S.E.R., and Hauser, M.D. (2004). A paradox in the evolution of primate vocal learning. *Trends Neurosci.* 27, 649–654. <https://doi.org/10.1016/j.tins.2004.08.009>.
- Elowson, A.M., and Snowdon, C.T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Anim. Behav.* 47, 1267–1277. <https://doi.org/10.1006/anbe.1994.1175>.
- Ey, E., and Fischer, J. (2009). The "acoustic adaptation hypothesis"—a review of the evidence from birds, Anurans and mammals. *Bioacoustics* 19, 21–48. <https://doi.org/10.1080/09524622.2009.9753613>.
- Fedurek, P., and Slocombe, K.E. (2011). Primate vocal communication: a useful tool for understanding human speech and language evolution? *Hum. Biol.* 83, 153–173. <https://doi.org/10.3378/027.083.0202>.
- Fedurek, P., Zuberbühler, K., and Dahl, C.D. (2016). Sequential information in a great ape utterance. *Sci. Rep.* 6, 38226. <https://doi.org/10.1038/srep38226>.
- Fedurek, P., Neumann, C., Bouquet, Y., Mercier, S., Magris, M., Quintero, F., and Zuberbühler, K. (2019). Behavioural patterns of vocal greeting production in four primate species. *R. Soc. Open Sci.* 6, 182181. <https://doi.org/10.1098/rsos.182181>.
- Fedurek, P., Tkaczynski, P.J., Hobaiter, C., Zuberbühler, K., Wittig, R.M., and Crockford, C. (2021). The function of chimpanzee greeting calls is modulated by their acoustic variation. *Anim. Behav.* 174, 279–289. <https://doi.org/10.1016/j.anbehav.2021.02.002>.
- Fischer, J., and Hammerschmidt, K. (2020). Towards a new taxonomy of primate vocal production learning. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190045. <https://doi.org/10.1098/rstb.2019.0045>.
- Fischer, J., Wheeler, B.C., and Higham, J.P. (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Curr. Biol.* 25, R1028–R1029. <https://doi.org/10.1016/j.cub.2015.09.010>.
- Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F., and Hammerschmidt, K. (2020). Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proc. R. Soc. B Biol. Sci.* 287, 20202531. <https://doi.org/10.1098/rspb.2020.2531>.
- Foester, S., Franz, M., Murray, C.M., Gilby, I.C., Feldblum, J.T., Walker, K.K., and Pusey, A.E. (2016). Chimpanzee females queue but males compete for social status. *Sci. Rep.* 6, 35404. <https://doi.org/10.1038/srep35404>.
- Fox, J., and Weisberg, S. (2011). *An (R) Companion to Applied Regression*, Second Edition (Sage).
- Geissmann, T. (1986). Mate change enhances duetting activity in the Siangam gibbon (*Hylobates syndactylus*). *Behaviour* 96, 17–27. <https://doi.org/10.1163/156853986X00199>.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biol. Rev.* 77, 57–76. <https://doi.org/10.1017/S1464793101005826>.
- Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A.D., Wittig, R.M., and Crockford, C. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Commun. Biol.* 5, 1–15. <https://doi.org/10.1038/s42003-022-03350-8>.
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior* (Harvard University Press).
- Goodall, J. (1989). *Glossary of Chimpanzee Behaviors* (Jane Goodall Institute).
- Gultekin, Y.B., Hildebrand, D.G.C., Hammerschmidt, K., and Hage, S.R. (2021). High plasticity in marmoset monkey vocal development from infancy to adulthood. *Sci. Adv.* 7, eabf2938. <https://doi.org/10.1126/sciadv.abf2938>.
- Hollén, L.I., Clutton-Brock, T., and Manser, M.B. (2008). Ontogenetic changes in alarm-call production and usage in meerkats (*Suricata suricatta*): adaptations or constraints? *Behav. Ecol. Sociobiol.* 62, 821–829. <https://doi.org/10.1007/s00265-007-0508-x>.
- Hollén, L.I., and Radford, A.N. (2009). The development of alarm call behaviour in mammals and birds. *Anim. Behav.* 78, 791–800. <https://doi.org/10.1016/j.anbehav.2009.07.021>.
- International Phonetic Association (2022). <https://www.internationalphoneticassociation.org/>.
- Hultsch, H., and Todt, D. (1989). Memorization and reproduction of songs in nightingales (*Luscinia megarhynchos*): evidence for package formation. *J. Comp. Physiol.* 165, 197–203.
- Hurford, J.R. (2012). *The Origins of Grammar: Language in the Light of Evolution II* (OUP).
- Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends Cogn. Sci.* 3, 272–279. [https://doi.org/10.1016/S1364-6613\(99\)01333-9](https://doi.org/10.1016/S1364-6613(99)01333-9).
- Janik, V.M., and Slater, P.J. (1997). Vocal learning in mammals. *Adv. Study Behav.* 26, 59–100.
- Janik, V.M., and Slater, P.J.B. (2000). The different roles of social learning in vocal communication. *Anim. Behav.* 60, 1–11. <https://doi.org/10.1006/anbe.2000.1410>.
- Janik, V.M., Sayigh, L.S., and Wells, R.S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proc. Natl. Acad. Sci. USA* 103, 8293–8297. <https://doi.org/10.1073/pnas.0509918103>.
- Jansen, D.A., Cant, M.A., and Manser, M.B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biol.* 10, 97. <https://doi.org/10.1186/1741-7007-10-97>.
- Jarvis, E.D. (2019). Evolution of vocal learning and spoken language. *Science* 366, 50–54. <https://doi.org/10.1126/science.aax0287>.
- Kalan, A.K., Mundry, R., and Boesch, C. (2015). Wild chimpanzees modify food call structure with

- respect to tree size for a particular fruit species. *Anim. Behav.* 101, 1–9. <https://doi.org/10.1016/j.anbehav.2014.12.011>.
- Kalan, A.K., Piel, A.K., Mundry, R., Wittig, R.M., Boesch, C., and Kühl, H.S. (2016). Passive acoustic monitoring reveals group ranging and territory use: a case study of wild chimpanzees (*Pan troglodytes*). *Front. Zool.* 13, 34. <https://doi.org/10.1186/s12983-016-0167-8>.
- Koda, H., Oyakawa, C., Kato, A., and Masataka, N. (2007). Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour* 144, 681–692.
- Laidre, M.E., and Johnstone, R.A. (2013). Animal signals. *Curr. Biol.* 23, R829–R833. <https://doi.org/10.1016/j.cub.2013.07.070>.
- Lameira, A.R., Delgado, R.A., and Wich, S.A. (2010). Review of geographic variation in terrestrial mammalian acoustic signals: human speech variation in a comparative perspective. *J. Evol. Psychol.* 8, 309–332. <https://doi.org/10.1556/jep.8.2010.4.2>.
- Lameira, A.R., Santamaria-Bonfil, G., Galeone, D., Gamba, M., Hardus, M.E., Knott, C.D., Morrogh-Bernard, H., Nowak, M.G., Campbell-Smith, G., and Wich, S.A. (2022). Sociality predicts orangutan vocal phenotype. *Nat. Ecol. Evol.* 1–9. <https://doi.org/10.1038/s41559-022-01689-z>.
- Laporte, M.N.C., and Zuberbühler, K. (2010). Vocal greeting behaviour in wild chimpanzee females. *Anim. Behav.* 80, 467–473. <https://doi.org/10.1016/j.anbehav.2010.06.005>.
- Lattenkamp, E.Z., Vernes, S.C., and Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *J. Exp. Biol.* 221, jeb180729. <https://doi.org/10.1242/jeb.180729>.
- Lemoine, S., Preis, A., Samuni, L., Boesch, C., Crockford, C., and Wittig, R.M. (2020). Between-group competition impacts reproductive success in wild chimpanzees. *Curr. Biol.* 30, 312–318.e3. <https://doi.org/10.1016/j.cub.2019.11.039>.
- Leroux, M., and Townsend, S. (2020). Call combinations in great apes and the evolution of syntax. *Anim. Behav. Cogn.* 7, 131–139.
- Leroux, M., Bosshard, A.B., Chandia, B., Manser, A., Zuberbühler, K., and Townsend, S.W. (2021). Chimpanzees combine pant hoots with food calls into larger structures. *Anim. Behav.* 179, 41–50. <https://doi.org/10.1016/j.anbehav.2021.06.026>.
- Marshall, A.J., Wrangham, R.W., and Arcadi, A.C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Anim. Behav.* 58, 825–830. <https://doi.org/10.1006/anbe.1999.1219>.
- McComb, K., and Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biol. Lett.* 1, 381–385. <https://doi.org/10.1098/rsbl.2005.0366>.
- Menzel, E.W. (1964). Patterns of responsiveness in chimpanzees reared through infancy under conditions of environmental restriction. *Psychol. Forsch.* 27, 337–365. <https://doi.org/10.1007/BF00421336>.
- Mielke, A., Samuni, L., Preis, A., Gogarten, J.F., Crockford, C., and Wittig, R.M. (2018). Bystanders intervene to impede grooming in Western chimpanzees and sooty mangabeys. *R. Soc. Open Sci.* 4, 171296. <https://doi.org/10.1098/rsos.171296>.
- Mitani, J.C., Hasegawa, T., Gros-Louis, J., Marler, P., and Byrne, R. (1992). Dialects in wild chimpanzees? *Am. J. Primatol.* 27, 233–243. <https://doi.org/10.1002/ajp.1350270402>.
- Mitani, J.C., Gros-Louis, J., and Macedonia, J.M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *Int. J. Primatol.* 17, 569–583. <https://doi.org/10.1007/BF02735192>.
- Mitani, J.C., Hunley, K.L., and Murdoch, M.E. (1999). Geographic variation in the calls of wild chimpanzees: a reassessment. *Am. J. Primatol.* 47, 133–151. [https://doi.org/10.1002/\(SICI\)1098-2345\(1999\)47:2<133::AID-AJP4>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1098-2345(1999)47:2<133::AID-AJP4>3.0.CO;2-1).
- Moran, S., McCloy, D., and Wright, R. (2012). Revisiting population size vs. phoneme inventory size. *Language* 88, 877–893.
- Neumann, C., Assahad, G., Hammerschmidt, K., Perwitasari-Farajallah, D., and Engelhardt, A. (2010). Loud calls in male crested macaques, *Macaca nigra*: a signal of dominance in a tolerant species. *Anim. Behav.* 79, 187–193. <https://doi.org/10.1016/j.anbehav.2009.10.026>.
- Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Inwan, A.M., Agil, M., Widdig, A., and Engelhardt, A. (2011). Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Anim. Behav.* 82, 911–921. <https://doi.org/10.1016/j.anbehav.2011.07.016>.
- Nishida, T., Kano, T., Goodall, J., McGrew, W.C., and Nakamura, M. (1999). Ethogram and ethnography of Mahale chimpanzees. *Anthropol. Sci.* 107, 141–188.
- Notman, H., and Rendall, D. (2005). Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Anim. Behav.* 70, 177–190. <https://doi.org/10.1016/j.anbehav.2004.08.024>.
- Nowicki, S., and Searcy, W.A. (2014). The evolution of vocal learning. *Curr. Opin. Neurobiol.* 28, 48–53. <https://doi.org/10.1016/j.conb.2014.06.007>.
- Podos, J., Nowicki, S., and Peters, S. (1999). Permissiveness in the learning and development of song syntax in swamp sparrows. *Anim. Behav.* 58, 93–103. <https://doi.org/10.1006/anbe.1999.1140>.
- R Core Team (2018). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
- Reader, S.M., and Laland, K.N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl. Acad. Sci. USA* 99, 4436–4441. <https://doi.org/10.1073/pnas.062041299>.
- Reynolds, V. (2005). *The Chimpanzees of the Budongo Forest: Ecology, Behaviour and Conservation* (Oxford University Press).
- Santos, E.S., Scheck, D., and Nakagawa, S. (2011). Dominance and plumage traits: meta-analysis and metaregression analysis. *Anim. Behav.* 82, 3–19.
- Schaik, C.P. van, Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S., and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science* 299, 102–105. <https://doi.org/10.1126/science.1078004>.
- Setchell, J.M., and Wickings, E.J. (2005). Dominance, status signals and coloration in male mandrills (*Mandrillus sphinx*). *Ethology* 111, 25–50. <https://doi.org/10.1111/j.1439-0310.2004.01054.x>.
- Sewall, K.B., Young, A.M., and Wright, T.F. (2016). Social calls provide novel insights into the evolution of vocal learning. *Anim. Behav.* 120, 163–172. <https://doi.org/10.1016/j.anbehav.2016.07.031>.
- Seyfarth, R.M., and Cheney, D.L. (1980). The ontogeny of vervet monkey alarm calling behavior: a preliminary report. *Z. Tierpsychol.* 54, 37–56. <https://doi.org/10.1111/j.1439-0310.1980.tb01062.x>.
- Seyfarth, R.M., and Cheney, D.L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain Lang.* 115, 92–100. <https://doi.org/10.1016/j.bandl.2009.10.003>.
- Shapiro, A.D., Slater, P.J., and Janik, V.M. (2004). Call usage learning in gray seals (*Halichoerus grypus*). *J. Comp. Psychol.* 118, 447.
- Smit, W. (1977). *The Behavior of Communicating: An Ethological Approach*.
- Snowdon, C.T., and Elowson, A.M. (1999). Pygmy marmosets modify call structure when paired. *Ethology* 105, 893–908. <https://doi.org/10.1046/j.1439-0310.1999.00483.x>.
- Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim. Behav.* 55, 673–687. <https://doi.org/10.1006/anbe.1997.0602>.
- Suzuki, T.N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nat. Commun.* 7, 10986. <https://doi.org/10.1038/ncomms10986>.
- Suzuki, T.N., Wheatcroft, D., and Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS Biol.* 16, e2006532. <https://doi.org/10.1371/journal.pbio.2006532>.
- Townsend, S.W., Slocombe, K.E., Emery Thompson, M., and Zuberbühler, K. (2007). Female-led infanticide in wild chimpanzees. *Curr. Biol.* 17, R355–R356.
- Townsend, S.W., Hollén, L.I., and Manser, M.B. (2010). Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. *Anim. Behav.* 80, 133–138. <https://doi.org/10.1016/j.anbehav.2010.04.010>.
- Townsend, S.W., Engesser, S., Stoll, S., Zuberbühler, K., and Bickel, B. (2018). Compositionality in animals and humans. *PLoS Biol.* 16, 1–7. <https://doi.org/10.1371/journal.pbio.2006425>.

- van Noordwijk, M.A., van Schaik, C.P., and Wich, S.A. (2006). Innovation in wild Bornean orangutans (*Pongo pygmaeus wurmbii*). *Behaviour* 143, 839–876. <https://doi.org/10.1163/156853906778017944>.
- van Schaik, C.P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S., and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science* 299, 102–105. <https://doi.org/10.1126/science.1078004>.
- Vernes, S.C., and Wilkinson, G.S. (2020). Behaviour, biology and evolution of vocal learning in bats. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190061. <https://doi.org/10.1098/rstb.2019.0061>.
- Vernes, S.C., Kriengwatana, B.P., Beeck, V.C., Fischer, J., Tyack, P.L., ten Cate, C., and Janik, V.M. (2021). The multi-dimensional nature of vocal learning. *Philos. Trans. R. Soc. B Biol. Sci.* 376, 20200236. <https://doi.org/10.1098/rstb.2020.0236>.
- Watson, S.K., Townsend, S.W., Schel, A.M., Wilke, C., Wallace, E.K., Cheng, L., West, V., and Slocombe, K.E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* 25, 495–499. <https://doi.org/10.1016/j.cub.2014.12.032>.
- Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E.G., Wrangham, R.W., and Boesch, C. (1999). Cultures in chimpanzees. *Nature* 399, 682–685. <https://doi.org/10.1038/21415>.
- Wilbrecht, L., and Nottebohm, F. (2003). Vocal learning in birds and humans. *Ment. Retard. Dev. Disabil. Res. Rev.* 9, 135–148. <https://doi.org/10.1002/mrdd.10073>.
- Wilson, M.L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I.C., Hashimoto, C., Hobaiter, C.L., Hohmann, G., Itoh, N., Koops, K., et al. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature* 513, 414–417. <https://doi.org/10.1038/nature13727>.
- Wittig, R.M., and Boesch, C. (2003). Food competition and linear dominance hierarchy among female chimpanzees of the Tai National Park. *Int. J. Primatol.* 24, 847–867. <https://doi.org/10.1023/A:1024632923180>.
- Wittig, R.M., Crockford, C., Langergraber, K.E., and Zuberbühler, K. (2014). Triadic social interactions operate across time: a field experiment with wild chimpanzees. *Proc. R. Soc. B Biol. Sci.* 281, 20133155. <https://doi.org/10.1098/rspb.2013.3155>.
- Wittig, R.M., and Boesch, C. (2019). Observation protocol and long-term data collection in Tai. In *The Chimpanzees of the Tai Forest: 40 Years of Research*, C. Boesch, R.M. Wittig, C. Crockford, L. Vigilant, T. Deschner, and F.H. Leendertz, eds. (Cambridge University Press), pp. 44–57.
- Wright, T.F. (1996). Regional dialects in the contact call of a parrot. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 867–872. <https://doi.org/10.1098/rspb.1996.0128>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
R version 3.6.2	R Core Team (2018)	https://cran.r-project.org/bin/windows/base/old/3.6.2/
Lme4 package in R	Bates et al. (2015)	https://cran.r-project.org/web/packages/lme4/index.html

RESOURCE AVAILABILITY

Lead contact

The data used for this study are available from the lead contact upon request, Dr. Cédric Girard-Buttoz (cedric.buttoz@gmail.com).

Materials availability

This study did not generate new unique reagents or genetic sequences.

Data and code availability

This paper does not report original code.

Any additional information required to reanalyse the data reported in this paper is available from the [lead contact](#) upon request, Dr. Cédric Girard-Buttoz (cedric.buttoz@gmail.com).

Study site and subjects: The study was conducted on four chimpanzee communities belonging to two populations. One community of eastern chimpanzees, the Sonso community in the Budongo Forest ([Reynolds, 2005](#)) and three western chimpanzee communities in the Taï National Park within the Taï Chimpanzee Project ([Wittig and Boesch, 2019](#)), the East, North and South communities.

CC, ML, RMW and TB recorded vocalisations from known adult male and female chimpanzees (>12 years of age, 15 females and 11 males in Budongo and 33 females and 19 males in Taï) in these four communities from two populations. In the Sonso community, ML collected recordings in 2007–2008 and 2012 and CC and RMW collected recordings in 2008–2009. In Taï, CC collected recordings in the North community in 1999 and 2000 and TB in the East, North and South communities in 2019 and 2020.

METHOD DETAILS

Vocalisation recordings

All observers recorded chimpanzee vocalisations from 5 to 20 m using either a short ME65, ME66 or MKH 416, or long gun ME67 or MKH 418 Sennheisser directional microphones attached to Sony WMD6C professional Walkman (1998–2000), a Marantz PMD-660 solid-state recorder (2008–2010), a Marantz PMD-661 solid-state recorder in 2012 or a Tascam DR-40X recorder (2019–2020). In all cases, the observers recorded vocalisations both during focal animal sampling and *ad libitum* sampling. For each vocalisation, we recorded the context of production, and in particular whether or not the call was uttered towards a higher-ranking individual, during an approach between two individuals, during a fusion between two parties and/or as part of communication between different parties (see below for definitions). We also recorded whether the recipient of the call (whenever identifiable) aggressed the caller before the call was uttered. Here we considered only aggression comprising locomotion or contact aggression (see details about the aggression included below in ‘Behavioural Observations’).

Vocal coding

For this study, we combined several population datasets for Taï 1998–2000 and 2018–2020 and Sonso 2007–2010 and 2012 (see above). The vocal data in these datasets were coded in the same way using

examinations of the spectrogram of each recording using PRAAT software. We differentiated call types on the spectrogram using their distinctive temporal and spectral features (see [Girard-Buttoz et al., 2022](#) for details). For the analysis, we considered only calls of high quality, with the lowest frequency band visible, recorded from the beginning to the end of the vocal sequence, and with the signaller ID defined. We considered all calls which were either pant-hoot emitted alone (i.e., not in a pant-hoot sequence with build-up phase and climax but only panted hoots, i.e. only the build-up phase of the pant-hoot sequence without the introduction or the climax, [Figure S1A](#)), pant-grunt emitted alone ([Figure S1B](#)) or the combination of the two calls into a greeting hoot sequence ([Figures 1](#) and [S2](#)). While we refer to PH and PG as single call type in our study, it is important to note that they both could be considered as sequences in themselves since they both comprise a repetition of a single call type (the hoo for PH and the grunt for PG) with inhaled pants produced in between. We defined a greeting hoot vocalisation as a vocal sequence containing PG and PH elements ([Figures 1](#) and [S2A–S2C](#)) adjacent and at the start of the sequence. PG and PH can occur in sequences with a range of other call types. Given that it is not yet clear how sequences containing these elements might influence contextual information conveyed, nor in which contexts they are associated, we only accepted these elements within vocal sequences if they occurred *after* the two elements of interest in this study, PH_PG + X or PG_PH + X (see examples [Figures S2D](#) and [S2E](#) and [Table S1](#) for an exhaustive list of the sequences included in the analyses), i.e., we excluded sequences when some other call types were produced at the beginning of the sequence or between the PG and PH elements (i.e. X + PH_PG or X + PG_PH). Following these selection criteria, our final dataset comprised 40 individuals who uttered greeting hoots (4 males and 5 females in Tai East, 2 males and 11 females in Tai North, 3 males and 8 females in Tai South and 1 male and 7 females in Sonso) and for whom we had all the information on the context of production, the party size at the time of calling and 38 individual for whom we also had information on the rate of aggressions received (see below).

Behavioral observations

In order to assess the social environment in each study population and in particular the risk of aggression, we used focal behavioural data collected by the local field assistants and CC, ML and TB during the same field seasons during which the vocal data were recorded. We complemented our aggression data using focal observations from MG collected in 2019 and 2020 in Tai. We compiled all the aggressions from or towards an adult individual. We considered only aggressions likely to lead to an injury as being likely to lead to changes in vocal behaviour i.e., those that comprised locomotion and or were contact aggression (chase, charge, directed display, hit, bite, jump on, and push following the definition by [Goodall, 1989](#); [Nishida et al., 1999](#)). Although observations were recorded from different sites and time periods, aggression comprising locomotion or contact aggression was an integral part of each researchers' data collection protocol. Furthermore, even though direct inter-observer reliability could not be conducted between the different observers in our study, these behaviors are prone to high levels of inter-observer reliability (over 0.8 across tens of different observers in Tai, [Wittig and Boesch, 2019](#)). Also, one researcher, CC, collected data in both sites using the same protocol. Based on the total focal observation time for each individual, we compiled a rate of aggression given and received per hour. In addition, during these focal observation each observer recorded every pant-grant vocalisation uttered by any individual in the party as well as the identity of the caller and recipient of the vocalisation.

Dominance hierarchy

For all the datasets, we compiled dominance hierarchies based on unidirectional submissive pant-grunt vocalization (given by the lower ranking of the two individual towards the higher ranking, [Bygott, 1979](#)). For Budongo and the Tai period from 1999 to 2000, we used the hierarchies published in the relevant studies (see [Laporte and Zuberbühler, 2010](#); [Wittig et al., 2014](#); [Wittig and Boesch, 2003](#) for details). In all these studies, the hierarchy was calculated using I&S method ([de Vries, 1998](#)). For the Tai period 2019–2020, we took advantage of the long term data of the project allowing us to have a long buffer period of several years before the study period to compile dominance rank using a modified version of the Elo-rating method [Neumann et al. \(2011\)](#) developed by [Foester et al. \(2016\)](#). Please see [Mielke et al., 2018](#) for details).

QUANTIFICATION AND STATISTICAL ANALYSIS

Context of production for PG, PH and GH

We used two Generalised Linear Mixed Models (GLMMs) to test for differences in context of production between pant-grunts alone (PG) and greeting hoot (GH, Model 1a) and between pant-hoot alone (PH)

and GH (Model 1b, see [Figure S1](#) for spectrograms of PH and PG produced alone). For each model, we used the call type as the response and a binomial error distribution. As the response, GH = 1 and PG = 0 (Model 1a) or PH = 0 (Model 1b). Each call sequence constituted a data point. In both models, we used four categorical test predictors corresponding to the four contexts of interest: 1) **fusion** (Y/N) whether the call was emitted during a fusion (when two parties that were apart for at least 1 h reunited), 2) **inter-party communication** (Y/N) whether the call was produced within 1 min of a call given by another party (i.e. by another individual out of sight of the focal individual), 3) **approach** (Y/N) whether the call was given while the caller approached within 5 m of another conspecific and 4) **caller subordinate to the receiver** (three levels, 'yes', 'no', and 'no receiver').

In a third model (Model 1c), we tested whether greeting hoot sequences (GH) starting with a PH differed in context from GH starting with a PG. For this model, we used the same predictor as for Model 1a and 1b and the response was 1 for greeting hoots with PH first and 0 for greeting hoots with PG first. In models 1a, 1b and 1c we included individual identity and community ID as a random factor to avoid pseudoreplication.

Group differences in ordering of PG and PH within GH

We used another GLMM to test for differences in the ordering of PG and PH when combined together into a greeting hoot sequence (GH, Model 2). Specifically, using a binomial response (yes/no) we tested whether PH was emitted before (yes) or after (no) the PG within the GH sequence. As test predictors, we included community identity with four levels (Sonso, Tai East, Tai North and Tai South), the average rate of aggression received by the caller during the study period and the party size at the time the call was uttered. We used 'community identity' rather than 'population' (i.e., rather than Budongo vs. Tai) to also test for community differences within the same population (within Tai) and establish if the difference in ordering are population or community specific. In this model we also included individual identity as a random factor to avoid pseudoreplication.

All analyses were conducted in R 3.6.2 ([R Core Team, 2018](#)) using the function *glmer* from the package "lme4" ([Bates et al., 2015](#)). In Models 1a, 1b, and 1c we could not include random slopes between our predictors and the random factor individual identity since less than half of the individuals had data in all the level of each predictor (e.g. less than half of the individuals had both calls recorded during fusion and non-fusion context or during approach and non-approach contexts etc.) In Model 2, we included the random slope of party size within individual identity ([Baayen et al., 2008](#); [Barr et al., 2013](#)) but not the correlation between the random intercept and the slope. We tested each full model against its corresponding null model using a Likelihood Ratio Test (LRT, [Dobson, 2002](#)). Since we did not have control predictors, for each model, the null model comprised only the intercept and the random effect of individual identity.

We then assessed the significance of each predictor variable using a LRT between the full model and a reduced model comprising all the variables except the one being evaluated. This process was repeated across all variables, one by one, using the *drop1* function of the "lme4" package.

Before fitting each model, we tested for collinearity issues between our predictor variables by computing the variance inflation factor (VIF) using the function *vif* from the package "car" ([Fox and Weisberg, 2011](#)). Collinearity was not an issue in any of the final models (VIF of all predictor variables <1.2). We also assessed model stability by removing one level of each random effect at a time and recalculating the estimates of the different predictors which revealed that the results were stable. For each model, we calculated the marginal R^2 (i.e., the variance explained by the fixed effects) and the conditional R^2 (i.e., the variance explained by the entire model including both fixed and random effects) using the function *r.squaredGLMM* of the package "MuMin" ([Barton, 2020](#)).