Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning

Julia Fischer[1,2,3*], Franziska Wegdell[1,3], Franziska Trede[1,4], Federica Dal Pesco[1,3] & Kurt Hammerschmidt[1,3]

[1]Cognitive Ethology Laboratory, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany

[2]Department of Primate Cognition, Georg August University Göttingen, Göttingen, Germany

[3]Leibniz ScienceCampus Primate Cognition, Göttingen, Germany

[4]Primate Genetics Laboratory, German Primate Center, Göttingen, Germany

Corresponding author: jfischer@dpz.eu

ORCIDs

Julia Fischer 0000-0002-5807-0074

Franziska Wegdell 0000-0002-3108-2999

Franziska Trede 0000-0003-3690-1006

Federica Dal Pesco 0000-0003-2326-1185

Kurt Hammerschmidt 0000-0002-3430-2993

ABSTRACT

The extent to which nonhuman primate vocalizations are amenable to modification through experience is relevant for understanding the substrate from which human speech evolved. We examined the vocal behaviour of Guinea baboons, *Papio papio*, ranging in the Niokolo Koba National Park in Senegal. Guinea baboons live in a multi-level society, with units nested within parties nested within gangs. We investigated whether the acoustic structure of grunts of 27 male baboons of two gangs varied with party/gang membership and genetic relatedness. Males in this species are philopatric, resulting in increased male relatedness within gangs and parties. Grunts from members were more similar within than across social levels (N = 351 dyads for comparison within and between gangs, and N = 169 dyads within and between parties), but the effect sizes were small. Yet, acoustic similarity did not correlate with genetic relatedness, suggesting that higher amounts of social interactions rather than genetic relatedness promote the observed vocal convergence. We consider this convergence a result of sensory-motor integration and suggest this to be an implicit form of vocal learning shared with humans, in contrast to the goal-directed and intentional explicit form of vocal learning unique to human speech acquisition.

KEY WORDS: *Papio*, primate communication, vocal learning, sensory-motor integration, speech evolution, implicit learning

INTRODUCTION

One of the key preconditions for the development of speech is the ability to adjust vocal output in response to auditory input. Humans are exceptionally proficient at vocal learning. Although effortless speech learning is confined to the early years [1], humans still possess the ability to imitate sounds voluntarily and acquire further languages throughout their lives. Numerous comparative studies have aimed at elucidating the evolutionary origins of vocal learning within the primate lineage, to uncover the extent to which nonhuman primates reveal evidence for vocal plasticity, and whether such plasticity may be conceived as a pre-adaptation for the evolution of speech [2,3].

Despite considerable research effort, it appears that the ability to learn sounds from auditory experience in most nonhuman primate species is limited. Unlike some humans or some songbird species, nonhuman primates are not obligatory vocal learners that require species-specific auditory input to develop their normal vocal repertoires [4,5]. Early attempts to train a young chimpanzee to produce speech sounds yielded disappointing results and prompted most of the ‘ape language’ projects to turn to another modality, using either symbol systems or sign languages [6]. Studies of the neural basis of vocal production in different monkey species found that the animals lack the neural connections necessary for the volitional control over the fine structure of vocalizations, although they exert greater control over the usage of calls [reviewed in 2]. One exception to the rule of limited vocal plasticity may be orangutans that have greater control over their vocal apparatus [7,8].

In addition to the limited ontogenetic plasticity, a range of comparative studies within different nonhuman primate species strongly suggests that the motor patterns underlying vocalizations are evolutionarily highly conserved within genera [reviewed in 2]. For instance, the structure of alarm calls of members of the genus *Chlorocebus* differs only marginally between East African vervets, *Chlorocebus pygerythrus* and West African green monkeys, *Chlorocebus sabaeus.* Moreover, in response to a drone naïve West African green monkeys

spontaneously uttered calls that structurally were highly similar to East African vervet “eagle alarm calls”, indicating that the link between the perception of a specific (potentially dangerous) stimulus and the activation of a given motor program is also conserved [9].

At the same time, subtle modifications in vocal output as a result of auditory experience appear to be possible. For instance, common marmosets, *Callithrix jacchus,* increase the amplitude of their calls in noisy environments [“Lombard effect”; ,10,11]. More importantly, a range of species shows group-specific variations or ‘dialects’ in their vocalizations [reviewed in 12], while Japanese macaques matched some of the acoustic features of calls presented in playbacks [13]. These instances of vocal plasticity have been described as “vocal accommodation” [12,14,15] or “social shaping” [15], similar to the observation that humans may involuntarily match the pitch, temporal patterning and prosody of the people they are talking to.

Following the idea that auditory input may lead to vocal convergence, subjects that interact more frequently with one another using vocalizations should produce calls that are more similar to each other than those that interact less frequently. A higher acoustic similarity may also result from genetic relatedness, however. For instance, highly related subjects may also have a similar morphology of the vocal production apparatus [16]. Before conclusions about the role of experience can be drawn, it is necessary to assess whether potential acoustic variation between individuals can (also) be explained by genetic distance.

To date, only few studies investigated the effects of genetic relatedness and interaction frequency at the same time. Lemasson and colleagues reported that interaction frequency but not genetic relatedness accounted for acoustic variation in the calls of Campbell’s monkeys, *Cercopithecus campbelli campbelli* [17]. The reported correlation of acoustic similarity with grooming frequency may be spurious, however, as data from two groups (with N = 6 and 4 females, respectively) were pooled, and the correlation was largely

driven by the differences between the groups. Levréro and colleagues [15] studied the acoustic structure of contact calls in 36 male and female mandrills living in three social groups. Both genetic relatedness and familiarity impacted acoustic similarity of the species' 'kakak' calls, while retaining cues to kin memberships: playback experiments showed that subjects responded significantly more strongly to calls recorded from related kin, irrespective of familiarity [15].

We here set out to assess the impact of social interaction while controlling for genetic relatedness by comparing the acoustic variation in the grunts of male Guinea baboons, *Papio papio*. Guinea baboons are an interesting model to examine the influence of auditory experience and social group membership, as they live in a nested multi-level society with male philopatry [18]. At the base of the society are 'units' comprising one adult male, 1-6 females, and young. A small number of units, together with bachelor males, form a 'party', and 2-3 parties make up a 'gang' (Fig. 1a). Assignments to parties and gangs are based on spatial proximity and affiliative interactions [19]. During affiliative interactions with other group members, males produce low-frequency tonal grunts (Fig. 1b). Their social structure allowed us to assess vocal convergence at two social levels, namely within parties and within gangs.

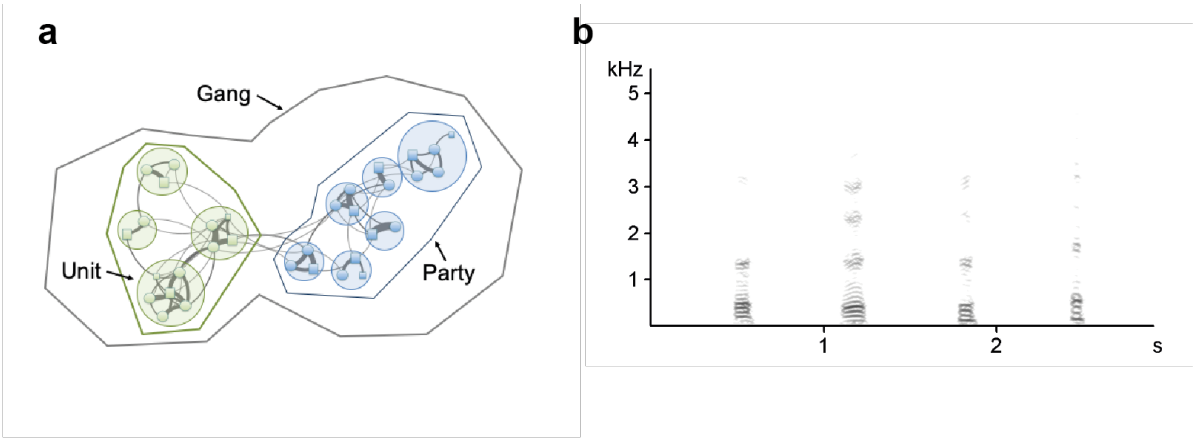

Figure 1. (a) The multilevel social organisation of Guinea baboons. Several units form a party, and two or more parties form a gang. (b) Spectrogram of grunts from four different males. Frequency (kHz) on the y-axis, time (s) on the x-axis. The spectrogram was created

using Avisoft-SASLab Pro 5.2 (1.024 pt FFT, sampling frequency: 11 kHz, time resolution: 2.9 ms, Flat Top window).

If the frequency of interaction affects the structure of calls, subjects that interact frequently with one another should produce calls that are more similar to each other. Thus, members of the same party should have the greatest similarity, while members of the same gang should produce calls that are more similar to each other than to calls produced by members of another gang. If genetic relatedness affects the vocal structure, dyads that are more highly related should reveal greater acoustic similarity. Note that these two effects (interaction frequency and relatedness) are not mutually exclusive.

METHODS

We obtained recordings of grunts from a total of 27 male baboons in 2010/11, 2014, and 2016. Thirteen of the males were members of the 'Mare' gang and 14 were members of the 'Simenti' gang. The Mare gang comprised two parties of 6 and 7 males each; the Simenti gang comprised two parties of 5 and 9 males each. Twenty-three of the 27 males were confirmed or assumed to be present throughout the study period (see Electronic Supplementary Material Tab. S1).

Vocalizations were recorded using Marantz PMD 661 recorders (D&M Professional, Longford, U.K.) with Sennheiser directional microphones (K6 power module + ME66 recording head; Sennheiser, Wedemark, Germany) equipped with Rycote windshields (RycoteWindjammer, Rycote Ltd., Stroud, U.K.). We used Avisoft SASLab Pro (Avisoft Bioacoustic, Berlin, Germany) to check the recording quality and to label and extract grunts with sufficient quality and low background noise. We only used calls recorded at a maximum distance of 3 m. To maximize the number of grunts per male, we included grunts from different contexts (Electronic Supplementary Material Tab. S2). In total, we included 756 grunts in the acoustic analysis. On average, we used 28 calls per subject in the analysis

(range: 5-127). The Mare and Simenti gang males were represented by 390 and 366 grunts, respectively. Ideally, one would have liked to include further gangs with additional subjects to assess whether the observed pattern holds beyond our study population, but adding further groups was beyond our capacities.

We reduced the sampling frequency from 44.1 to 5 and 1 kHz to obtain an appropriate frequency resolution for the estimation of acoustic features and calculated two 1024 pt Fast Fourier Transformation (FFT), one resulting in a frequency range of 2500 Hz (frequency resolution 5 Hz, temporal resolution 6.4ms) and a second FFT resulting in a frequency range of 500 Hz (frequency resolution of 1 Hz, and a temporal resolution of 16 ms). Calculating two FFTs allowed us to maximize the temporal resolution for the entire call type, and estimate the fundamental frequency at a higher frequency resolution. The resulting frequency-time spectra were analysed with a custom software program LMA 2019, which allows visual control of the accuracy of parameter estimation [20,21]. LMA outputs a total of 82 acoustic parameters.

To identify which parameters would be informative to distinguish between individuals (and thus, social levels), we entered all 82 acoustic features from the LMA output into a stepwise discriminant function (DFA) with subject identity as a grouping variable. The selection criterion for acoustic features to enter the discriminant function was $P_{in} = 0.05$ and to remove $P_{out} = 0.1$. The DFA used 31 acoustic features for individual discrimination (Electronic Supplementary Material Tab. S3). To quantify the acoustic distance between males, we used the average pairwise F-value from the discriminant function analysis as a dissimilarity score for each dyad. This approach has been applied in different studies examining relationships between acoustic structure and genetic or geographic distance [22,23]. The discriminant function analysis was performed using IBM SPSS version 26.0 (IBM, Armonk, NY). To assess whether the classification result of the individual discrimination of male grunts is higher than would be expected by chance, we additionally

performed a permuted DFA [24], which controls for variation in individual contributions of grunts.

We extracted DNA from fecal samples using the First-DNA all tissue kit (Genial®) and characterized genetic variation by assessing the individual allele variation on 24 polymorphic autosomal microsatellite markers. The 24 markers were amplified using the Multiplex PCR Kit (QIAGEN) and fluorescent-labelled primers. PCR products were separated and detected through capillary gel electrophoresis on an ABI 3130xL Genetic Analyzer (Applied Biosystems®, USA). Microsatellite allele sizes were evaluated using Gene Mapper 5 (Applied Biosystems®). One locus (D1s548) showed signs of null alleles and significant deviations from Hardy-Weinberg equilibrium and was therefore excluded, resulting in a total of 23 loci included in the relatedness estimation (calculated with MICRO-CHECKER version 2.2.3 [25] and the PopGenReport R package version 3.0.0 [26]. We used the R package "related" version 1.0[27,28] to estimate relatedness using R version 3.4.4 and RStudio version 1.1.456. The Wang estimator (hereafter "W") appeared to be most suitable for the present analysis (see Electronic Supplementary Material Tab. S4). W ranges from -1 to 1. Negative values indicate that dyads are less related than on average, while positive values indicate that they are more highly related than on average [see 29 for a detailed description of the analysis].

These and the following statistical analyses were conducted in the R environment version 3.6.3 [30], using the RStudio interface version 1.3.959 [31]. We used a mantel matrix correlation test (package "vegan"; version 2.5.6) to test the correlation between acoustic and genetic variation. To test whether calls within a gang were more similar to each other than between gangs, we applied a categorical mantel test, using 'same gang membership' (Yes/No) as the categorical predictor variable, and W or F (transformed as ln(1+F)) as the continuous variable. The analysis of the effect of gang membership was based on 351 dyads. To study the effect of party membership, we also used a categorical mantel test, but

only considered pairs of males that lived in the same gang (e.g., SNE and BAA, both members of the Mare gang, or BEN and WLD, both members of the Simenti gang; total number of dyads within both gangs N = 169). We used a restricted permutation approach where males were permuted between parties within gangs. We used 1,000 permutations in all analyses, except the one for the variation between parties within gangs, where we used 10,000 permutations. Effect sizes were calculated with the package “effsize” version 0.8.0. The data and code for statistical analysis are deposited at https://osf.io/h7q5r/?view_only=e24fb7b53e7a4c57a1fe67db5d2452ab.

ETHICS STATEMENT

This research complies with the Association for the Study of Animal Behaviour Guidelines for the Use of Animals in Research (Animal Behaviour, 2018, 135, I-X), the legal requirements of the country in which the work was carried out, and all institutional guidelines. The research has been approved by the Direction des Parcs Nationaux of the Republic of Senegal (22/04/19).

RESULTS

Confirming previous analyses, males were more highly related within gangs than between gangs (Categorical Mantel Test, $P = 0.001$, $N = 351$, Figure 2a). The effect size (Cohen’s d) was 0.52 (CI_{lower} -0.73, CI_{upper} -0.31; medium effect size). Within gangs, males in the same party were more highly related on average than males that were not members of the same party ($P = 0.035$, $N = 169$, Figure 2b), with a small effect size ($d = 0.24$, CI_{lower} -0.54, CI_{upper} -0.07).

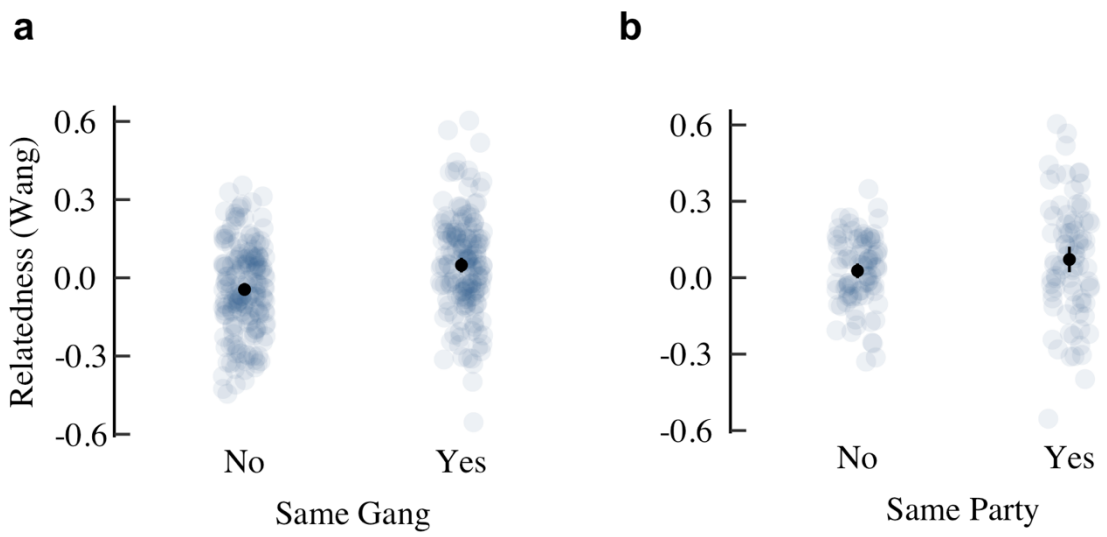

Figure 2. Genetic relatedness between male dyads that belong to (a) different gangs or the same gang, and (b) different parties or the same party within a gang. Light grey dots represent dyadic values, black dots the mean with 95% confidence interval.

Grunts could be assigned to the correct individual significantly more frequently than by chance, with an average correct assignment of 34.5 % using the procedure in SPSS (chance level 3.7%, leave-one-out validation: 21.0 % correct classification). The classification in the pDFA with a reduced set of N = 135 calls (see Electronic Supplementary Material Tab. S5 yielded an average classification of 11.2%, $P < 0.001$). Acoustic similarity did not correlate with genetic similarity (r = -0.006, P = 0.515). Because of the inherent uncertainty with which dyadic relatedness can be estimated [32], we ran an additional analysis in which we compared the acoustic dissimilarity of dyads in the top quartile (W > 0.125) vs. the bottom quartile (W < -0.117). Again, we found no effect of relatedness on acoustic dissimilarity (categorical Mantel test, P = 0.933, Fig. 3).

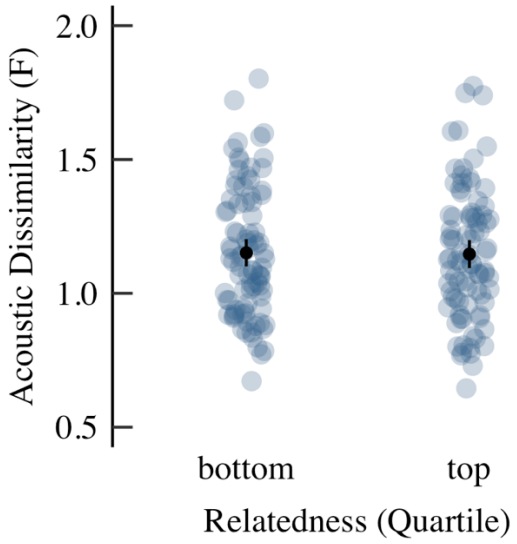

Figure 3. Relation between acoustic dissimilarity and genetic relatedness (top and bottom quartiles of the Wang estimator W) for N = 175 dyads. Light grey dots represent dyadic values, black dots the mean with 95% confidence interval. There were no significant differences between unrelated and related individuals (P = 0.933).

Grunts of males within gangs were more similar to each other than between gangs (categorical Mantel test, P = 0.012, Fig. 4a), and grunts of males within a party were also more similar to each other than between parties in the same gang (*P* = 0.001, Fig. 4b). The effect sizes were modest, however (*d* = 0.177, CI_{lower} -0.03, CI_{upper} 0.38 between gangs and 0.152, CI_{lower} -0.15, CI_{upper} 0.46 between parties, respectively). When we compared the mean acoustic dissimilarity of males that resided in the same party (mean logF = 0.29) to those that were part of a different gang (mean logF = 0.33), the effect size was small (*d* = 0.24, CI_{lower} -0.02, CI_{upper} 0.50). Grunts varied with social level (party/gang) mostly in parameters that are related to articulatory movements (Electronic Supplementary Material Tab. S6). The fundamental frequency or duration did not vary systematically between social levels.

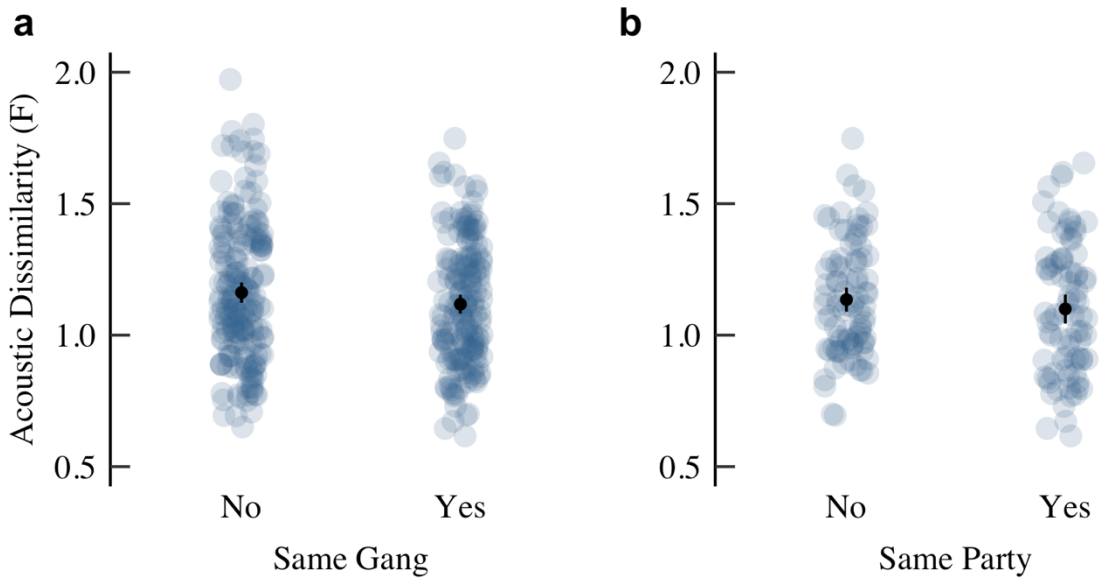

Figure 4. Acoustic dissimilarity of dyads that belong to (a) different gangs or the same gang, and (b) different parties or the same party within a gang. Light grey dots represent dyadic values, black dots the mean with 95% confidence interval. Calls from males in the same gang and the same party were on average more similar to each other than between gangs or parties.

DISCUSSION

The structure of male grunts varied between members of different gangs, and also between members of parties within a gang. The effect sizes of these two comparisons were modest, however. Males in the same gang were also more highly related to one another, but this did not account for the acoustic variation between parties and gangs, as evidenced by the lack of an effect of genetic relatedness on acoustic similarity. In this regard, the Guinea baboons differ from mandrills, where both relatedness and interaction frequency predicted the structure of the vocalizations [15].

It may seem puzzling at first that genetic relatedness did not account for the higher vocal similarity in Guinea baboons despite the fact that genetic relatedness and acoustic similarity were both higher within parties and gangs than between. This can be explained by

287 a combination of the fact that not all dyads within a social level are indeed more highly
288 related than across these social levels, and that acoustic similarity appears mainly to be
289 driven by social interaction, which is not restricted to highly related dyads. To a certain
290 degree, relatedness and acoustic similarity vary independently of one another.
291
292 How may auditory input affect vocal production? One mechanism that may support
293 the reported minor adjustments in vocal output with experience may rest on sensory-motor
294 integration [33]. According to the idea of a ‘common coding’ framework, specific sensorimotor
295 areas represent both sensory input and motor commands generating that corresponding
296 pattern [34]. In humans, neuroimaging studies identified specific motor activations when
297 subjects listened to speech sounds [35]. If such sensory-motor integration exists in the
298 auditory-vocal domain of nonhuman primates, the exposure to specific auditory input may
299 increase the likelihood to produce the corresponding motor pattern via co-activation.
300
301 A recent study provided compelling evidence for the integration of auditory input with
302 vocal output in a nonhuman primate species. In common marmosets, activity in the auditory
303 cortex directly affected the monkeys’ control of vocal production [36]. Firstly, a shift in the
304 auditory feedback of the monkey's vocalization led to compensatory changes in the
305 frequency patterns of the subsequent vocalizations. Secondly, microstimulation of the
306 auditory cortex during vocalization led to abrupt shifts in vocal frequency [36]. In a
307 translocation experiment, common marmosets (N = 4) adjusted the structure of their
308 vocalizations in response to auditory input from conspecifics, even if the individuals do not
309 interact directly [37]. Beyond the immediate effects of auditory experience, there is also
310 evidence that feedback from parents affects the trajectory of vocal development in
311 marmosets [38–40].
312
313 It has been argued that the human ability to imitate the utterances of others gradually
314 evolved from the vocal plasticity observed in nonhuman primates [17,41]. We contend that

vocal learning may be based on a variety of different mechanisms, including vocal convergence, ‘learning from success’, a form of usage learning that comprises the use of specific call variants because they are more likely to yield the desired response, as well as the spontaneous imitation of a recently formed auditory template [42]. Instead of conceiving vocal learning capacities as a continuum [43], we agree with other authors that vocal learning may be supported by a variety of different mechanisms [44]. Future studies should aim to distinguish between these mechanisms, and also consider the effect size of vocal plasticity.

Taking the extent of plasticity as well as the mechanisms that support them into account will contribute to overcoming futile debates about whether or not nonhuman primates reveal evidence for vocal learning [45]. The vast majority of studies in nonhuman primates that reported evidence for vocal convergence observed only minor changes within the species-specific range of calls. Thus, the small effect sizes reported here are an important aspect of the results. Humans, instead, are not only able to work on their accents, as Eliza Doolittle in ‘My fair lady’, but they can also sing “supercalifragilisticexpialidocious” with Mary Poppins. The spontaneous imitation of new words is open-ended, while it is much more difficult to change one’s accent once a certain age has been reached. Vocal convergence in nonhuman primates appears to be more similar to the formation of an accent than the acquisition of novel phonetic combinations that make up new words. An interesting open question is whether vocal convergence is simply a by-product of the sensory experience or whether it has been selected for, since it may signal ‘in-group’ membership and thus have an important social function [46].

Irrespective of whether vocal convergence has been selected for or not, we propose that it constitutes an implicit form of motor learning shared between nonhuman primates and humans, while speech production constitutes an explicit form of motor learning. Implicit and explicit processes are not entirely dichotomous; explicitly acquired motor skills can become automatic (as when you learn to drive a car), while implicit processes may be made explicit

[47]. Yet, it has proven useful to distinguish between implicit and explicit forms of knowledge and knowledge acquisition [48]. Taatgen suggested that implicit learning is a by-product of general learning mechanisms, while explicit learning is tied to learning goals and thus intentionality [47]. This definition appears useful for the distinction between vocal convergence as a result of sensory-motor integration on the one hand and the goal-directed acquisition of the patterns that result in the production of speech, on the other.

A further open question is whether the observed acoustic variation is salient to the animals themselves. In a previous study [49], we tested male responses to the playbacks of grunts of males that share the same home range as the study males (“neighbours”) vs. to grunts of males living 50 km away (“strangers”). As a control, we played back the grunts of males from their gang. Surprisingly, males responded strongly only to the grunts from males of their gang, but largely ignored neighbour or stranger males' calls. In principle, these responses could be explained by the recognition of the males' characteristics. Yet, it might also be the case that males recognize the 'sound' of their subgroup. Playbacks presenting artificially created grunts bearing the own gang’s characteristics vs. another gang’s characteristics would be needed to test this conjecture.

In summary, we find evidence for a moderate degree of vocal convergence in the grunts of male Guinea baboons. The magnitude of the change is difficult to compare to those of other studies on nonhuman primates mentioned above, given the differences in methodological approaches, but broadly appears to be in a similar range. Our findings add to the body of evidence that within species-specific constraints, subtle and potentially meaningful variation can be found in nonhuman primate vocalizations. This variation does not compare to the open-ended possibility of vocal imitation found in human speech, however.

Acknowledgments

We thank the Diréction des Parcs Nationaux (DPN) and Ministère de l'Environnement et de la Protéction de la Nature (MEPN) de la République du Sénégal for permission to work in the Niokolo Koba National Park. Roger Mundry provided invaluable statistical advice. This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 254142454 / GRK 2070, the Leibniz ScienceCampus Primate Cognition of the Leibniz Association, and the Christian-Vogel Fonds of the Gesellschaft für Primatologie.

Author contributions

JF, KH, and FW conceived the study, FW collected the data, FW and KH conducted the acoustic analysis, FDP and FT conducted the genetic analysis, JF and KH ran the statistical analyses and prepared the figures, JF wrote the manuscript with input from all authors.

References

1. Westermann G, Mani N. 2017 *Early Word Learning (Current Issues in Developmental Psychology Series)*. Oxford: Taylor & Francis.
2. Fischer J, Hage SR. 2019 Primate vocalization as a model for human speech: Scopes and limits. In *Human Language: From Genes and Brains to Behavior* (ed P Hagoort), pp. 639–656. Cambridge MA: MIT Press.
3. Jarvis ED. 2019 Evolution of vocal learning and spoken language. **54**, 50–54.
4. Nowicki S, Searcy WA. 2014 The evolution of vocal learning. *Current Opinion in Neurobiology* **28**, 48–53. (doi:10.1016/j.conb.2014.06.007)
5. Tyack PL. 2020 A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20180406-undefined.
6. Wallman J. 1992 *Aping Language*. Cambridge: Cambridge University Press.
7. Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker RW. 2009 A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–64. (doi:10.1007/s10329-008-0117-y)
8. Lameira AR, Hardus ME, Bartlett AM, Shumaker RW, Wich SA, Menken SBJ. 2015 Speech-like rhythm in a voiced and voiceless orangutan call. *PLoS ONE* **10**, e116136. (doi:10.1371/journal.pone.0116136)
9. Wegdell F, Hammerschmidt K, Fischer J. 2019 Conserved alarm calls but rapid auditory learning in monkey responses to novel flying objects. *Nature Ecology & Evolution* **3**, 1039–1042. (doi:10.1038/s41559-019-0903-5)
10. Brumm H, Voss K, Köllmer I, Todt D. 2004 Acoustic communication in noise: Regulation of call characteristics in a New World monkey. *Journal of Experimental Biology* **207**, 443–448. (doi:10.1242/jeb.00768)
11. Egnor SER, Hauser MD. 2006 Noise-induced vocal modulation in cotton-top tamarins (*Saguinus oedipus*). *American Journal of Primatology* **68**, 1183–1190.
12. Ruch H, Zürcher Y, Burkart JM. 2018 The function and mechanism of vocal accommodation in humans and other primates. *Biological Reviews* **93**, 996–1013. (doi:10.1111/brv.12382)
13. Sugiura H. 1998 Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour* **55**, 673–687.
14. Fischer J. 2008 Transmission of acquired information in nonhuman primates. In *Learning and Memory: A Comprehensive Reference* (eds R Menzel, J Byrne), pp. 299–313. Oxford: Elsevier. (doi:10.1016/B978-012370509-9.00055-3)
15. Levréro F, Carrete-Vega G, Herbert A, Lawabi I, Courtiol A, Willaume E, Kappeler PM, Charpentier MJE. 2015 Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communications* **6**, 7609. (doi:10.1038/ncomms8609)
16. Hauser MD. 1992 Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production? *The Journal of the Acoustical Society of America* **91**, 2175–2179. (doi:10.1121/1.403676)
17. Lemasson A, Ouattara K, Petit EJ, Zuberbühler K. 2011 Social learning of vocal structure in a nonhuman primate? *BMC Evolutionary Biology* **11**, 362-undefined. (doi:10.1186/1471-2148-11-362)
18. Fischer J *et al.* 2017 Charting the neglected West: The social system of Guinea baboons. *American Journal of Physical Anthropology* **162**, 15–31. (doi:10.1002/ajpa.23144)
19. Patzelt A, Kopp GH, Ndao I, Kalbitzer U, Zinner D, Fischer J. 2014 Male tolerance and male-male bonds in a multilevel primate society. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 14740–14745. (doi:10.1073/pnas.1405811111)
20. Fischer J, Noser R, Hammerschmidt K. 2013 Bioacoustic field research: A primer to acoustic analyses and playback experiments with primates. *American Journal of Primatology* **75**, 643–663. (doi:10.1002/ajp.22153)

21. Schrader L, Hammerschmidt K. 1997 Computer-aided analysis of acoustic parameters in animal vocalizations: A multi-parametric approach. *Bioacoustics* **7**, 247–265. (doi:10.1080/09524622.1997.9753338)
22. Thinh VN, Hallam C, Roos C, Hammerschmidt K. 2011 Concordance between vocal and genetic diversity in crested gibbons. *BMC Evolutionary Biology* **11**, 36. (doi:10.1186/1471-2148-11-36)
23. Meyer D, Hodges JK, Rinaldi D, Wijaya A, Roos C, Hammerschmidt K. 2012 Acoustic structure of male loud-calls support molecular phylogeny of Sumatran and Javanese leaf monkeys (genus *Presbytis*). *BMC Evolutionary Biology* **12**, 16. (doi:10.1186/1471-2148-12-16)
24. Mundry R, Sommer C. 2007 Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour* **74**, 965–976. (doi:10.1016/j.anbehav.2006.12.028)
25. van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004 MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538. (doi:10.1111/j.1471-8286.2004.00684.x)
26. Gruber B, Adamack AT. 2015 landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers. *Molecular Ecology Resources* **15**, 1172–1178. (doi:10.1111/1755-0998.12381)
27. Wang J. 2011 Coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11**, 141–145. (doi:10.1111/j.1755-0998.2010.02885.x)
28. Pew J, Muir PH, Wang J, Frasier TR. 2015 related: An R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources* **15**, 557–561. (doi:10.1111/1755-0998.12323)
29. Dal Pesco F, Trede F, Zinner D, Fischer J. 2020 Analysis of genetic relatedness and paternity assignment in wild Guinea baboons (*Papio papio*) based on microsatellites. *protocols.io* **dx.doi.org/10.17504**, undefined. (doi:dx.doi.org/10.17504/protocols.io.bfdmji46)
30. R Core Team. 2020 R: A Language and Environment for Statistical Computing.
31. RStudio Team. 2020 RStudio: Integrated Development for R.
32. van Horn RC, Altmann J, Alberts SC. 2008 Can't get there from here: inferring kinship from pairwise genetic relatedness. *Animal Behaviour*. **75**, 1173–1180. (doi:10.1016/j.anbehav.2007.08.027)
33. Pickering MJ, Garrod S. 2013 An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* **36**, 329–347. (doi:10.1017/S0140525X12001495)
34. Prinz W, Neumann O. 1990 A common-coding approach to perception and action. In *Relationships between perception and action: Current approaches*, pp. 167–201. Berlin & New York: Springer.
35. Pulvermüller F, Fadiga L. 2010 Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* **11**, 351–360. (doi:10.1038/nrn2811)
36. Eliades SJ, Tsunada J. 2018 Auditory cortical activity drives feedback-dependent vocal control in marmosets. *Nature Communications* **9**, 2540-undefined. (doi:10.1038/s41467-018-04961-8)
37. Zürcher Y, Willems EP, Burkart JM. 2019 Are dialects socially learned in marmoset monkeys? Evidence from translocation experiments. *PLoS ONE* **14**. (doi:10.1371/journal.pone.0222486)
38. Gultekin YB, Hage SR. 2017 Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nature Communications* **8**, 14046. (doi:10.1038/ncomms14046)
39. Gultekin YB, Hage SR. 2018 Limiting parental interaction during vocal development affects acoustic call structure in marmoset monkeys. *Science Advances* **4**, 11. (doi:10.1126/sciadv.aar4012)

40. Takahashi DY, Liao DA, Ghazanfar AA. 2017 Vocal learning via social reinforcement by infant marmoset monkeys. *Current Biology* **27**, 1844-1852.e6. (doi:http://dx.doi.org/10.1016/j.cub.2017.05.004)
41. Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, West V, Slocombe KE. 2015 Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology* **25**, 495–499. (doi:10.1016/j.cub.2014.12.032)
42. Fischer J, Hammerschmidt K. 2020 Towards a new taxonomy of primate vocal production learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20190045. (doi:10.1098/rstb.2019.0045)
43. Arriaga G, Jarvis ED. 2013 Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language* **124**, 96–116. (doi:10.1016/j.bandl.2012.10.002)
44. Martins PT, Boeckx C. 2020 Vocal learning: Beyond the continuum. *PLoS Biology* **18**, e3000672. (doi:10.1371/journal.pbio.3000672)
45. Wirthlin M *et al.* 2019 A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron* **104**, 87–99. (doi:10.1016/j.neuron.2019.09.036)
46. Crockford C, Herbinger I, Vigilant L, Boesch C. 2004 Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**, 221–243. (doi:10.1111/j.1439-0310.2004.00968.x)
47. Taatgen NA. 1999 *Learning without limits: from problem solving towards a unified theory of learning*. Groningen: University of Groningen.
48. Frensch PA, Rünger D. 2003 Implicit learning. *Current Directions in Psychological Science*. **12**, 13–18. (doi:10.1111/1467-8721.01213)
49. Maciej P, Patzelt A, Ndao I, Hammerschmidt K, Fischer J. 2013 Social monitoring in a multilevel society: A playback study with male Guinea baboons. *Behavioral Ecology and Sociobiology* **67**, 61–68. (doi:10.1007/s00265-012-1425-1)