

## Research



**Cite this article:** Raimondi T, Di Panfilo G, Pasquali M, Zarantonello M, Favaro L, Savini T, Gamba M, Ravignani A. 2023 Isochrony and rhythmic interaction in ape duetting.

*Proc. R. Soc. B* **290**: 20222244.

<https://doi.org/10.1098/rspb.2022.2244>

Received: 7 November 2022

Accepted: 8 December 2022

**Subject Category:**

Neuroscience and cognition

**Subject Areas:**

behaviour, cognition, neuroscience

**Keywords:**

song, gibbon, synchrony, rhythm, music

**Authors for correspondence:**

Teresa Raimondi

e-mail: [teresa.raimondi@unito.it](mailto:teresa.raimondi@unito.it)

Andrea Ravignani

e-mail: [andrea.ravignani@mpi.nl](mailto:andrea.ravignani@mpi.nl)

<sup>†</sup>Co-first authors.

<sup>‡</sup>Co-senior authors.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6350493>.

# Isochrony and rhythmic interaction in ape duetting

Teresa Raimondi<sup>1,†</sup>, Giovanni Di Panfilo<sup>1,†</sup>, Matteo Pasquali<sup>1</sup>,  
Martina Zarantonello<sup>1</sup>, Livio Favaro<sup>1</sup>, Tommaso Savini<sup>2,‡</sup>, Marco Gamba<sup>1,‡</sup> and  
Andrea Ravignani<sup>3,4,‡</sup>

<sup>1</sup>Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

<sup>2</sup>Conservation Ecology Program, King Mongkut University of Technology Thonburi, School of Bioresources and Technology, Bangkok, Thailand

<sup>3</sup>Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

<sup>4</sup>Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Denmark

LF, 0000-0002-8698-472X; AR, 0000-0002-1058-0024

How did rhythm originate in humans, and other species? One cross-cultural universal, frequently found in human music, is isochrony: when note onsets repeat regularly like the ticking of a clock. Another universal consists in synchrony (e.g. when individuals coordinate their notes so that they are sung at the same time). An approach to biomusicology focuses on similarities and differences across species, trying to build phylogenies of musical traits. Here we test for the presence of, and a link between, isochrony and synchrony in a non-human animal. We focus on the songs of one of the few singing primates, the lar gibbon (*Hylobates lar*), extracting temporal features from their solo songs and duets. We show that another ape exhibits one rhythmic feature at the core of human musicality: isochrony. We show that an enhanced call rate overall boosts isochrony, suggesting that respiratory physiological constraints play a role in determining the song's rhythmic structure. However, call rate alone cannot explain the flexible isochrony we witness. Isochrony is plastic and modulated depending on the context of emission: gibbons are more isochronous when duetting than singing solo. We present evidence for rhythmic interaction: we find statistical causality between one individual's note onsets and the co-singer's onsets, and a higher than chance degree of synchrony in the duets. Finally, we find a sex-specific trade-off between individual isochrony and synchrony. Gibbon's plasticity for isochrony and rhythmic overlap may suggest a potential shared selective pressure for interactive vocal displays in singing primates. This pressure may have convergently shaped human and gibbon musicality while acting on a common neural primate substrate. Beyond humans, singing primates are promising models to understand how music and, specifically, a sense of rhythm originated in the primate phylogeny.

## 1. Introduction

Rhythm permeates human life across physiological, behavioural and social domains: heartbeats, neural oscillations, spoken language and music are all built on precise rhythmic patterns, which act as building blocks of physiological and communicative processes [1]. This is true for humans, but how widespread is endogenous rhythm production in the animal kingdom [2]? The simplest, lowest-entropy rhythmic structure is isochrony, namely a pattern where time intervals between successive onsets of a signal all have roughly equal durations [2]. Isochronous rhythm can be produced at an individual level: the heartbeat, walking and singing are some examples. Rhythm also plays a crucial role in interactive processes, as in human music and animal coordinated displays.

Inter-individual interaction has an adaptive value in animal evolution [3] rhythm, as a tool, may help individuals to coordinate by predicting the interlocutor's turn and preparing and adapting their own. This can be mediated through two opposite strategies [3,4]. The first is synchrony, where the overlap between individuals' phonation is maximized; this happens, for example, in human joint singing. The second is turn-taking, where the overlap is minimized on the advantage of rapid exchange of short turns, as, for example, in speech.

The spontaneous production of isochronous vocal patterns seems relatively rare in other vertebrates; it has only been reported in two bird species [5,6], a bat [7], rock hyraxes [8] and one primate [9]. The empirical connection between isochrony and synchrony is even less clear because the two are often studied in isolation [10]. This lack of data strikingly contrasts with the hypothesis that a link between isochrony and synchrony, potentially already present in our hominoid ancestors' loud-calls, may be the mechanism that shaped our coordinated communication [4]. In fact, every human culture shows collective, pulse-based rhythmic singing and dancing displays, in which individuals entrain to an isochronous, predictable pulse, allowing group coordination [11]. This purported role of cooperative vocal interaction in enhancing rhythmic regularity has been detected not only in collective dance and music but also in vocalizations [12]: in humans, synchrony boosts isochrony.

Are humans the only apes capable of flexibly orchestrating isochrony and synchrony in their behaviour? To date, the joint presence of, and connection between, isochrony and synchrony in other mammals has never been found. Here we show the first evidence for isochrony, rhythmic flexibility mediated by co-singing, and synchrony in the vocalizations of an ape, the lar gibbon (*Hyllobates lar*; figure 1a). Gibbons are closely related to humans and share with us an unusual form of vocal communication: the song (figure 1b,c). Primate song plays a crucial role in various social and sexual dynamics (e.g. territorial defence, hierarchies and partnership assessment, courtship, social bonding, emotions sharing). Singing primates, in general, may thus represent a convenient animal model to unravel the origins and mechanisms shaping the evolution of speech and musicality [13,14]. Gibbons, and lar gibbons specifically, are ordinarily found in *ex-situ* contexts, making controlled collection of acoustic data possible. Similarities exist between gibbon songs and human music, for instance the collective context of emission of the song [15] and the association of ritualized locomotor displays to the song [16,17]. Because of this, gibbons may be good models for unravelling the biological origin of musicality in our *taxon* [16].

We searched for isochronous patterns in lar gibbons' songs, specifically in female and male contributions to duet (figure 1c) and male solos (figure 1b). We hypothesized that gibbons may flexibly deploy and modulate rhythmic isochrony based on social context [4,12,18]. We extracted the inter-onset-intervals ( $t_k$ ), namely the duration of the interval separating the onset of a song element from the next one. Then, we calculated the ratios between two adjacent intervals ( $r_k = t_k / (t_k + t_{k+1})$ ), analysing their distribution to detect the existence of categorical rhythms (density peaks on specific, small-integer ratios; e.g. a 1 : 1 peak represents isochrony) [5,9]. We then probed the biological substrates of isochrony in this species, testing the effect of biomechanical constraints, in terms of call rate, in explaining rhythmic patterns of gibbons' duet: calling fast is energetically expensive and a

potential signal of signaller's quality [19]. Because duetting seems a coordinated display [18], we also investigated the role of interaction in shaping the songs' rhythmic structure. To do so, we first assessed the causality of one individual's phonation on the one of the co-singer. Then, as synchrony may underpin coordinated displays [4,14], we quantified the overlapped phonation between duetting females and males, predicting that shared advertisement purposes enhance co-singer's synchrony. Finally, we tested for a relationship between isochrony and synchrony in the two sexes. We predicted that isochrony and synchrony may be strictly linked, similarly to humans. We also expected that the sexual dimorphism in form and function of gibbons' song may translate into a sex-specific trade-off between isochrony and synchrony.

## 2. Material and methods

### (a) Animals and recordings

Six habituated gibbon familiar groups were followed, with a total of 12 individuals and 215 songs, and specifically 157 female contributions to duet, 157 male contributions to duet and 58 male solos. Four groups were inhabiting the forests of Huai Kha Khaeng Wildlife Sanctuary (Thailand), and two other *ex situ* groups were living at the Cappeller faunistic park (Corvigliano, Italy) and the Falconara zoo park (Ancona, Italy). The contribution of each group and gibbon to the final dataset is provided in the electronic supplementary material, table S1.

We recorded the animals between 6:00 AM and 12:00 AM using a Sennheiser ME67 microphone connected to a solid-state digital audio recorder Tascam DR-100MKII (44.1 KHz sampling rate). All vocalizations were recorded at 5–50 m distance from the animals, aiming the microphone toward the individual vocalizing to maximize recording quality.

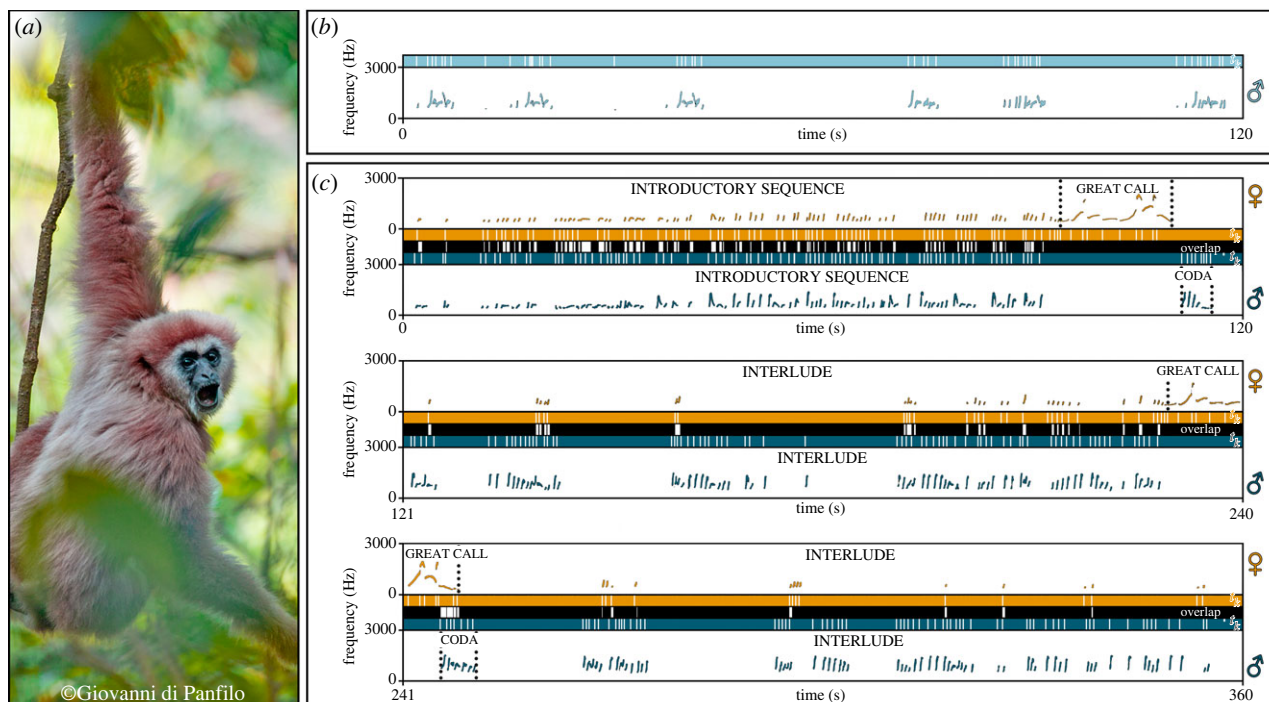
### (b) Acoustic analyses

We edited the songs using Praat 6.0.14 and saved them as WAV audio files [20]. Field notes and video recordings allowed us to recognize and separate individual contributions to each song; each annotated contribution was subsequently saved as a single Praat *TextGrid*, an object featuring onset and offset of each note. A computing cluster (OCCAM [21]) processed all 196 768 vocal units via a custom Praat script and exported all onsets of song units from separate *TextGrids* into one .csv datasheet. We calculated the temporal interval between an onset and the next one, which defines an inter-onset interval ( $t_k$ ). We focused on all  $t_k \leq 5$  sec, as this value is typically hypothesized as upper limit for meter perception and performance in humans and there is no quantitative evidence about an upper threshold on other apes [22]. We calculated the tempo frequency as the inverse of peak values of  $t_k$  per song type (Hz). The ratio ( $r_k$ ) was then calculated between a  $t_k$  and the next one,  $t_{k+1}$ , as  $t_k / (t_k + t_{k+1})$ .

### (c) Statistical analyses

#### (i) Testing isochrony per song type

To test the significance of the peaks of the  $r_k$  density distribution falling in the vicinity of isochrony (corresponding to  $r_k = 0.5$ ) we followed the methodology in Roeske and colleagues [5] and De Gregorio *et al.* [9]. On-isochrony ratio ranges were centred on a 1 : 1 ratio (i.e.  $r_k = 0.500$ ), while the off-isochrony ones correspond to the peripheral range at the left and right sides of the on-isochrony range. Specifically, we took on/off-isochrony boundaries at 0.400, 0.440, 0.555, 0.600 ( $r_k$  values) and counted, per individual contribution and per song type (female contribution to duet, male contribution to duet, male solo), the number of  $r_k$  instances falling



**Figure 1.** Organization of lar gibbon's songs. (a) Male lar gibbon singing in the Huai Kha Khaeng Wildlife Sanctuary (Thailand). (b) Spectrogram and inter-onset-interval graphs of the male solo. The fundamental frequency is highlighted in light blue on the spectrogram. The coloured bar indicates inter-onset intervals ( $t_k$ ) of the solo singing male, where solid white lines on the bar represent the onsets. (c) Spectrograms and inter-onset-interval graphs of the reproductive couple's whole duet. The fundamental frequency of individuals' contributions is highlighted on the spectrogram in dark blue for the male contribution to duet and dark yellow for the female contribution to duet. The sections of the song are labelled in the upper part of the spectrogram and separated with dotted lines. Coloured bars indicate inter-onset intervals ( $t_k$ ) of the contributions of each individual with white lines again corresponding to the onsets. Black bars turn white when the co-singers overlap. Notice how rhythmicity unfolds heterogeneously throughout the duet, alternating periods of higher and lower overlap. Note clusters onsets of the duetting gibbons influence each other (see also figure 4), with introductory sequences and interludes showing higher levels of synchrony, while great calls and codas partly overlap.

into the off-isochrony versus on-isochrony sectors of the curve. We calculated an *isochrony rate*, as the ratio between on- and off-isochrony observations per contribution.

Using a generalized linear mixed model (GLMM; R package *glmmTMB* [23]), we tested whether  $r_k$  observation counts differed by song type in interaction with the type of interval (on- or off-isochrony interval). The response variable was the observations count ( $r_k$ ), which followed a Poisson distribution, and the individual contribution code was used as a random factor (electronic supplementary material, table S3). To test the significance of the full model, we built a null model comprising only the random factors and compared the full and the null with a likelihood ratio test (Anova with 'Chisq' argument [24]). We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons for each level of the explanatory variables with the *emmeans* [25] package ( $p$ -values adjustment with Tukey method).

We then tested the effect of song types (fixed factor) on *isochrony rate* (response variable), which is the ratio between the number of on-isochrony and the number of off-isochrony observations per contribution, using a GLMM (*lme4* package [26]). The song code was entered as random factor. *isochrony rate* was log-transformed and followed a normal distribution. Full versus null models were compared with a likelihood ratio test (Anova with 'Chisq' argument [24]). We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons for each level of the explanatory variables with *emmeans* [25] package ( $p$ -values adjustment with Tukey method).

## (ii) Call rate calculation and assessment of its effect on isochrony

We calculated the *call rate* (number of onsets/10 s) and recalculated *isochrony rate* on contributions chunks of 10 s. A GLMM tested whether song type (predictor) affected *call rate* values

(response variable), with individual contribution code as random factor (electronic supplementary material, table S4). The response variable, once transformed as  $1/(call\ rate + 1)$ , fit a Gaussian distribution. We built (*lme4* [26] package) a full model and a null model comprising only the random factors. We compared them with a likelihood ratio test (Anova with 'Chisq' argument [24]). We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons for each level of the explanatory variables with *emmeans* [25] package ( $p$ -values adjustment with Tukey method).

We tested the effect of *call rate* in interaction with song type (fixed factors) on *isochrony rate* (response variable) with a GLMM (*lme4* [26] package), where the individual contribution code was a random factor (electronic supplementary material, table S5). The response variable, once log-transformed, followed a Gaussian distribution. We built and compared full and null models with a likelihood ratio test [24]. We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons and  $p$ -adjustment (Tukey method) with *emmeans* [25] and slope estimation with *lsmeans* [27] packages.

We wanted to test whether high versus low *call rate* affected isochrony. To test this hypothesis, we partitioned the *call rate* values between low and high *call rate*, depending on whether these were below or above the median of the specific song type. We counted how many  $r_k$  observations (response variable) for each song type (predictor) fell on or off peaks (predictor). We then built a GLMM model testing whether  $r_k$  counts could be explained not only by song type and on or off peaks—similarly to previous models—but also whether this purported relationship was modulated by an interaction with low versus high *call rate* (*glmmTMB* [23] package; see electronic supplementary material, table S6). We compared full and null models with a likelihood ratio test [24]. We obtained  $p$ -values for each predictor using the

R summary function and performed pairwise comparisons and  $p$ -adjustment (Tukey method) with *emmeans* [25].

### (iii) Assessing causality between male and female's contributions

We converted each individual contribution into binary series. We binned contributions into intervals of 10 ms. We mapped the interval to a new time series containing a 1 when phonated, 0 otherwise. We performed a Granger causality test (GrangerTest function in R – lag = 5, order = 500), which assesses whether a time series is efficient in predicting another, between the female and male binary time series of the same song, in both directions (F → M, M → F). We then counted the number of occurrences where  $p < 0.001$ , for F → M and M → F causality directions, to infer whether one of the two sexes affects more than the other the partner's timing of vocalizations. We limited the causality analysis to songs recorded in the wild to minimize the effect of potential external disturbances that often occur in captivity (e.g. other species' singing, visitors' presence).

### (iv) Is duets' overlap different from chance?

For every contribution recorded in the wild, we split songs' TextGrids into chunks of 3 min and calculated the overlapped phonation, summing the TextGrids' tiers with a custom Python [28] script. We then normalized it in two ways (electronic supplementary material, figure S7). First, we measured the rate (from 0, where no overlap is present, to 1, where all phonation is overlapped) of overlapped phonation on the total phonation of an individual contribution. Second, we also calculated the *normalized overlap*, by dividing the overlap duration in seconds by the total duration of phonation of the two co-singers (duration of phonation of the male + phonation of the female). We simulated random duets permuting each chunk against the others. For biological plausibility, the randomized permutation was carried out only between different sexes. Random duets were randomly paired songs either from real couples or from female-male non-coupled individuals. In simulations, we calculated the overlap and the *normalized overlap* as for real duets. We then tested whether the real overlap (REAL) was different from chance (SIMULATED permutations) with a GLMM (*lme4* [26] package; electronic supplementary material, table S8). The response variable, the *normalized overlap*, was normally distributed. We used the group code of the two singers as the random factor. We compared *full* and *null* model with a likelihood ratio test [24]. We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons and  $p$ -adjustment (Tukey method) with *emmeans* [25].

### (v) Correlation between overlap and isochrony

We investigated the effect of overlap on the *isochrony rate* with a GLMM (electronic supplementary material, table S9). Both *isochrony rate* and *normalized overlap* were calculated on the whole individual contribution. *Isochrony rate* was the response variable, while *normalized overlap* and sex, in interaction, were fixed factors. We tested for the interaction between overlap and sex, expecting potentially dimorphic effects of those variables on isochrony. We used the song code as a random factor. The *isochrony rate* followed a beta distribution, so we used the *glmmTMB* package to build the *full* and *null* models and compared them with a likelihood ratio test [24]. We obtained  $p$ -values for each predictor using the R *summary* function and performed pairwise comparisons and  $p$ -adjustment (Tukey method) with *emmeans* [25]. We performed slope estimation using the *lsmeans* [27] package.

For every model, we assessed normality, homogeneity (via function provided by R. Mundry), and number of the residuals; the *performance* [29] package tested for collinearity among fixed factors. Effect plots were produced with *sjPlot* [30] package.

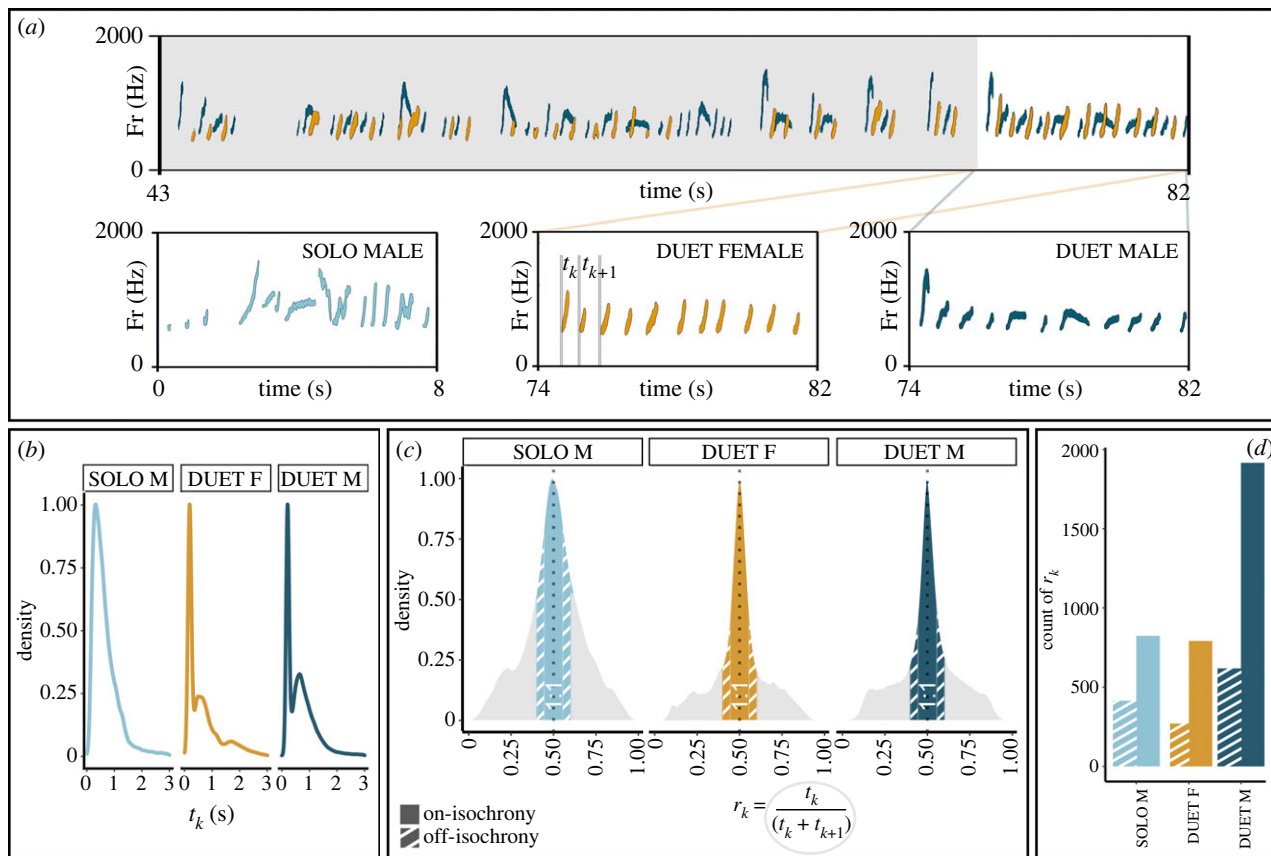
## 3. Results

### (a) Lar gibbons deploy isochrony in both duets and solos

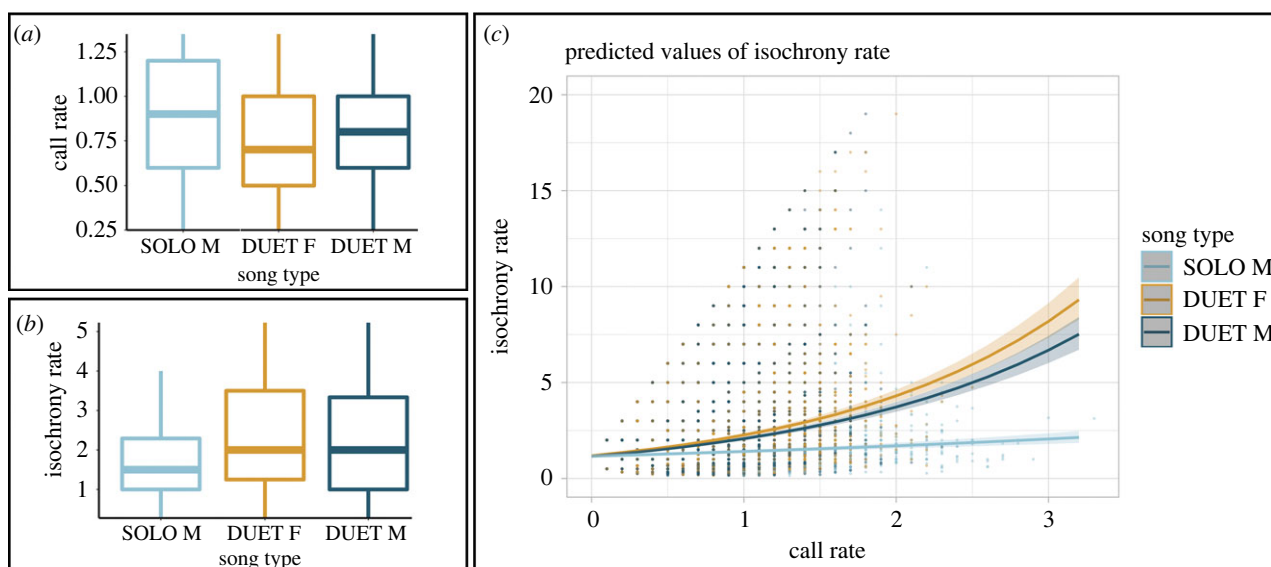
The distribution of raw inter-onset intervals ( $t_k$  – figure 2a; electronic supplementary material, figure S2) values shows two peaks for female and male duet contributions (figure 2b), with maxima corresponding to  $t_k$  values of 0.181 s (5.525 Hz) and 0.503 s (1.988 Hz) for the female duet, at 0.204 s (4.902 Hz) and 0.637 s (1.570 Hz) for the male duet. Male solos show only one peak at 0.322 s (3.105 Hz). Density plots of  $r_k$  values (figure 2c) show a clear peak on the small integer-ratio of 1 : 1, namely isochrony. At least visually, there seems to be strong rhythmic regularity. We quantified this intuition by computing and comparing the count of  $r_k$  values falling into on/off-isochrony boundaries [9] (figure 1d). Through our model (*full* versus *null*: df = 5, Chisq = 24989.590,  $p < 0.001$ ; electronic supplementary material, table S3), we found that the  $r_k$  peaks around isochrony were statistically significant in the three song types (*female contribution to duet*, off- versus on-isochrony  $r_k$  count: estimate =  $-1.168$ ,  $t$ -value =  $-95.220$ ,  $p < 0.001$ ; *male contribution to duet*, off- versus on-isochrony  $r_k$  count: estimate =  $-1.067$ ,  $t$ -value =  $-106.792$ ,  $p$ -value  $< 0.001$ ; *male solo* off- versus on-isochrony  $r_k$  count: estimate =  $-0.580$ ,  $t$ -value =  $-34.293$ ,  $p < 0.001$ ), meaning that the number of  $r_k$  values falling into on-isochrony range of the curve were significantly higher than those falling in the off-isochrony range.

### (b) Call rate, isochrony and their relationship are context-dependent

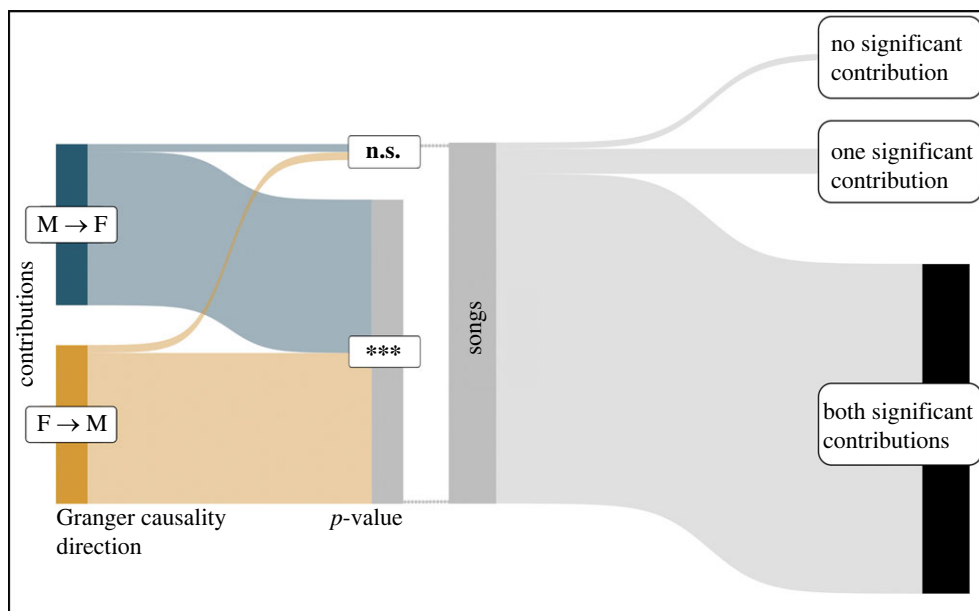
We obtained average *call rate* values ( $\pm$ s.d.) of  $0.768 \pm 0.360$  Hz for the *female contribution to duet*,  $0.774 \pm 0.315$  Hz for the *male contribution to duet* and  $0.943 \pm 0.468$  Hz for the *male solo*. Our model testing the effect of song type on *call rate* values (*full* versus *null*: Chisq = 65.469, d.f. = 2,  $p < 0.001$ ; electronic supplementary material, table S4) showed no rate difference between female and male contributions to the duet, but significantly higher values of *call rate* for the male solo (*post-hoc* comparisons; *female contribution to duet* versus *male contribution to duet*: estimate = 0.007,  $z$  ratio = 1.495,  $p = 0.293$ ; *female contribution to duet* versus *male solo*: estimate = 0.046,  $z$  ratio = 8.056,  $p < 0.001$ ; *male contribution to duet* versus *male solo*: estimate: 0.039,  $z$  ratio = 7.028,  $p < 0.001$  – electronic supplementary material, figure S2 and table S4; *call rate* values showed in figure 3a). Our model (*full* versus *null*; Chisq = 1569.052, d.f. = 5,  $p < 0.001$  – electronic supplementary material, table S5) testing the effect of *call rate* in interaction with the song type on the *isochrony rate* (figure 3b) showed no differences between female and male contributions in a duet (*post-hoc* comparisons; *female contribution to duet* versus *male contribution to duet*: estimate: 0.075,  $z$  ratio = 2.363,  $p$ -value = 0.048; *female contribution to duet* versus *male solo*: estimate = 0.383,  $z$  ratio = 9.442,  $p$ -value  $< 0.001$  – electronic supplementary material, table S5). Conversely, we found significantly lower isochrony in the *male solo* compared to both the female and male duets (*post-hoc* comparisons; *male contribution to duet* versus *male solo*: estimate = 0.309,  $z$  ratio = 7.742,  $p$ -value  $< 0.001$  – electronic supplementary material, figure S2 and table S5). Isochronous regularity differs between sexes in a duet—females are more isochronous—and either sex has



**Figure 2.** Rhythmic features of the lar gibbon's songs. (a) Zoomed-in portions of the spectrograms of both duet contributions and male solo. The fundamental frequency of individual contributions is highlighted in light-blue for the *male solo*, in dark-blue for the *male contribution to duet* and dark-yellow for the *female contribution to duet* (both detailed in specific portions indicated by the coloured lines). A representation of the onsets (grey lines) and the relative inter-onset intervals ( $t_k$ ) is reported on the spectrogram of the female contribution. (b) Probability density function representing the distributions of  $t_k$  per song type.  $t_k$  values are calculated on 12 adult individuals for a total of 372 individual contributions to songs. (c) Probability density function representing the distributions of rhythm ratios ( $r_k$ ) per song type, suggesting a difference between *male solos* and *male* and *female contributions to duets*. Solid coloured sections of the curves indicate on-isochrony  $r_k$  ranges, striped sections indicate off-isochrony  $r_k$  ranges. (d) Histogram of the counts for  $r_k$  values falling within on-isochrony versus off-isochrony ranges of the density function (depicted in panel c), per song type. For all song types, on-isochrony observations (solid bars) are significantly more numerous than off-isochrony ones (striped).



**Figure 3.** *call rate* and *isochrony* covariation in lar gibbon's song. (a) Boxplots representing the *call rate* per song type, calculated on chunks of 10 s for every individual contribution. The solo of the male shows significantly higher values of *call rate* than the male and female duets. Outliers corresponding to the 10% of higher and lower values of the variable are excluded from the plot. (b) Boxplots depicting the *isochrony rate* (on-isochrony counts/off-isochrony counts) per song type, calculated on chunks of 10 s for every individual contribution. The *isochrony rate* does not differ between male and female duet contributions, while the *isochrony rate* of the male singing solo is significantly lower than the one of the male and female in the duet. Outliers corresponding to the 10% of higher and lower values of the variable are excluded from the plot but included in the statistical computations. (c) Effect plot showing the predicted values derived from a GLMM looking at the effect of *call rate* on *isochrony rate*. The *female contribution to duet* and the *male contribution to duet* have a positive effect on *isochrony rate*, but both show a significantly steeper slope to the one of the *male solo*. Shaded areas indicate confidence intervals.



**Figure 4.** Granger causality in lar gibbon's duet—aggregated Sankey plots showing: on the left, the *proportion of contributions* having a significant causality ( $***p < 0.001$ ) on the onsets of the co-singer, for both directions (male on female and female on male) and non-significant ones (n.s.); on the right, the *proportion of songs* showing a reciprocal causality (both male on female and female on male), having significant causality in only one-direction (male on female or female on male) or no significant causality at all.

more isochrony in duets than males have in solos. Male solos have higher call rates but lower isochrony. These two results showed a context-dependent level of isochrony in the male, with the male singing more regularly when duetting than when singing alone; they also suggest that higher levels of isochrony *cannot* exclusively derive from higher call rates. Our model shows that call rate has a significant positive effect on isochrony; however, song type modulates this effect (i.e. the solo shows a significantly less steep slope, and thus significantly weaker effect, than the duet contributions of male and female—figure 3c, electronic supplementary material, table S4). To summarize, male solos show higher call rates, lower isochrony and weaker correlation between these two variables than the other song types.

We tested whether high versus low levels of *call rate* modulated isochrony. All the isochronous peaks were still significant at both high and low values of *call rate* (electronic supplementary material, table S5), confirming that *call rate* alone cannot explain isochrony levels.

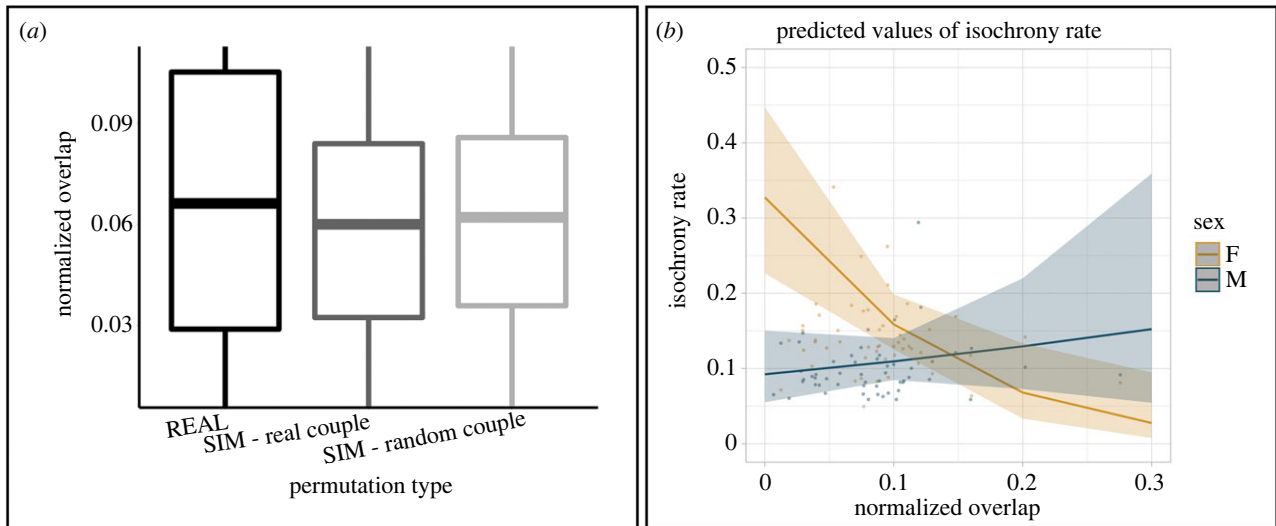
### (c) Male and female influence each other's phonation onsets

Since *call rate* was not sufficient to explain context-dependent isochrony levels, we investigated the presence of rhythmic interaction, by testing whether an individual's vocal rhythm affects its partner's. For this we used Granger Causality, a test probing whether future vocal onsets of one individual can be better predicted by considering past onsets of the co-singer (as opposed to past vocal onsets of the first individual). We found that 95% of the individual contributions to duets affect (i.e. Granger-caused at  $p < 0.001$ —figure 4) the partner's phonation. Moreover, this effect emerged bidirectionally, both for the female on male's contribution ( $F \rightarrow M$  in figure 4) and the male on female's contribution ( $M \rightarrow F$  in figure 4). At the song level, 91.39% of the songs showed a reciprocal highly

significant causality ( $p < 0.001$ ) between male and female contributions, 6.89% showed at least one significant contribution, 1.72% showed no significant contribution. This provided compelling evidence that the two sexes influence each other's onsets during the duet.

### (d) Synchrony: duet overlap is higher than chance

Since we found a potential for rhythmic interaction, we then quantified overlap in the duets through two different metrics. The first metric, normalized on the individual total phonation, showed that females presented a rate (from 0 to 1) of overlap (mean  $\pm$  s.d.) of  $0.177 \pm 0.145$  while males of  $0.158 \pm 0.131$ . The second metric, the *normalized overlap* (normalized on the sum of phonations of both co-singers – a rate from 0 to 0.5), in actual duets (REAL; mean  $\pm$  s.d. =  $0.079 \pm 0.037$ ) assessed the degree of synchrony between actual singing partners (electronic supplementary material, table S8). To have baselines and quantify randomness, the *normalized overlap* was also calculated on the randomly permuted opposite-sex individuals' contributions, for both existing couples (SIMULATED: real couple; mean  $\pm$  s.d. =  $0.060 \pm 0.037$ ; electronic supplementary material, table S8) and randomly paired individuals (SIMULATED: random couple; mean  $\pm$  s.d. =  $0.062 \pm 0.035$ ; electronic supplementary material, table S8). When testing the difference in overlap (response variable) between real duets and randomly simulated ones (figure 5a; GLMM, full versus null;  $\text{Chisq} = 28.711$ ,  $p < 0.001$ ; electronic supplementary material, table S9), we found that real duets showed significantly higher overlap than both types of simulated ones (SIMULATED random couple – REAL estimate =  $-0.018$ ;  $z$  value =  $-2.416$ ;  $p = 0.037$ ; SIMULATED real couple – REAL: estimate =  $-0.016$ ;  $z$  value =  $-5.347$ ;  $p < 0.001$ ). The two types of simulated duets showed no difference in their rate of *normalized overlap*. These results indicate that the overlap, and thus synchrony, heard in actual duets is not perfect but significantly higher than chance.



**Figure 5.** Synchrony and isochrony covariation in lar gibbon's song. (a) Boxplots showing the amount of *normalized overlap* (sum of all durations of overlapping female-male phonation / sum of all durations of female-male phonation), measuring the degree of synchrony. *Normalized overlap* is a rate going from 0 (no overlap between the co-singers) to 0.5 (co-singers perfectly overlapped). The amount of overlap is calculated for the real duets (REAL) and simulated ones (SIM - two types of simulations: simulated duets of real couples and simulated duets with random couples composed of opposite-sex individuals from the sample). Real duets show significantly higher overlap than simulated ones, displaying that the animals are statistically more synchronous than expected by chance. (b) Effect plot showing the predicted values taken from the GLMM looking at the effect of *normalized overlap* on *isochrony rate*. The two regression lines show the sex-specific trends that link the two variables. Shaded areas indicate confidence intervals.

### (e) The effect of synchrony on isochrony is sex-specific

Our model looking at the effect of synchrony (*normalized overlap*) and sex on isochrony (*isochrony rate*—full versus null;  $\text{Chisq} = 30.526$ ,  $\text{d.f.} = 3$ ,  $p < 0.001$ ; figure 5b; electronic supplementary material, table S10) showed a significant effect of synchrony on the *isochrony rate* (estimate =  $-9.484$ ,  $t$ -value =  $-2.416$ ,  $p = 0.001$ ; electronic supplementary material, table S10). In particular, we found a significant interaction between overlap and sex (estimate =  $11.383$ ,  $t$ -value =  $-5.347$ ,  $p < 0.001$ ; electronic supplementary material, table S10). There was a significant negative effect for females (confidence interval: lower cl. =  $-15.26$ ; upper cl. =  $-3.71$ ; electronic supplementary material, table S10)—which means higher isochrony corresponding to lower synchrony—and a non-significant positive trend for males (confidence interval: lower cl. =  $-3.37$ ; upper cl. =  $7.17$ ; electronic supplementary material, table S10 figure S5b): the effect of synchrony on isochrony is overall significant but sex-specific.

## 4. Discussion

We provide compelling evidence for isochrony and synchrony in the song of an ape and elucidate the existing link between the two. Our results show how lar gibbon's songs are organized into highly isochronous, though plastic, rhythmic patterns; this isochrony is only partially explained by physiological constraints linked to call rate. Male inter-onset intervals become less variable and more isochronous from solo to duet, suggesting that the interaction between co-singers may shape the song's rhythmic structure. Furthermore, the timing of phonation of each individual in a duet is predicted by the timing of the co-singer's vocalizations, suggesting a potential rhythmic interaction. The exact coincidence of events in time can be quantified as the overlap between two individuals' vocalizations, namely their synchrony, which we showed is higher than chance across singers when compared with artificial dyads. Finally, we searched and found a link between

synchrony and isochrony, specifically a significant statistical interaction indicating different trends across sexes. Females show a significant negative relationship between rhythm regularity and overlap, suggesting a potential trade-off between individual and group rhythms.

### (a) Isochrony and its role in coordinated displays

The production of isochronous rhythm in spontaneous animal vocalizations is rare and poorly studied [2,31]. Isochrony may have important biological functions: for instance, rock hyraxes producing more isochronous rhythmic patterns have higher fitness, in terms of reproductive success [8]. Here we provide evidence for above-chance occurrence of isochrony in an ape species. This occurs in both male and female songs.  $t_k$  peak values in lar gibbons are dimorphic and bimodal (similarly to another primate, *Indri indri* [9]), and differ between song types. Gibbons' isochrony transcends absolute temporal intervals: raw  $t_k$  values may vary but nonetheless lead to the same isochronous 1 : 1 ratio across sexes and song types. That is to say, as in human music, isochrony emerges even when tempo or note duration vary [11]. At the same time,  $t_k$  values are not randomly distributed, but mainly clustered around short intervals that produce the isochronous patterns: gibbon songs show two dominant tempi ( $1/t_k$  peak values) in the duet contributions which are tentatively close to two human behaviours: music and speech. The slower tempo—at 1.988 Hz for the female and 1.570 Hz for the male—is closer to Western music, where 2 Hz is the preferred tempo [32] and human locomotion [33]. The faster tempo we found—at 5.525 Hz for the female and 4.902 Hz for the male—falls in the range of speech, which relies on the association of vocal output and facial posture in the 3–7 Hz range [34]. Similarly, macaques' 'lip smacks' [35] and orangutans' 'clicks/faux-speech' [36] have tempi close to speech. Speech-like rhythm in nonhuman primate vocalizations, often associated with facial expressions, may underpin—according to some—an ancestral original audiovisual rhythmic feature, still detectable

in some branches of the primate phylogeny [35,36]. Gibbon tempi are numerically bounded by those of human music and speech. Beyond tempo, music and speech differ in rhythm regularity: music—like gibbon song—has it, while speech does not [37]. Human music and speech may have branched, says one hypothesis, from an ancestral ‘musical proto-language’ [38,39]; our finding of human-compatible song tempi in another ape provides indirect support to this hypothesis. Looking ahead, the sound-facial expressions link in gibbons is an open topic for future research, and so is the similarity between their songs’ tempo and human music or speech.

The use of isochrony to engage in group coordination, as for choruses and dance, was deemed a uniquely human trait [4]. Some features of gibbons’ song are considered inherited and developmentally fixed [16,40]. Nonetheless, lar gibbons can instantly and flexibly adjust their vocal contribution; they can start and stop singing depending on subtle temporal and spectral variations in the co-singer utterances [18]. This may involve a duet coordination mechanism that potentially requires mutual learning and fine-tuned adaptation. Supporting this hypothesis, established pairs show higher organization and coordination when compared to new ones [41]. Our data also support this learning and adaptation hypothesis. We found a reciprocal causal link between the timing of the vocal emission of duetting individuals: the onset of one individual determines the onset of the co-singer. Moreover, males show higher isochrony in duets than in solos. We thus show that isochrony is deployed differently depending on social context, potentially fulfilling a need for tuning rhythmic structure via tempo regularity. Speculatively, if reciprocal rhythmic adaptation were a flexible and expensive learning process, it might represent an explicit energy commitment; this rhythm-advertised commitment may enhance the strength of the pair-bond and lower the risk of partner desertion [42].

Human experiments show that both music and speech display temporal regularity when people mutually share a coordination purpose [12]: regular rhythm may support cooperative interaction. Our results showing enhanced isochrony in gibbons duetting might imply that they share with us some signalling aims, such as cooperatively communicating and then being subjected to similar selective pressures. If so, they may also rely on homolog or analogue neural mechanisms that allow rhythm and coordination in humans.

Finally, physiological constraints in vocal emission may also partly explain our results. Solo songs are, on average, longer [15] and show a higher call rate than duets. It is likely that a longer vocal display and a high call rate demand higher energy investment and depend on breathing constraints [19]: vocal fatigue may determine less regular rhythmic patterns in prolonged bouts [8,43]. At the same time, a costly display like the song is an honest signal of the emitter quality [44], making solo songs highly subjected to sexual selection [45]. The flexibility of rhythmic structure may thus not only serve vocal interaction between co-singers, but also be the outcome of a trade-off between quality signalling and physiological constraints.

### (b) Overlap versus turn-taking: parallels with human music and speech

Across species, synchronous coordination stems from two primary purposes: either minimization of the overlap in favour of

turn-taking or maximization of the overlap toward signal amplification [3,4]. We found a higher than chance rate of overlap in the songs. This finding dovetails with studies suggesting that overlap in the songs may mediate mechanisms such as mate attraction, anti-predatory purposes (e.g. gibbons [43], anurans [45]) and signalling the cohesion of an alliance in mammals (e.g. lemurs [46], dolphins [47]). In bottlenose dolphins, cooperative context enhances motor and vocal synchrony: a shared function may enhance synchrony and coordination [47]. Similarly, male and female gibbons jointly advertise their presence in the forest, and signal amplification through synchrony enhances transmission over long distances [14,48] which Merker *et al.* [4] hypothesized crucial in our ancestors’ long-distance calls. Long-range signal transmission is needed to localize conspecifics and reduce the costs of territorial conflicts [49,50]. The adaptiveness of group signalling more than individual broadcasting, and the preponderance of synchrony rather than turn-taking, may be the result of a monogamous mating system [51]. The avoidance of overlap seems to be preeminent in morphologically dimorphic species subjected to higher levels of sexual selection [52]. Our results showing a higher than chance level of overlap in a monogamous, non-morphologically dimorphic species corroborate the idea that species subjected to weaker sexual selection [53] benefit from synchrony more than from turn-taking.

Human speech and music are notably different in the amount of overlap versus turn-taking occurring during interactions [54]. Both speech and music rely on rhythm. In conversations, speakers tend to minimize silence and avoid overlap: turn-taking is enhanced through spectro-temporal clues like prosody. Cross-cultural work showed that overlap avoidance and silence minimization, quantified via response latency in conversation, are the norm across languages [55]. Conversely, ensemble music often shows high levels of overlap: in many musical cultures (near-)synchrony is enhanced. Both mechanisms involve flexibility, adjustment and anticipation that allow the two domains to serve their different, specific adaptive functions. Our gibbons exhibit an average overlap rate of 16–18%, with higher than chance level of synchrony. It is difficult to compare our overlap rates (obtained by dividing duration overlap by duration of phonation of the singer) to those of human conversations (often expressed as duration of silent gaps in seconds) or music. However, qualitatively, we would argue that gibbon overlap rates are well above what observed in human spoken conversations, and below those characteristics of several musical genres [55].

### (c) Sex-specific relationship between isochrony and synchrony

Our results highlighted a complex relationship linking synchrony and isochrony. We found an overall negative effect of synchrony on isochrony on the whole duet, and dimorphic effects of synchrony on isochrony. Synchrony did not enhance isochrony in males, but significantly decreased isochrony in females. This supports the idea that songs serve multiple functions that vary depending on dimorphic selection pressures acting on the duet contributions [13,56–58]. We propose that, for both sexes, isochrony may function to increase the redundancy of the signal during communication [2] and to mediate the overlap between the co-singers aiming at signal summation [4]. A trade-off mechanism may explain how females’ isochrony decreases with higher synchronization:



besides the interest in advertising the couple's mated status to neighbouring groups, females may broadcast their quality to higher quality males and female competitors. Previous work suggests that females' great calls are both an index of their physical condition [59] and the least overlapped part of the song [18]. Hence, we propose that female advertisement can be achieved by minimizing synchrony, hence changing the rhythmic structure of a bout, thus making it less predictable for their mate. In other words, the female would be the most rhythmically flexible [60] of the two in a race to the arms with its partner [61].

Some hypothesized human music originated via an ancestral form of loud call, shared with primates closely related to us [4,16]. Numerous extant ape species show vocal displays, e.g. loud calls and songs, that share features with human music and may also derive from a proto-musical loud call of our last common primate ancestor. Such features include loudness, for long-distance communication, tonal notes organized into higher-level structures, inked locomotor displays coordinated to vocal utterances and an inherited component [16]. After the divergence of the human branch, our species coordinated vocal displays specialized and acquired unique characteristics [16,62,63]: more flexibility in song structure improvisation and new conventions, a crucial role of learning shaping all those processes, the presence of a stable beat structuring rhythmic music structure. Our results offer indirect support to this hypothesis.

Isochrony is a 'statistical universal' feature of human music, meaning it is found more often than not across diverse musical cultures [11,64]. Some argue it may be rare in other species, or even a human prerogative [62]. Others maintain that isochrony's function for group coordination may be the one feature exclusive to our species [4]. Indeed, whether endogenous isochrony can be driven by exogenous (e.g. environmental, social) factors, in other species, is unknown [2], even in the only non-human mammals (*Indri indri*) shown to date to have rhythmic categories [9]. Our results not only support the idea that a primate closely related to us can spontaneously produce isochronous rhythms, but also that this rhythmic pattern can be shaped by exogenous *stimuli* linked to interaction. Endogenous and exogenous mechanisms coexist and are not mutually exclusive in determining rhythm [65]; consequently, rhythm should be investigated in terms of both underlying neural circuitry [66] and species-specific social and communicative adaptations [9]. From a comparative perspective, rhythmic capacities result from a mosaic of anatomic and functional changes throughout evolution. This perspective is in line with the *gradual audiomotor evolution* hypothesis, suggesting a coupling between the auditory and motor system in non-vocal-learners primate species [67]; the closer a primate is to humans, the more developed their rhythmic skills should be in terms of coordination and entrainment

skills. Rhythmic capacities in gibbons support for the *gradual audiomotor evolution hypothesis*: our similar neural substrate, and potentially convergent pressures for singing, may grant similar rhythmic capacities. The rhythmic features we found are only a first step towards understanding potential rhythmic entrainment in gibbons.

Our results all together support the idea that evolution, by convergence or shared ancestral traits, may have selected isochrony as an adaptive trait for collective, coordinated vocal displays. Still, the role of isochrony in mediating synchronization in animal species remains widely unexplored. Now that individual building blocks of animal rhythms are increasingly studied, one can start probing their interaction. We suggest that elucidating the link between synchrony and isochrony may represent a substantial step in reconstructing musicality evolution and its meaning for our and other species.

**Ethics.** We complied with all the good practice protocols for field primatology of the International Primatological Society (2014) to work with wild species of primates. NRCT (National Research Council of Thailand) provided us with research permits for collecting non-invasive acoustic recordings of wild gibbons, while according to Italian law, no permits were required on captive groups. We never used playback stimuli to avoid any modification of gibbons' behaviour. All procedures follow the guidelines of the Association for the Study of Animal Behaviour for the care and use of animals for research activities (2021).

**Data accessibility.** Timing data used for this study are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.wpzgmsbrm> [68]. Code and custom-written scripts are available from the corresponding author, T.R., upon request.

Additional tables and plots are provided in electronic supplementary material [69].

**Authors' contributions.** T.R.: conceptualization, formal analysis, investigation, visualization, writing—original draft, writing—review and editing; G.D.P.: investigation; M.P.: investigation; M.Z.: investigation; L.F.: investigation; T.S.: investigation; M.G.: conceptualization, investigation, supervision, writing—review and editing; A.R.: conceptualization, investigation, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare we have no competing interests.

**Funding.** Center for Music in the Brain is funded by the Danish National Research Foundation (DNRF117). The Comparative Bioacoustics Group is funded by Max Planck Group Leader funding to A.R.

**Acknowledgements.** Our thanks go to NRCT (National Research Council of Thailand) for the research permits (No. 0002/1356 - 11th February 2019) and DNP (Department of National Parks, Wildlife and Plants Conservation) people of Khao Nang Rum Research Station and to the CEG (Conservation Ecology Group) of the King's Mongkut University of Bangkok. We would also like to thank Parco Faunistico Cappeller (Cartigliano, Vicenza, Italy) and Parco Zoo Falconara (Falconara Marittima, Ancona, Italy) for giving us the opportunity to study their animals. We are grateful to Olivier Friard for his help during the scripting and performing of the quantitative statistical analysis using OCCAM.

## References

- Patel AD. 2003 Rhythm in language and music. *Ann. N Y Acad. Sci.* **999**, 140–143. (doi:10.1196/annals.1284.015)
- Ravignani A, Madison G. 2017 The paradox of isochrony in the evolution of human rhythm. *Front. Psychol.* **8**, 1–13. (doi:10.3389/fpsyg.2017.01820)
- de Reus K, Soma M, Anichini M, Gamba M, De Heer Kloots M, Lense M, Bruno JH, Trainor L, Ravignani A. 2021 Rhythm in dyadic interactions. *Phil. Trans. R. Soc. B* **376**, 20200337. (doi:10.1098/rstb.2020.0337)
- Merker BH, Madison GS, Eckerdal Gamla Kyrkv P. 2009 On the role and origin of isochrony in human rhythmic entrainment. *Cortex* **45**, 4–17.
- Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical rhythms are shared between songbirds and humans. *Curr. Biol.* **30**, 3544–3555.e6. (doi:10.1016/j.cub.2020.06.072)
- Norton P, Scharff C. 2016 'Bird Song Metronomics': isochronous organization of zebra finch song

- rhythm. *Front. Neurosci.* **10**, 309. (doi:10.3389/fnins.2016.00309)
7. Burchardt LS, Norton P, Behr O, Scharff C, Knörnschild M. 2019 General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *R. Soc. Open Sci.* **6**, 181076. (doi:10.1098/rsos.181076)
  8. Demartsev V, Haddas-Sasson M, Ilany A, Koren L, Geffen E. 2022 Male rock hyraxes that maintain an isochronous song rhythm achieve higher reproductive success. *J. Anim. Ecol.* **00**, 1–12. (doi:10.1111/1365-2656.13801)
  9. De Gregorio C, Valente D, Raimondi T, Torti V, Miarretsoa L, Friard O, Giacoma C, Ravignani A, Gamba M. 2021 Categorical rhythms in a singing primate. *Curr. Biol.* **31**, R1379–R1380. (doi:10.1016/j.cub.2021.09.032)
  10. Greenfield MD, Honing H, Kotz SA, Ravignani A. 2021 Synchrony and rhythm interaction: from the brain to behavioural ecology. *Phil. Trans. R. Soc. B* **376**, 20200324. (doi:10.1098/rstb.2020.0324)
  11. Savage PE, Brown S, Sakai E, Currie TE. 2015 Statistical universals reveal the structures and functions of human music. *Proc. Natl Acad. Sci. USA* **112**, 8987–8992. (doi:10.1073/pnas.1414495112)
  12. Bowling DL, Herbst CT, Fitch WT. 2013 Social origins of rhythm? Synchrony and temporal regularity in human vocalization. *PLoS ONE* **8**, e80402. (doi:10.1371/journal.pone.0080402)
  13. Clarke E, Reichard UH. 2006 The syntax and meaning of wild gibbon songs. *PLoS ONE* **10**, e73. (doi:10.1371/journal.pone.0000073)
  14. Gamba M, Torti V, Estienne V, Randrianarison RM, Valente D, Rovara P, Bonadonna G, Friard O, Giacoma C. 2016 The indris have got rhythm! Timing and pitch variation of a primate song examined between sexes and age classes. *Front. Neurosci.* **10**, 249. (doi:10.3389/fnins.2016.00249)
  15. Raemaekers JJ, Raemaekers PM, Haimoff EH. 1984 Loud calls of the Gibbon (*Hylobates lar*): repertoire, organisation and context. *Behaviour* **91**, 146–189. (doi:10.1163/156853984X00263)
  16. Geissmann T. 1999 Gibbon Songs and Human Music from an Evolutionary Perspective. In *The origins of music* (eds S Brown, B Merker, C Wallin). Cambridge, MA: The MIT Press.
  17. Fan P-F, Ma C-Y, Garber PA, Zhang W, Fei H-L, Xiao W. 2016 Rhythmic displays of female gibbons offer insight into the origin of dance. *Sci. Rep.* **6**, 34606. (doi:10.1038/srep34606)
  18. Terleph TA, Malaivijitnond S, Reichard UH. 2018 Male white-handed gibbons flexibly time duet contributions. *Behav. Ecol. Sociobiol.* **72**, 1–13. (doi:10.1007/s00265-017-2432-z)
  19. Catchpole CK, Slater PJB. 2003 *Bird song: biological themes and variations*. Cambridge, UK: Cambridge University Press.
  20. Boersma P. 2001 Praat, a system for doing phonetics by computer. *Glottol.* **5**, 341–345.
  21. Aldinucci M, Bagnasco S, Lusso S, Pasteris P, Rabellino S, Vallero S. 2017 OCCAM: a flexible, multi-purpose and extendable HPC cluster. *J. Phys.: Conf. Ser.* **898**, 082039. (doi:10.1088/1742-6596/898/8/082039)
  22. London J. 2004 *Hearing in time: psychological aspects of musical meter*. Oxford, UK: Oxford University Press.
  23. Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Machler M, Bolker BM. 2017 *glmmTMB* balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R. J.* **9**, 378–400. (doi:10.32614/RJ-2017-066)
  24. Dobson AJ, Barnett AG. 2018 *An introduction to generalized linear models*. London, UK: Chapman and Hall/CRC.
  25. Lenth R. 2020 Emmeans: estimated marginal means, aka least-squares means. R package version 1.5.0. See <https://rdrr.io/cran/emmeans/>.
  26. Bates D, Mächler M, Bolker B, Walker S. 2014 Fitting linear mixed-effects models using lme4. *arXiv:1406.5823* [stat].
  27. Lenth RV. 2016 Least-squares means: the R package lsmeans. *J. Stat. Soft.* **69**, 1–33. (doi:10.18637/jss.v069.i01)
  28. Van Rossum G, Drake FL. 2009 *Python/C API manual—python 3*. n.p.: CreateSpace.
  29. Lüdecke D, Makowski D, Waggoner P, Patil I. 2020 Performance: assessment of regression models performance. R package version 0.4. See <https://dominiquemakowski.github.io/publication/ludecke2021performance/ludecke2021performance.pdf>.
  30. Lüdecke D. 2022 Sjplot: data visualization for statistics in social science. R package version 2.8.11. See <https://CRAN.Rproject.org/package=sjplot>.
  31. Bouwer FL, Nityananda V, Rouse AA, ten Cate C. 2021 Rhythmic abilities in humans and non-human animals: a review and recommendations from a methodological perspective. *Phil. Trans. R. Soc. B* **376**, 20200335. (doi:10.1098/rstb.2020.0335)
  32. Stevens C. 2002 *ICMPC7: 7th international conference on music perception & cognition: conference proceedings*. Adelaide, Australia: Causal Productions for AMPS.
  33. MacDougall HG, Moore ST. 2005 Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *J. Appl. Physiol.* **99**, 1164–1173. (doi:10.1152/jappphysiol.00138.2005)
  34. Chandrasekaran C, Trubanova A. 2009 The natural statistics of audiovisual speech. *PLoS Comput. Biol.* **5**, e1000436. (doi:10.1371/journal.pcbi.1000436)
  35. Ghazanfar AA, Takahashi DY. 2014 Facial expressions and the evolution of the speech rhythm. *J. Cogn. Neurosci.* **26**, 1196–1207. (doi:10.1162/jocn\_a\_00575)
  36. 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)
  37. Brown S, Weishaar K. Speech is 'heterometric': the changing rhythms of speech. See [https://iscaspeech.org/archive\\_v0/sp2010/sp10\\_074.html](https://iscaspeech.org/archive_v0/sp2010/sp10_074.html).
  38. Ravignani A, De Boer B. 2021 Joint origins of speech and music: testing evolutionary hypotheses on modern humans. *Semiotica* **2021**, 169–176. (doi:10.1515/sem-2019-0048)
  39. Brown S. 2017 A joint prosodic origin of language and music. *Front. Psychol.* **8**, 1894. (doi:10.3389/fpsyg.2017.01894)
  40. Geissmann T. 1984 Inheritance of song parameters in the Gibbon Song, analysed in 2 Hybrid Gibbons (*Hylobates pileatus* X *H. lar*). *Folia Primatol.* **42**, 216–235. (doi:10.1159/000156165)
  41. Haimoff EH. 1984 Acoustic and organizational features of gibbon songs. In *The lesser apes: evolutionary and behavioral biology* (eds H Preuschoft, DJ Chivers, WY Brockelmann, N Creel), pp. 333–353. Edinburgh, UK: Edinburgh University Press.
  42. Wickler W. 1980 Vocal duetting and the pair bond: I. coyness and partner commitment. A hypothesis. *Zeitschrift für Tierpsychologie* **52**, 201–209. (doi:10.1111/j.1439-0310.1980.tb00711.x)
  43. Brockelman WY, Reichard U, Treesucon U, Raemaekers JJ. 1998 Dispersal, pair formation and social structure in gibbons (*Hylobates lar*). *Behav. Ecol. Sociobiol.* **42**, 329–339. (doi:10.1007/s002650050445)
  44. Barelli C, Mundry R, Heistermann M, Hammerschmidt K. 2013 Cues to androgens and quality in male Gibbon songs. *PLoS ONE* **8**, e82748. (doi:10.1371/journal.pone.0082748)
  45. Legett HD, Aihara I, Bernal XE. 2021 The dual benefits of synchronized mating signals in a Japanese treefrog: attracting mates and manipulating predators. *Phil. Trans. R. Soc. B* **376**, 20200340. (doi:10.1098/rstb.2020.0340)
  46. Torti V, Gamba M, Rabemananjara ZH, Giacoma C. 2013 The songs of the indris (Mammalia: Primates: Indridae): contextual variation in the long-distance calls of a lemur. *Ital. J. Zool.* **80**, 596–607. (doi:10.1080/11250003.2013.845261)
  47. Moore BL, Connor RC, Allen SJ, Krützen M, King SL. 2020 Acoustic coordination by allied male dolphins in a cooperative context. *Proc. R. Soc. B* **287**, 20192944. (doi:10.1098/rspb.2019.2944)
  48. Rehberg-Besler N, Doucet SM, Mennill DJ. 2016 Overlapping vocalizations produce far-reaching choruses: a test of the signal enhancement hypothesis. *Behav. Ecol.* **28**, 494–499. (doi:10.1093/beheco/arw176)
  49. Bonadonna G, Torti V, Randrianarison RM, Martinet N, Gamba M, Giacoma C. 2014 Behavioral correlates of extra-pair copulation in *Indri indri*. *Primates* **55**, 119–123. (doi:10.1007/s10329-013-0376-0)
  50. Reichard U. 1995 Extra-pair copulations in a monogamous Gibbon (*Hylobates lar*). *Ethology* **100**, 99–112. (doi:10.1111/j.1439-0310.1995.tb00319.x)
  51. Haimoff EH. 1986 Convergence in the duetting of monogamous Old World primates. *J. Hum. Evol.* **15**, 51–59. (doi:10.1016/S0047-2484(86)80065-3)
  52. Deputte BL. 1982 Duetting in male and female songs of the white-cheeked gibbon (*Hylobates concolor leucogenys*). In *Primate communication* (eds CT Snowdon, CH Brown, MR Peterson). Cambridge, UK: Cambridge University Press.

53. Leigh SR, Shea BT. 1995 Ontogeny and the evolution of adult body size dimorphism in apes. *Am. J. Primatol.* **36**, 37–60. (doi:10.1002/ajp.1350360104)
54. de Reus K. et al. 2020 *Rhythm in dyadic interactions*. See <https://osf.io/9yrkv> (doi:10.31234/osf.io/9yrkv)
55. Stivers T et al. 2009 Universals and cultural variation in turn-taking in conversation. *Proc. Natl Acad. Sci. USA* **106**, 10 587–10 592. (doi:10.1073/pnas.0903616106)
56. Cowlishaw G. 1992 Song function in Gibbons. *Behavior* **121**, 131–153. (doi:10.1163/156853992X00471)
57. Mitani JC. 1988 Male Gibbon (*Hylobates agilis*) singing behavior: natural history, song variations and function. *Ethology* **79**, 177–194. (doi:10.1111/j.1439-0310.1988.tb00710.x)
58. Geissmann T. 2002 Duet-splitting and the evolution of gibbon songs. *Biol. Rev. Camb. Philos. Soc.* **77**, 57–76. (doi:10.1017/S1464793101005826)
59. Terleph TA, Malaivijitnond S, Reichard UH. 2016 Age related decline in female lar gibbon great call performance suggests that call features correlate with physical condition. *BMC Evol. Biol.* **16**, 1–13. (doi:10.1186/s12862-015-0578-8)
60. De Gregorio C, Zanolli A, Valente D, Torti V, Bonadonna G, Randrianarison RM, Giacoma C, Ravignani A. 2019 Female indris determine the rhythmic structure of the song and sustain a higher cost when the chorus size increases. *Cur. Zool.* **65**, 89–97. (doi:10.1093/cz/zoy058)
61. Greenfield MD, Rolzen I. 1993 Katydid synchronous chorusing is an evolutionarily stable outcome of female choice. *Nature* **364**, 618–620. (doi:10.1038/364618a0)
62. Fitch WT. 2005 The evolution of music in comparative perspective. *Ann. N Y Acad. Sci.* **1060**, 29–49. (doi:10.1196/annals.1360.004)
63. Hauser MD, McDermott J. 2003 The evolution of the music faculty: a comparative perspective. *Nat. Neurosci.* **6**, 663–668. (doi:10.1038/nn1080)
64. Jacoby N, McDermott JH. 2017 integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* **27**, 359–370. (doi:10.1016/j.cub.2016.12.031)
65. Merker B. 1999 Synchronous chorusing and the origins of music. *Musicae Sci.* **3**, 59–73. (doi:10.1177/102986490000305105)
66. Honing H, Bouwer FL, Prado L, Merchant H. 2018 Rhesus Monkeys (*Macaca mulatta*) sense isochrony in rhythm, but not the beat: additional support for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* **12**, 475. (doi:10.3389/fnins.2018.00475)
67. Merchant H, Honing H. 2014 Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* **7**, 274. (doi:10.3389/fnins.2013.00274)
68. Raimondi T, Di Panfilo G, Pasquali M, Zarantonello M, Favaro L, Savini T, Gamba M, Ravignani A. 2023 Data from: Isochrony and rhythmic interaction in ape duetting. Dryad Digital Repository. (doi:10.5061/dryad.wpzgmsbrm)
69. Raimondi T, Di Panfilo G, Pasquali M, Zarantonello M, Favaro L, Savini T, Gamba M, Ravignani A. 2023 Isochrony and rhythmic interaction in ape duetting. Figshare. (doi:10.6084/m9.figshare.c.6350493)