Measuring rhythms of vocal interactions: a proof of principle in harbour seal pups

PHILOSOPHICAL TRANSACTIONS
OF THE ROYAL SOCIETY B
BIOLOGICAL SCIENCES

Measuring rhythms of vocal interactions: A proof of principle in harbour seal pups

Journal:	*Philosophical Transactions B*
Manuscript ID	RSTB-2021-0477.R1
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Anichini, Marianna; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group; Norwegian University of Science and Technology, Biological Sciences; Hanse-Wissenschaftskolleg; Carl von Ossietzky University of Oldenburg, Department for Neuroscience de Reus, Koen; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group; Vrije Universiteit Brussel Brussels Education Center, Artificial Intelligence Lab; Radboud University Donders Institute for Brain Cognition and Behaviour Hersh, Taylor; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group Valente, Daria; University of Turin, Life Sciences and Systems Biology Salazar-Casals, Anna; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group; Sealcentre Pieterburen, Research Berry, Caroline; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group Keller, Peter; The MARCS Institute for Brain Behaviour and Development; Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Ravignani, Andrea; Max Planck Institute for Psycholinguistics, Comparative Bioacoustics Group; Center for Music in the Brain, Clinical Medicine, Aarhus University & The Royal Academy of Music
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	FACE2FACE
Subject:	Behaviour < BIOLOGY, Cognition < BIOLOGY, Evolution < BIOLOGY
Keywords:	Behavioural interaction, asynchrony, circular statistics, categorical rhythms, time series analysis, interactive vocal rhythm

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
Yes

Statement (if applicable):
This research was conducted following the principles of animal care and the Dutch and EU regulations for the ethical treatment of animals in behavioural research and was subject to the health and needs of the animals. All recordings complied with the guidelines outlined by the Veterinary Department of the Sealcentre Pieterburen.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
The datasets and codes for statistical analysis used in this article have been uploaded as part of the Supplementary Material. Please follow this link to find the related files:
https://osf.io/8m4yv/?view_only=2549b6002bbc4dbd9a86735a7e8393d0

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Measuring rhythms of vocal interactions: A proof of principle in harbour seal pups

Marianna Anichini[1,2,3,4], Koen de Reus[1,5,6], Taylor A. Hersh[1], Daria Valente[7], Anna Salazar-Casals[1,8], Caroline Berry[1], Peter E. Keller[9,10#] and Andrea Ravignani[1,10#]

[1] Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
[2] Department of Biological Sciences, Faculty of Natural Sciences, Norwegian University of Science and Technology Ålesund, Norway.
[3] Hanse-Wissenschaftskolleg Institute for Advanced Study, 'Brain' Research Area, Delmenhorst, Germany.
[4] Division Animal Physiology & Behaviour, Department for Neuroscience, Carl von Ossietzky University Oldenburg, Oldenburg, Germany.
[5] Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium.
[6] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
[7] Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.
[8] Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands.
[9] MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Australia.
[10] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark.
[#]Joint senior authors

Correspondence:
Marianna Anichini
e-mail: marianna.anichini@mpi.nl
Andrea Ravignani
e-mail: andrea.ravignani@mpi.nl

ORCID ID: MA, 0000-0002-2697-0297; KdR, 0000-0002-6129-9726; TAH, 0000-0002-7891-596X; DV, 0000-0001-6086-5135; ASC, 0000-0002-3722-5436; CB, 0000-0003-4918-5846; PK, 0000-0001-7579-6515; AR, 0000-0002-1058-0024

Abstract

Rhythmic patterns in interactive contexts characterise human behaviours such as conversational turn-taking. These timed patterns are also present in other animals, and often described as rhythm. Understanding fine-grained temporal adjustments in interaction requires complementary quantitative methodologies. Here, we showcase how vocal interactive rhythmicity in a non-human animal can be quantified using a multi-method approach. We record vocal interactions in harbour seal pups (*Phoca vitulina*) under controlled conditions. We analyse these data by combining analytical approaches, namely categorical rhythm analysis, circular statistics, and time series analyses. We test whether pups' vocal rhythmicity varies across behavioural contexts depending on the absence or presence of a calling partner. Four research questions illustrate which analytical approaches are complementary vs. orthogonal. For our data, circular statistics and categorical rhythms suggest that a calling partner affects a pup's call timing. Granger causality suggests that pups predictively adjust their call timing when interacting with a real partner. Lastly, the ADaptation and Anticipation Model (ADAM) estimates statistical parameters for a potential mechanism of temporal adaptation and anticipation. Our analytical complementary approach constitutes a proof of concept; it shows feasibility in applying typically unrelated techniques to seals to quantify vocal rhythmic interactivity across behavioural contexts.

Keywords: Behavioural interaction, asynchrony, circular statistics, categorical rhythms, time series analysis, interactive vocal rhythm

1. Introduction & methodological approaches

1.1. Rhythmic interaction and turn-taking

Humans adopt precise signalling behaviours to exchange information [1,2]. No matter the signal modality (e.g., acoustic, visual), an interactive event between sender and receiver is governed by a timed structure [3–6]. The structured exchange of communicative turns (i.e., turn taking) characterises our capacity for social interaction, enabling us to communicate with others [7,8]. The study of interactive rhythms—how two (or more) individuals coordinate their signalling in time—is an emerging field of research, with more recent work extending structural analyses of communication signals, once restricted to human spoken conversation, to other species [3,9]. Turn-taking in communication has been documented in primates [10] and in other non-primate species [6,11–13]. For example, call exchanges in adult common marmoset monkeys (*Callithrix jacchus*) exhibit coupled oscillator dynamics, like those observed in human turn-taking [14]. Two key components of turn-taking are the flexible organisation and distribution of turns, and the temporal relationship between adjacent turns [11]. In humans, face-to-face interactions require participants to be, among others, both socially and temporally sensitive [15]. Does behavioural context also affect signal timing in non-human animals? Comparative investigations on interactive vocal timing in mammals may help uncover shared turn-taking mechanisms, potentially providing more insights into their evolution.

1. 2. Methodological approaches and challenges for studying rhythmic interaction

Expanding the human turn-taking framework to other species is currently hindered by, among other things, lack of suitable methodological approaches [11,16]. Cross-species frameworks exist [17], but open questions still remain in animal face-to-face interaction, such as: Which analytical methodologies used to investigate turn-taking in humans may reveal temporal adjustments in other species? Can turn-taking arise from non-cooperative behavioural interactions? Can methods developed for individual rhythm analyses be used to study rhythmic interaction? Can parametric models for human rhythmic prediction and reaction detect similar features in other species? These open questions require suitable animal models and quantitative methods.

Different forms of vocal rhythmic interaction, such as synchronous chorusing and turn-taking, have been mainly studied in mammals within a cooperative dynamic, like parent-infant and male-female dyads. Time series analyses like Granger causality have shown temporal interdependence between vocalisations in male-female pairs [18,19] and movements [20] of non-human primates. Circular statistics is another method to study timing adjustments in interactions, and has been used in previous animal work, including a seal pup playback experiment [21] and a study on parent-infant monkey interaction [22]. Categorical rhythms—those for which the temporal intervals between signal onsets are distributed categorically rather than uniformly—are a universal characteristic of human

music, which is often produced in an interactive context [23]. Similar rhythms are also present in non-human animal songs (e.g., thrush nightingales [23], indris [24], but whether such rhythms characterise other interactive non-song vocalisations, such as animal calls, is unknown [23,24]. Roeske and colleagues [23] hypothesised that categorical rhythms play a role in calls produced to attract and hold conspecific attention, by making sequences of vocalisations more predictable to listeners. Categorical rhythm analyses could therefore be an interesting method to test the predictability of vocal sequences in non-human animal interactions. Lastly, the ADaptation and Anticipation Model (ADAM), originally developed to model the mechanisms for interpersonal coordination in humans [25], has been adopted to probe sensorimotor and cognitive mechanisms underlying temporal dynamics in interaction [26,27]. Although ADAM is designed for 'simultaneous chorusing', it could also be used for a mixture of bouts of synchrony, turn-taking, and other regimes [21,28].

In this proof of concept study, we showcase how these methodological tools— Granger causality, circular statistics, categorical rhythm analysis, and ADAM—can be applied to a new animal model: the harbour seal (*Phoca vitulina*). To illustrate the utility and compatibility of these different analytical methods, we show how they can be used to better understand the rhythmic communication of a small sample of harbour seal pups in different behavioural contexts.

1.3. Our animal model

The "vocal learning-beat perception and synchronisation" (VL-BPS) hypothesis states that only vocal learning species—those capable of producing new vocalisations or modifying existing ones based on auditory experience—may possess advanced rhythmic abilities [29,30]. This hypothesis is inherently cross-modal: it suggests a strong link between audition and timed movement. For example, Snowball, a sulphur-crested cockatoo (*Cacatua galerita eleonora*), was shown to perceive auditory rhythms at different tempi and to predictively synchronise his body movements to them [31]. Parrots are phylogenetically distant from humans and, among mammals, pinnipeds (seals, sea lions, and walruses) are one of the vocal learning groups (besides humans, bats, elephants, and cetaceans). Pinnipeds may well be the best mammalian model for testing the VL-BPS hypothesis—the ability to extract a beat from periodic acoustic stimuli and entrain to it in a predictive and adaptive manner—since some species showed vocal mimicry and plasticity [32,33] and others can keep a beat [34]. These characteristics, paralleling human abilities, make pinnipeds an ideal animal clade for comparative research on the origins of rhythmic communicative behaviour.

Harbour seals exhibit both vocal flexibility [33,35] and rhythmic interactivity [21], and are particularly vocal in the first few weeks of life [36]. During the lactation period, harbour seal pups emit 'mother attraction calls' (hereafter 'calls') to draw their mothers' attention [37]. Mothers are silent and use the individual vocal signatures in these calls to recognize their pups [36,38]. Against

the acoustically complex backdrop of large mother-pup rookeries, rhythmically tuned pup calls could constitute a socio-ecologically selected trait that allows individual pups to avoid conspecific call overlap by adjusting the timing of their own call onsets. Such timing plasticity could allow a pup to be more acoustically conspicuous and increase its chances of successful reunions with its mother. Unlike cooperative types of turn-taking (e.g., in humans and in common marmosets (*Callithrix jacchus*) [39]) harbour seal pups' interactions are a by-product of neighbouring pups vocalising to attract their silent mothers, and are thus probably competitive.

To date, only two papers studied vocal rhythms in harbour seals, crucially both focusing on single individuals [21,28]. The first study was a playback experiment in which a pup vocally interacted with sounds broadcasted from a loudspeaker [21]. The pup adjusted the timing of its calls in an asynchronous manner by responding to the broadcasted conspecific calls with a non-uniformly distributed response phase whose mean approximated 90º [21]. The second study looked at the presence and development of vocal rhythms in three harbour seal pups [28]. Complementary analytical approaches showed how the pups' *individual* calling patterns gained more rhythmic structure over time [28]. However, a major limitation of both studies was the lack of sociality (i.e., individuals were tested alone) and, by extension, interactivity (i.e., the stimuli did not adapt to the response of the tested animals).

1.4. Aims & research questions

In this work, we show how vocal interactive rhythmicity in non-human animals can be quantified using a multi-method approach spanning various research domains (e.g., temporal, social, cognitive) (Table 1). We illustrate this approach through four research questions, all of which relate back to whether harbour seal vocal interactive rhythmicity varies in different behavioural contexts (Table 1). While our sample sizes are too small to enable species-wide inferences, they are sufficient to illustrate how methods typically used to study human communication can be adopted to study interactivity in animal communication. The goal of this paper is thus to outline a quantitative roadmap that future research can follow. Circular statistics and categorical rhythm analysis are used to address the first question about temporal adjustment in interaction: *"does the presence of a calling partner affect the call timing of individual pups?"* (Q1). The next two questions consider the effect of behavioural context on temporal adjustment: "*does the type of calling partner (real or broadcasted) affect the call timing of individual pups?*" (Q2) and "*when the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing?*" (Q3). We answer these questions using circular statistics (Q2 and Q3) and Granger causality tests (Q2). Lastly, ADAM is used to investigate the fourth question about the cognitive processes involved in temporal adjustment: *"which timing mechanisms are used by pups during vocal interactions?"* (Q4).

1.5. Subjects, housing conditions, and behavioural contexts

We recorded nine wild-born pups (A-I) calling in different behavioural contexts while housed at Sealcentre Pieterburen (the Netherlands) (Method S1/S2). During the recordings, each pup was housed in an enclosure with a swimming pool and a resting platform (Figure S1). One pup was housed alone (I) while the others were housed in pairs (A/B, C/D, E/F, G/H). Note that the enclosures were physically but not acoustically isolated from each other, meaning that pups could hear other pups in neighbouring enclosures.

We analysed focal pup vocalisations during four different behavioural contexts (Figures 1, Table S1): (i) when the focal pup was alone (pup I), (ii) when the focal pup heard a playback of conspecific calls (pup I), (iii) when the focal pup's partner was silent (pups A-H), and (iv) when the focal pup's partner was also vocalising (pups A-H). Hereafter, we refer to these conditions as: (i) alone, (ii) 1-way interaction with a broadcasted partner, (iii) silent partner, and (iv) 2-way interaction with a real partner. Notice that only some pups entered each condition and vice-versa (Table S1).

Our sample size was affected by the unpredictable arrival of animals at the Sealcentre, which varies seasonally. Only medically stable and healthy pups were included in this study. The number of daily recording sessions per pup pair varied based on veterinary staff recommendations at the rehabilitation centre (Table S1); we did not record pups with signs of disease. Data from pup I (i.e., the alone and 1-way contexts both with and without playback stimuli) have been re-analysed from previous studies [28,21, respectively] and combined with unpublished data from pups A-H (the silent partner and 2-way contexts) (Table S1). The contribution of each pup to each analysis is shown in Table S2.

1.6. Extraction of temporal variables, definition of call bout and vocal interactions

We first extracted the onsets and offsets of each pup call recorded in each behavioural context (Figure 1, Method S3). From these values, we calculated rhythmic metrics such as call duration, inter-onset intervals (IOIs), ratios of adjacent IOIs, and inter-call intervals (ICIs; i.e., silent gaps). Each IOI was obtained by subtracting the *onset* of call n from the *onset* of call $n+1$, while the ICI was calculated by subtracting the *offset* of call n from the *onset* of call $n+1$ (i.e., IOI minus duration of call n). Calls were organised into bouts, defined here as a series of at least three subsequent calls that were separated from adjacent bouts by a period greater than 1.5 times the median ICI of the recorded individual(s) calls (Figure 1). The response phase was computed as the ratio of the 'response IOI' (i.e., time interval between the call onset of the partner and the call onset of the pup's response) and the previous IOI, multiplied by 360, resulting in a unit vector with an angle on a circle. A vocal interaction was defined as a group of three calls within the same bout, which includes two calls from the broadcasted/real

partner followed by the response of the focal individual (Figure 1). Following previous methodology [28,38], we calculated the IOI ratio, r_k, for each pair of adjacent IOIs, t_k and t_{k+1}, in a bout as:

$$r_k = \frac{t_k}{t_k + t_{k+1}}$$

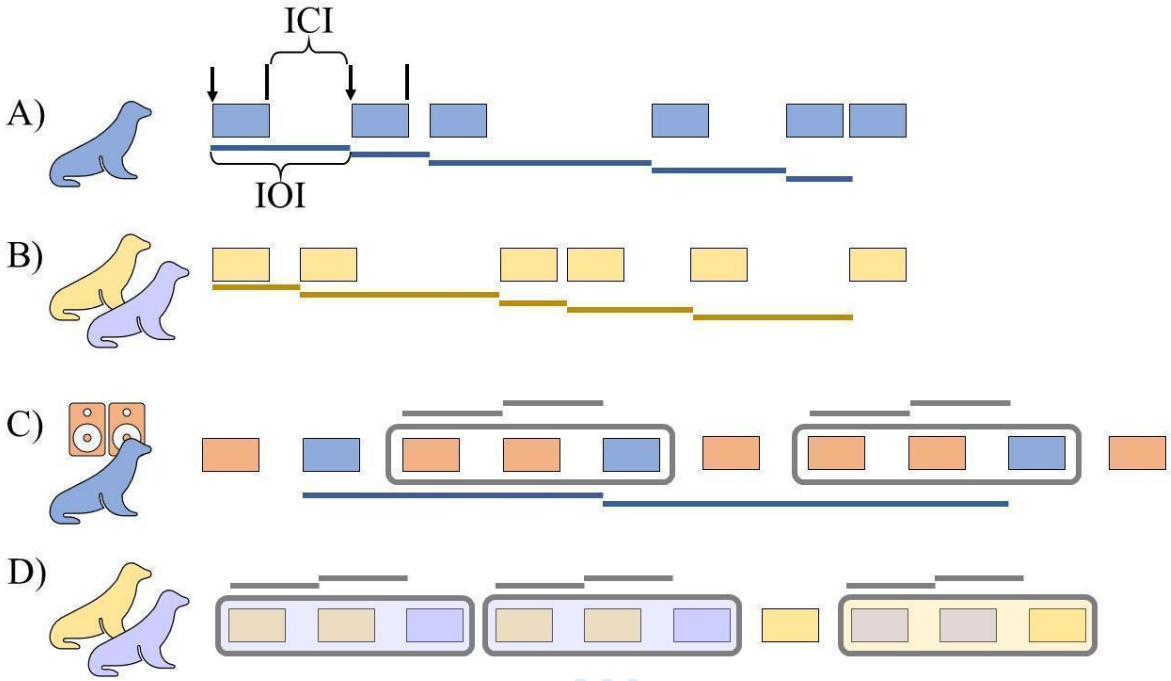

Figure 1. Schematic of experimental conditions and key measurements. Each of the four rows represents one bout and indicates a different behavioural context: (A) alone, (B) silent partner, (C) 1-way interaction, and (D) 2-way interaction. Boxes are coloured to represent the call source and grey rectangles denote vocal interactions. For bouts (C) and (D), calls are isochronously spaced for ease of visualisation, but empirical patterns were not necessarily isochronous. Horizontal bars indicate how IOIs were calculated for different analyses. The call onsets (black arrows) and offsets (black lines) are shown for the first two calls in (A).

1.7. Descriptions of analytical approaches

Circular statistics, wherein periodic measures are converted to angles on a circle and compared to distributions of interest [40], were used to investigate rhythmic periodicities in pup call response phases (Method S4). Following [21], we considered the values of the response phases as circular data falling between 0° and 360°. We obtained the circular mean (μ) (i.e., the average direction of the response phases calculated from the pup calls; Table S3). Then, we ran Rayleigh z-tests to investigate whether the distribution of response phases was uniform (e.g., arousal hypothesis) or showed a unimodal peak (Table S4) [21]. Subsequently, we tested for uniformity against a specified mean direction for the unimodal peak using a V-test [40–43].

Our data met the assumptions for circular statistics. We tested if the response phases in all four behavioural contexts followed a von Mises distribution using one-sample Watson tests (Table S5). With deviations from uniformity (null hypothesis von Mises distribution rejected), we used Kuiper's

test, Watson's test, and Rao's spacing test, to confirm the p-value obtained from the Rayleigh test (Table S6) as suggested by Landler et al. [40]. More details are shown in Method S4.

We then tested whether response phase distributions varied depending on the presence of the calling partner. We expected that vocally interacting pups would adjust their responses to broadcasted (1-way interaction) or real (2-way interaction) conspecific calls to avoid overlap and, hence, their response phases would show a unimodal distribution. Following previous work [21] and applying the V-test, we tested the null hypothesis of call response phase uniformity against two alternative unimodal departures: 0° (i.e., synchrony) and 90° (i.e., asynchrony). Using Watson's two-sample U^2 test [44], we also compared the call phase distributions of 1) a pup calling alone vs. when responding to a broadcasted partner (pup I) and 2) a pup calling in the presence of a silent partner vs. when their partner was also calling (pups A-H). For interacting (1-way or 2-way) pups, we applied Watson's two-sample U^2 test to assess whether the type of partner (i.e., real or broadcasted) differentially affected the pups' response timing. We predicted that pups interacting with a real partner would show more adaptive call timing, thanks to potential communicative cues from other modalities. Lastly, we compared the distributions of call phases of the single pup vocalising alone to those of the paired pups when their partner was silent to test whether the simple presence of a silent partner affects individual call timing. In both behavioural contexts, we predicted that calling patterns for pups without a responsive partner would show a different rhythmic structure to those observed in interaction.

Categorical rhythm analysis tests whether the temporal intervals between signal onsets, as inferred from IOI ratios, are distributed categorically rather than uniformly. We predicted that empirical and simulated null ratio distributions (i.e., the expected distribution if no rhythmic categories exist) will not differ when a pup is alone or with a silent partner but will differ when a pup is vocally interacting. For vocally interacting pups, we predicted a significant peak in empirical ratio distributions at the 4:1 small integer ratio based on the lone seal in [21], which called at approximately one-quarter of the playback's period. All categorical rhythm analyses were done within bouts following previous methodology (section Method S5) [23,24], with IOIs calculated in various ways depending on the behavioural context (Table S8, Figure 1). We used one-sample Kolmogorov-Smirnov (KS) tests to determine whether empirical IOI ratio distributions significantly differed from simulated null IOI ratio distributions. Our data met the one-sample KS test assumptions, namely that the sample is random and the theoretical distribution is continuous and fully defined. When the empirical and simulated distributions were significantly different, we also looked for evidence of small integer ratio (SIR) categorical rhythms—specifically at the 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, and 4:1 ratios—which have been found in other species' vocalisations [23,24]. In these analyses, the empirical ratio distributions were divided into "on-integer" and "off-integer" ratio bins (Table S7). On- and off-integer bin counts for each SIR were normalised by bin size and compared using a paired Wilcoxon

signed-rank test (a non-parametric test that allows for non-normality in the population data and assumes paired differences are continuous, symmetrically distributed, and mutually independent). When sample sizes allowed, we used two-sample KS tests (having met the assumption of mutual independence of measurements within samples) to determine whether the ratio distributions of individual pups differed across behavioural contexts.

The **Granger causality** test investigates whether the values of a time series A are better predicted when considering the values from a second time series B, as opposed to only using values from time series A [45]. Here, we assessed whether the call timing of a pup partaking in a 1-way or 2-way vocal interaction can be predicted using the call timing of its partner. More specifically, to investigate whether the call timing of a pup differed in relation to the type of partner, we tested whether there is a difference in predicting the time series of the pup interacting with a broadcasted partner vs. time series of the pups interacting with a real partner. Previous work showed that individuals respond to conspecific calls with a non-random pattern [18–20,46]. We therefore expected that the time series of a pup can be better predicted considering the time series of a vocalising partner rather than those of a broadcasted signal. In both the 1-way and 2-way interactive scenarios, we considered Granger causality at two levels: (1) the entire recording, regardless of the length of the pauses between consecutive calls, and (2) different bouts within each recorded session. We restricted the analysis on the different bouts to call sequences that were long enough to generate accurate estimates (i.e., a minimum of five paired calls [47]). The bouts included in this analysis range from 5 to 20 calls. We conducted the Granger causality test using call onsets and different *lag* measures, from one to five (Method S6), testing whether the previous one to five onsets in the first time series can be used to better predict the second time series (Figures S5/S6, Table S11). For the 1-way interaction, we performed a one-way analysis, considering whether the pup's timing could be predicted using the playback timing. For the 2-way interaction, we performed a two-way analysis to assess whether the two interacting pups influenced the timing of each other's calls.

We used **ADAM** to test for evidence of reactive error correction and predictive processes in the 1-way and 2-way interactive scenarios. ADAM consists of three computational modules that interact via internal models of 'self' and 'other' that support one's own action planning and external event prediction, respectively (Figure S2). The *adaptation module* compensates for synchronisation errors by implementing error correction processes that alter the phase and/or period of an internal timekeeper controlling for action (here, call) timing. These error correction processes determine the provisional timing of the next planned action by providing input to an internal model of the 'self'. The *anticipation module* computes the expected timing of upcoming events based on the weighted sum of two processes: the linear extrapolation of previous IOIs in the sequence and the copying (or 'tracking') of the previous IOI, with the output informing temporal predictions generated by the

'other' internal model. Finally, a *joint module* integrates and compares the output of the adaptation and anticipation modules and compensates for discrepancies by implementing an anticipatory error correction process before the next motor command is issued. The joint module thus reduces potential temporal mismatches between action plans in 'self' internal models and temporal predictions in 'other' internal models, thereby regulating the balance between the integration (merging) and segregation (distinction) of information about 'self' and 'other' [48,49]. Each process instantiated in ADAM is controlled by an independent parameter, and the value of these parameters can be estimated for a particular individual by fitting the model to behavioural time series data [26,27,50,51]. Parameter estimates were obtained for both the adaptation-only version of ADAM—which includes phase correction and period correction—and the full ('joint') version—including period correction, temporal prediction/tracking, and anticipatory error correction. Both versions of ADAM were applied to each interactive context because it is not possible to know a priori whether the pups' call sequences (real or broadcasted) have a steady base tempo (for which adaptation is sufficient) or a systematically changing tempo (which benefits from both anticipation and adaptation) (Method S7).

Table 1. Summary table showing in order: research questions, analyses, contributing pups, predictions, whether the data supports each prediction, statistical test(s) used, and result(s) obtained. The column "Supported by data?" has three possible answers: results support the prediction (Y), results only partially support the prediction (Partial), and results do not support the prediction (N). Due to sample size and/or analytical requirements, not all pup data could be used in each analysis. The acoustic variables of interest for the different analytical approaches were: response phases (circular statistics), IOI ratios (categorical rhythms), IOIs (Granger causality, ADAM), and asynchronies (ADAM).

Research question	Analytical approaches	Pups	Prediction	Supported by data?	Statistical test(s) and result(s)
Temporal domain, Q1: *Does the presence of a calling partner affect the call timing of individual pups?*	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
			Pups will call in asynchrony to avoid overlap during vocal interactions.	Y	V-test: Pup calls start at one-quarter of the partner's period
			Response phases will be affected by the presence of a vocalising partner.	Y	Watson's U^2 test: Response phase distributions differ between non-interactive and interactive contexts.

	Categorical rhythms	A, B, C, E, H, I	Empirical and chance ratio distributions will only significantly differ when pups are vocally interacting.	Partial	1-sample KS tests: Simulated and empirical ratio distributions are rarely significantly different (exceptions: pup I alone, pup I 1-way, pup B 2-way).
		A, B, I	Vocally interacting pups will have a significant peak at the 4:1 SIR.	N	Paired Wilcoxon signed-rank test: No significant peaks at any of the tested small integer ratios.
			Ratio distributions of individual pups will differ across behavioural contexts.	Y	2-sample KS tests: Ratio distributions significantly differed for the same individuals in different behavioural contexts.
Social domain, Q2: *Does the type of calling partner (real or broadcasted) affect the call timing of individual pups?*	Circular statistics	A-I	Pups interacting with a real partner will show more adaptive call timing than the pup interacting with a broadcasted partner.	N	Watson's U^2 test: No difference in response phases between 1-way and 2-way interactions.
	Granger causality	A, B, C, D, E, F, I	The time series of a pup will be better predicted considering the time series of a vocalising partner rather than the time series of a broadcasted signal.	Y	Bidirectional & unidirectional causality: Interaction with a real partner impacted the pup's vocal behaviour more than the playback. Mutual temporal adaptation among pairs of vocally interacting pups.
Social domain, Q3: *When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing?*	Circular statistics	A-I	The calling pattern of pups will show similar rhythmic structure in both the alone and silent partner conditions, as no vocal interaction is taking place in both cases.	Y	Watson's U^2 test: No difference between response phase distributions of pup calling alone and pups calling with a silent partner.

Cognitive domain, Q4: *Which timing mechanisms are used by pups during vocal interactions?*	ADAM	A, B, I	In the 1-way interaction, pups may show sensitivity to (non-)interactivity which would be reflected by parameter changes over repeated sessions. The 2-way interaction may be mediated by basic temporal adaptation and possibly higher-level anticipatory timing.	Partial	1-way interaction: Temporal anticipation, and to a lesser extent also adaptation, decreased across the playback sessions, and was absent in the final session. 2-way interaction: Clear evidence for temporal adaptation, with differing parameter estimates for each seal pup suggesting the emergence of different interactive roles.

3. Results

3.1. Does the presence of a calling partner affect the call timing of individual pups? (Q1)

All pups' data entered the circular statistics analysis. Running the Rayleigh test, we found that the response phase distribution was uniform for pup I which was recorded alone ($z = 0.04$, $p = 0.254$; Figure 2A), whereas it was non-uniform for pups A-H which were recorded with a silent partner ($z = 0.11$, $p < 0.001$; Figure 2A; Table 1). This non-uniformity may have been driven by the individual contributions of pups B and C, which had non-uniformly distributed response phases (Table S4), whereas the other six pups had a uniform distribution. The Rayleigh tests run anew in the interactive contexts, showed that the response phase distributions of pup calls were unimodal in both the 1-way (pup I: $z = 0.39$, $p < 0.001$) and 2-way (pups A - H: $z = 0.41$, $p < 0.001$) interactions (Figure 2B; Table 1). Applying the V-test in both contexts, the direction of the response phases did not statistically match 0º (1-way: $z = -0.02$, $p = 0.587$; 2-way: $z = 0.06$, $p = 0.110$), suggesting that pups did not synchronise with their partner (real or broadcasted). However, the response phase direction did match 90º (1-way: $z = 0.38$, 2-way: $z = 0.41$, $p < 0.001$), supporting the previously reported evidence of asynchronous calling behaviour [21]. A Watson's two-sample U^2 test confirmed that the response phase distributions significantly differed between the alone vs. 1-way interaction context for pup I (U^2 1.76, $p < 0.001$), and between the silent partner vs. 2-way interaction context for pups A-H ($U^2 = 0.78$, $p < 0.001$, Figure S3; Table 1). Finally, the circular standard deviation values were higher for the alone and silent partner contexts compared to both interactive vocal contexts, indicating a larger dispersion of the response phases for the former conditions. This outcome is also confirmed by the values for the mean resultant length.

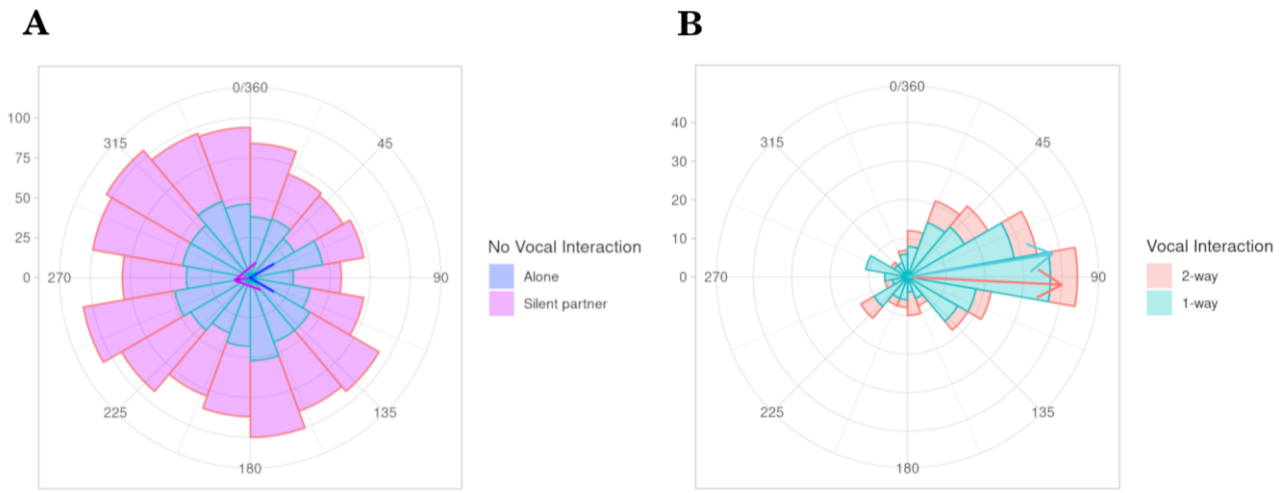

Figure 2. Circular histogram plots (bin width = 20˚) showing response phases in (A) both types of behavioural contexts with no vocal interaction (alone/silent partner), and (B) in both types of vocal interaction contexts (1-way/2-way). Angles are measured in degrees starting from 0˚ and going clockwise to 360˚. The arrows indicate the circular mean (μ) and colours correspond to the different behavioural contexts. The length of the arrow corresponds to the value of the mean resultant length (ρ).

In the categorical rhythm analyses, the empirical ratio distribution was significantly different from chance when pup I was recorded alone (Figure 3A) and during the 1-way interaction (Figure 3C). In both contexts, there was no evidence of significant peaks at any of the tested ratios (Table 1). When the playback calls were disregarded from IOI calculations (Figure 3B), there was no significant difference in empirical and simulated ratio distributions for pup I. However, pairwise KS tests showed that the ratio distributions significantly differed when comparing each of the three behavioural contexts (alone vs. 1-way interaction disregarding playback vs. 1-way interaction when pup I responds) to each other (Table S10; Table 1). For the five pups that were well-sampled in the silent partner context (pups A, B, C, E, and H; Table S8), the empirical ratio distributions did not significantly differ from chance (Table S9, Figures 3/S4, Method S4). Finally, when considering 2-way interactions, only pups A and B were well-sampled enough (i.e., had at least ten ratios for both the silent partner and 2-way interaction contexts) to compare, but the empirical and simulated ratio distributions were significantly different only when pup B was the responder (Figure 3; Table S8; Table S9). Once again, there were no significant peaks at any of the tested ratios for pup B. For both pups A and B, the empirical ratio distributions significantly differed when comparing different behavioural contexts (Table S10). Collectively, there was thus little evidence of SIR rhythmic categories in pup calls, but IOI ratios did significantly differ when looking at the same individuals in different behavioural contexts (Table 1).

Thus, the results from both analyses suggest that the *presence* of a calling partner does affect the call timing of the focal pup, in terms of both call response phases and IOI ratios.

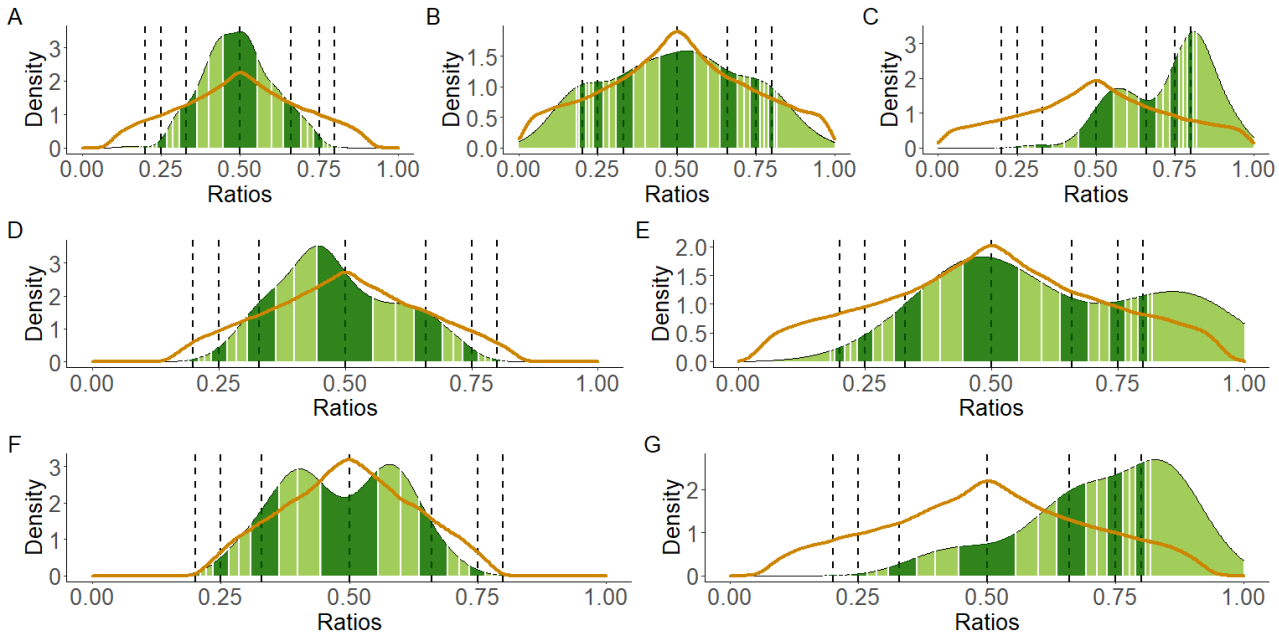

Figure 3. IOI ratio density plots for pups in different behavioural contexts. Pup I: (A) alone, (B) 1-way interaction (disregarding playback), (C) 1-way interaction (responding to playback). Pup A: (D) silent partner, (E) 2-way interaction (responding to partner). Pup B: (F) silent partner, (G) 2-way interaction (responding to partner). For each plot, the dashed vertical lines indicate, from left to right, 1:4, 1:3, 1:2, 1:1 (i.e., isochrony), 2:1, 3:1, and 4:1 small integer ratios. On-integer ratio ranges are in dark green and off-integer ratio ranges are in light green, with white lines and black dashed lines denoting bin boundaries. The orange curves indicate the ratio distribution expected under a uniform distribution if no rhythm categories exist. The empirical ratio distribution significantly differed from the simulated ratio distribution for panels (A), (C), and (G) only. Note that the scale of the y-axes differ.

3.2. Does the type of calling partner (real or broadcasted) affect the call timing of individual pups? (Q2)

To address this research question, we once again used circular statistics and the response phase distributions of all nine pups. Specifically, we compared calls from pup I during the playback (1-way interaction) with calls of pups A-H when their partner was also calling (2-way interaction). A Watson's two-sample U^2 test statistically confirmed that the response phase distributions did not differ between the 1-way and 2-way vocal interactions ($U^2 = 0.07$, $p > 0.10$; Table 1).

Interestingly, however, the Granger causality results from seven pups (A, B, C, D, E, F, I) showed that call timing behaviour differed depending on the type of partner (Table 1). For the 1-way interaction, five different playback sessions featuring pup I were considered (ranging from 34 to 121 paired calls) and the timing of the pup's calls were never significantly predicted by the timing of the playback (Table S11, Figure S5). For the 2-way interaction, five different recording sessions were considered (ranging from 8 to 71 paired onsets). We found that in two sessions, the timing of the first pup did not significantly predict the timing of the second, in either direction (pair A>B and B>A; pair E>F and F>E; Table S11, Figure S6B/S6D). However, in two other sessions, the timing of a pup was significantly predicted by the calling partner in both directions, across different lag values (A>B: lag-

1, lag-3, lag-4, lag-5; B>A: lag-3, lag-4, lag-5; C>D and D>C: lag-2 and lag-3; Table S11, Figure S6B/S6C). In the last session, the timing of pup A could be predicted by that of pup B, while we found no indication of temporal adjustment for pup B. At the bout level, we restricted our analyses to four series of paired onsets featuring pups A and B (as they were the only pup pair that met the sample size requirements). In this last scenario, only the timing of pup B was significantly influenced by pup A (Figure S6A).

While the circular statistics results thus suggest that the type of calling partner (real vs. broadcasted) does not affect focal pup call timing, the Granger causality results suggest that in certain 2-way (but not 1-way) interactions, focal pup call timing can be predicted by the partner's call timing.

3.3. When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing? (Q3)

Using circular statistics, we compared the calls of pup I recorded alone with those of the 8 other pups (A-H) recorded with a silent partner present. Watson's two-sample U^2 test results show that the response phase distributions did not statistically differ between the two contexts ($U^2 = 0.10$, $p > 0.05$). In other words, having a silent partner was essentially the same as having no partner—in terms of the effect on focal pup response phase distributions—for the pups in our study. This is intuitive, given that in both behavioural contexts, there is no acoustic stimuli to "respond" to.

3.4. Which timing mechanisms are used by pups during vocal interactions? (Q4)

ADAM parameter estimation was conducted on call data from all five playback sessions featuring pup I (1-way, Figure 4 A/B), but was restricted to pups A and B for the 2-way interactions due to sample size limitations (Figure 4 C/D). Interactive vocal bouts were concatenated to obtain a time series length which would provide reliable ADAM parameter estimates. A simulation test then ensured that the estimates were not compromised by differing sequence lengths or by the concatenation procedure (Method S7). The reliability of observed parameter estimates was tested by comparison against corresponding values for randomly permuted data (Method S7). The quality of the fits to the data did not differ significantly between versions of ADAM (Method S7). Results for the 1-way interaction featuring pup I were remarkable with regards to typical human data (e.g. [26,27]) as most significant parameters were negative in sign (Figure 4A/B, Tables S13/S14). Negative phase and period correction estimates indicate that calling earlier will lead to a shortening of the next IOI, while calling later will lead to a lengthening of the next IOI. Negative prediction/tracking estimates mean that when the playback's IOIs increase (i.e., deceleration), the pup's IOIs will decrease (i.e., acceleration), and vice versa. This systematically enhances the timing distinction between calls, possibly testing the responsiveness of the partner (i.e., playback) by

introducing timing asynchronies and gauging their effects. It is worth noting that evidence of such behaviour generally decreased across the five playback sessions, with no significant parameter estimates emerging in the final session.

For the 2-way interaction (Figure 4C/D, Tables S13/S14), the parameter estimation procedure was run twice, each time with one of the two pups serving as the focal individual and the other as the external reference. With this procedure, similar parameter estimates for each pup would indicate a symmetrical pattern of influence, whereas different estimates would indicate asymmetrical influence. Results for this type of vocal interaction showed evidence for temporal adaptation. In particular, parameter estimates were consistent with pup A engaging in period correction while pup B engaged in both phase and period correction. Period correction estimates were positive in sign for pup A and negative for pup B. Pup A's positive period correction estimates suggest that calling earlier will lead to a lengthening of the next IOI, whereas calling later will lead to a shortening of the next interval. Pup B exhibited the opposite effect but to a lesser degree. Pup A thus implemented a timing mechanism that increased the similarity in their calling rates, while pup B implemented a timing mechanism that increased the distinction in calling rates. It should be noted, however, that the observed parameter estimates do not necessarily reflect individual call styles per se, but rather the roles that dynamically emerge within the context of this specific pairing of pups.

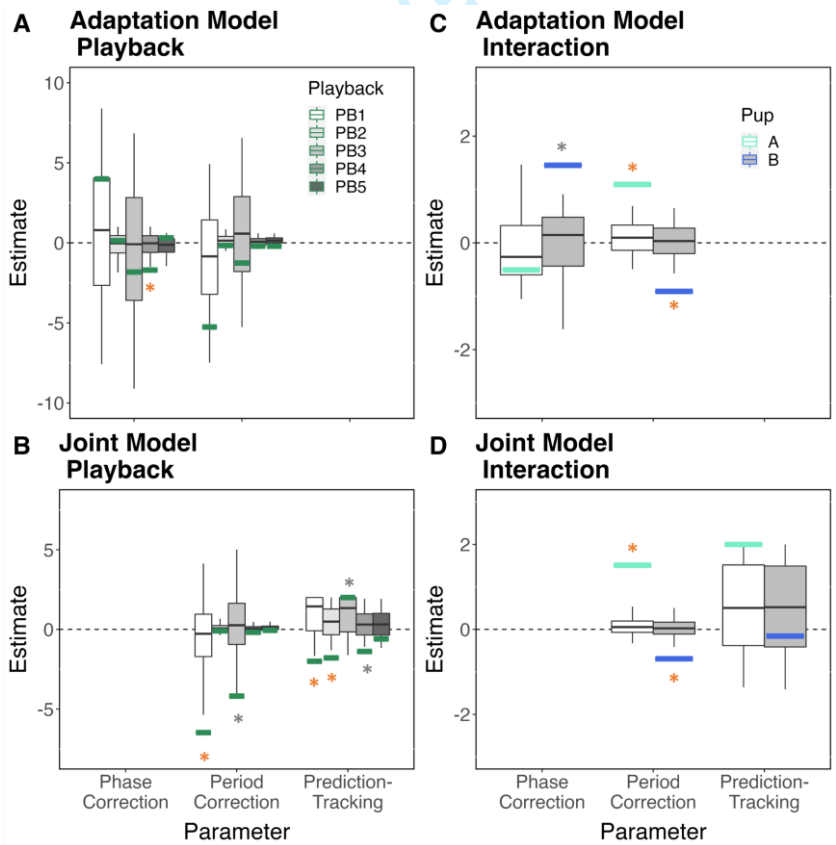

Figure 4. ADAM parameter estimates for seal pups A, B, and I. Left panels (A & B) show parameter estimates for pup I in five playback sessions (1-way interaction); right panels (C & D) show estimates for pups A and B (2-way interaction). Top panels (A & C) show phase correction and period estimates obtained with the adaptation-only version of ADAM; bottom panels (B & D) show period correction and prediction-tracking estimates for the full 'joint' model (anticipatory error correction estimates are not shown). Parameter estimates are shown separately for pup A relative to pup B (aquamarine bars) and pup B relative to pup A (blue bars) in the recording session, and for pup I (green bars) relative to playback calls in separate playback sessions (PB1-PB5). Corresponding parameter estimates for randomly permuted data from each session are represented by box plots, with the central horizontal line indicating the median, the bottom and top edges of the box indicating the 25th and 75th percentiles, and the whiskers indicating the 5th and 95th percentiles. Real estimates with asterisks are significantly different from permuted data estimates at the 2-tailed (orange) or 1-tailed (grey) level.

4. Discussion

This study sought to provide a methodological proof of concept for quantifying vocal rhythmicity in non-human animal interactive communication. Particularly, we showed how complementary methodologies can be used to test whether the (Q1) presence and (Q2) type of a vocalising partner, or the (Q3) presence of a silent partner affect patterns in animal communication (Table 1). We also investigated which (Q4) underlying cognitive mechanisms potentially play a role in vocal interactions (Table 1). The different analytical techniques proved fruitful; combining and contrasting their results could highlight nuances in rhythmic behaviour. Nonetheless, our sample size is undeniably small, and our opportunistic design (e.g., mixing within- and between-individual comparisons) cannot provide strong inference. We believe that our experimental setup, combined with the approaches we present, can be adapted and expanded in future work to better understand the temporal, social, and cognitive processes underlying interactive communication in animals.

Q1. The presence of a vocalising partner influences rhythm in vocal interactions

The prediction of overlap avoidance was supported by circular statistics, which showed that the distributions of response phase angles were unimodal. When vocally interacting, pups responded at approximately one-quarter of the playback/real partner calling period (90◦) and showed phase angles significantly different from 0◦, supporting previous results from one individual (pup I) [21]. Pups responding to conspecifics seem to time their calls to avoid overlap, consistently de-synchronizing their call onsets. When pups were recorded alone or with a silent partner, the distribution of phase angles was uniform, with no significant rhythmic pattern. Our findings confirm and extend previous outcomes [21]; in our limited sample, pups not only react to playbacks with asynchrony, but also respond to conspecific partner calls in an asynchronous manner.

Regarding categorical rhythms, the empirical ratio distribution did not significantly differ from the distribution expected by chance if IOIs were uniformly distributed for most pup/behavioural context combinations (7/10). For the remaining three combinations (pup I alone; pup I, 1-way interaction; pup B, 2-way interaction), there was no significant evidence that call timing exhibited

categorical rhythms at the seven tested small integer ratios. This includes the 4:1 ratio, which we hypothesised would frequently occur when pups were vocally interacting with playback stimuli or partners based on previous work [21]. Importantly, however, the empirical ratio distributions significantly differed across behavioural context conditions for the three pups (A, B, I) for whom such comparisons were possible. When alone or with a silent partner, pup calls generally showed unimodal ratio distributions centred around isochrony, whereas calls emitted by vocally interacting pups showed a clear right-shifted peak, or a second peak to the right of isochrony, indicating that the pup's IOIs during vocal interactions were generally shorter than the playback's/partner's IOIs. This context-dependent asymmetry bolsters results related to Q4, namely that the presence of vocalising partner can significantly impact pup vocal behaviour, with interacting individuals trying to prevent call overlap. Roeske and colleagues [23] hypothesised that categorical rhythms may be prevalent in calls meant to attract and hold conspecific attention, such as the calls pups use to attract the attention of their mothers. We did not find evidence of rhythm categories at seven tested ratios; this negative result has methodological value, highlighting that not all species that produce attention-seeking vocalisations show integer ratio categories.

A vocal interaction with non-adaptive playback stimuli (1-way) may represent a limitation to studying spontaneous vocal production. The extent of this limitation can be gauged by comparing results to data from vocal interactions with a real partner (2-way), and with the use of complementary analytical methodologies, such as categorical rhythm analysis and circular statistics.

Q2. Type of calling partner (real or broadcasted) partly affects rhythm in interaction.

Findings from Q2 show contrasting results. Call phases were statistically the same no matter if pups could interactively adjust their call timing to each other (2-way) or not (in 1-way). Partly in contrast with this, the Granger causality analysis showed how vocally interacting with a real individual impacted the pup's vocal behaviour more than interacting with a broadcasted partner Indeed, we found evidence of mutual temporal adaptation among pairs of vocally interacting pups and, consistent with a recent hypothesis [6], conspecific interactions can be facilitated by the reciprocal adjustment of timing behaviour. In contrast, we found little evidence that a pup adjusts the time series of its calls to a playback series. Indeed, in most cases, the pup's call timing could not be predicted by the playback's call timing.

Together, findings from circular statistics (Q1) and Granger causality (Q2) point towards a directionality-overlap avoidance relationship, which has also been suggested for other non-human species (e.g., primates [18,52]; birds [53]; amphibians [54]; seals [55]). Interestingly, when infant marmosets interact with their parents, the probability that their vocalisations will overlap with those

of adults decreases over time [22], suggesting that turn-taking in some mammals is a learned vocal behaviour scaffolded by active parental feedback [39].

Q3. A silent partner does not trigger call rhythmicity

Circular statistics indicated that pups did not show any periodic pattern both, either when calling alone or with a silent partner present (Q1). This is consistent with the fact that wild pups produce calls to attract a silent mother [36]: if no other calling conspecifics are present, there is no need to adjust one's call timing and vocalisations are produced with a random onset. Moreover, given that the response phase distributions did not differ between pups recorded with a silent partner or alone, it suggests that the presence of a silent partner did not trigger variable calling behaviour in our study population.

Q4. Purported timing mechanisms for vocal interactions

The ADAM analysis suggests that seal pups may perceive temporal patterns [56], which arise between their calls and those of others, and react to them by adopting different mechanisms for temporal adaptation. The negative parameter estimates that we observed in the 1-way interaction sessions (consistent with enhancing the distinction between calls) could reflect attempts to lead the temporal interaction or even to test the responsiveness of the (broadcasted) partner by introducing timing perturbations and gauging their effects. The decrease of temporal adaptation and/or anticipation observed across playback sessions is consistent with a gradual process of habituation, with the pup possibly learning that the playback is non-interactive. Harbour seals are capable of acoustic recognition based on habituation paradigms; they can discriminate among different stimuli and selectively habituate to them [57]. The lack of temporal adaptation to the playback stimulus we observed in a seal pup may entail similar habituation processes. From a comparative perspective, these findings also suggest a sensitivity to social contingency that may be analogous to capacities in human infants, who become disinterested and display fewer signs of positive affect during vocal interactions with non-responsive or delayed video recordings of their mothers [58–62]. Future studies could address the role of temporal contingency by using interactive playback sequences [63–65].

Ours constitutes the first attempt to apply ADAM to non-human animals. On the technical side, this necessitated the validation of an approach where brief interactive vocal bouts were concatenated to produce longer time series and thereby reduce the risk of model overfitting (Discussion S1). Demonstrating the utility of this procedure opens the door to applying the model in a wider range of behavioural contexts. Nevertheless, caveats are necessary when interpreting the seal data in light of previous work with ADAM in humans, where individuals intentionally produce movements, whose sensory effects occur simultaneously with rhythmically regular sounds (e.g. [26,27]). Assumptions

about intentionality and simultaneity may not apply to seal pup vocal interactions or rhythmic interactions in other animals [6,9]. Points of convergence in the main outcomes of the complementary analysis techniques suggest that ADAM, like categorical rhythm analysis, circular statistics, and Granger causality, may also be robust and informative under such conditions.

5. Conclusions

Crucially, our work highlights the efficacy of combining multiple methods to study rhythmic vocal behaviour. Our approaches vary in the degree to which they capture global temporal characteristics across events vs. local temporal dependencies between events. Global measures (e.g., from circular statistics or categorical rhythm analysis) reveal predominant rhythmic features of a vocal interaction while local time series measures (e.g., from Granger causality or ADAM) provide information about how these features might arise. Through this approach we could tease apart rhythm nuances in our dataset, further develop harbour seals as a model species, and demonstrate how certain analyses often restricted to humans, such as categorical rhythms (but see [23,24]) and ADAM, can be applied to non-human animals. The categorical rhythm and circular statistical analyses tackled similar questions from different angles, namely whether the distributions of IOI ratios (the former) or call response phases (the latter) significantly differed across behavioural contexts. The categorical rhythm analyses also sought to determine whether the rhythmicity of pup calling behaviour conforms to small integer ratios. Our negative result is, to our knowledge, the first published case of a species for which categorical rhythms are clearly absent from vocalisations, which adds to understanding of how, why, and when such rhythms evolve in communication systems. Time series analyses such as Granger causality allowed testing for timing adjustment. In the case of ADAM, mechanisms of temporal adaptation and anticipation that have previously been used to describe rhythmic behaviour in humans [66] were used to describe rhythmic behaviour in seals. Interestingly, the ADAM model provides an empirical warning about potential seal pup habituation effects when vocally interacting with a recorded partner.

Though our sample size is small, studies on single individuals are not unusual in comparative research [20,67]. Nevertheless, it is possible that the lack of adjustment to a playback, the adaptation to a real individual, or both, reflect a peculiar vocal behaviour of the individuals we tested and cannot be generalised to the species as a whole. The pups in this study were in a temporarily captive setting, albeit in acoustic proximity to other individuals, similar to conditions they would experience in nature. Unfortunately, vocal development in harbour seal pups has not been studied in wild colonies, hence we do not know whether captivity affects their vocal development. We do know, however, that pups vocally interact with neighbouring pups in the colony and not with their silent mothers; hence

the turn-taking behaviour observed in our captive conditions might extend to the same behaviour in wild conspecifics.

Motivation for an individual to respond and engage in an interaction, with the closest partner in the colony, may depend on the degree of participation signalled by the partner. This, in turn, may be triggered by individual-specific behaviours or by cues from other modalities. Multimodal communication should be the target of future studies since we cannot assume that such interactions are limited to acoustic cues. Despite these limitations, our study shows that adopting multiple complementary approaches can be a fruitful way to study rhythmic interactivity in non-human animal communication.

Acknowledgements

We are grateful to the editor and the two anonymous reviewers who provided insightful comments and suggestions on an earlier version of this manuscript. We thank all the volunteers, seal care members and veterinarians of Sealcentre Pieterburen who helped during the experiment and in the rehabilitation process of the seal pups. We are thankful to Chiara De Gregorio for sharing the R code used for the categorical rhythm analyses and to Laura Verga for providing helpful comments on a draft of this manuscript.

References

1. Dusenbery DB. 1992 *Sensory ecology: How organisms acquire and respond to information.* Freeman, New York.
2. Stegmann UE. 2013 *Animal Communication Theory: Information and Influence.* Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9781139003551)
3. Anichini M, de Heer Kloots M, Ravignani A. 2020 Interactive rhythms in the wild, in the brain, and in silico. *Can. J. Exp. Psychol.* **74(3)**, 170-175. (doi: 10.1037/cep0000224)
4. Ravignani A, Bowling DL, Fitch WT. 2014 Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology.* **5**,1118. (doi:10.3389/fpsyg.2014.01118)
5. Greenfield MD, Aihara I, Amichay G, Anichini M, Nityananda V. 2021 Rhythm interaction in animal groups: selective attention in communication networks. *Phil. Trans. R. Soc. B* **376(1835)**, 20200338. (doi:10.1098/rstb.2020.0338)
6. 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)
7. Levinson SC. 2006 Cognition at the heart of human interaction. *Discourse studies*. **8(1)**, 85-93.
8. Levinson SC. 2016 *The Interaction Engine hypothesis.* The Language in Interaction Summerschool on Human Language: From Genes and Brains to Behavior. Berg en Dal, The Netherlands.
9. Ravignani A, Verga L, Greenfield MD. 2019 Interactive rhythms across species: the evolutionary biology of animal chorusing and turn-taking. *Annals of the New York Academy of Sciences,* **1453(1)**, 12-21. (doi:10.1111/nyas.14230)
10. Levinson SC. 2016 Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences*. **20(1)**, 6-14. (doi.org/10.1016/j.tics.2015.10.010)
11. Pika S, Wilkinson R, Kendrick KH, Vernes SC. 2018 Taking turns: bridging the gap between human and animal communication. *Proc. R. Soc. B* **285(1880)**, 20180598. (doi:10.1098/rspb.2018.0598)
12. Demartsev V, Strandburg-Peshkin A, Ruffner M, Manser M. 2018 Vocal turn-taking in meerkat group calling sessions. *Curr. Biol*. **28(22)**, 3661-3666. (doi: 10.1016/j.cub.2018.09.065)
13. Schulz TM, Whitehead H, Gero S, Rendell L. 2008 Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Anim. Behav.* **76(6)**, 1977-1988. (doi:10.1016/j.anbehav.2008.07.032)
14. Takahashi DY, Narayanan DZ, Ghazanfar AA. 2013 Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys, *Curr. Biol.* (doi:10.1016/j.cub.2013.09.005)
15. Levinson SC. 2020 On the human" interaction engine". In *Roots of human sociality*, pp. 39-69. Routledge.
16. Yoshida S, Okanoya K. 2005 Evolution of turn-taking: a bio-cognitive perspective. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, **12(3)**, 153-165. (doi:10.11225/jcss.12.153)
17. Henry MJ, Cook PF, de Reus K, Nityananda V, Rouse AA, Kotz SA. 2021 An ecological approach to measuring synchronization abilities across the animal kingdom. *Phil. Trans. R. Soc. B*. **376(1835)**, 20200336. (doi:10.1098/rstb.2020.0336)
18. 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)
19. Clink DJ, Tasirin JS, Klinck H. 2020 Vocal individuality and rhythm in male and female duet contributions of a nonhuman primate. *Curr. Zool.* **66(2)**, 173–186. (doi:10.1093/cz/zoz035)
20. Fuhrmann D, Ravignani A, Marshall-Pescini S, Whiten A. 2014 Synchrony and motor mimicking in chimpanzee observational learning. *Sci. Rep.* **4**, 1–7. (doi:10.1038/srep05283)
21. Ravignani A. 2019 Timing of antisynchronous calling: A case study in a harbour seal pup (Phoca vitulina). *J Comp Psychol.* **133(2)**, 272–277. (doi:10.1037/com0000160)

22. Takahashi DY, Fenley AR, Ghazanfar AA. 2016 Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Phil. Trans. R. Soc. B: Biological Sciences*. **371(1693)**, 20150370. (doi:10.1098/rstb.2015.0370)

23. Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical rhythms are shared between songbirds and humans. *Curr. Biol.* **30**, 3544–3555. (doi:10.1016/j.cub.2020.06.072)

24. De Gregorio C, Valente D, Raimondi T, Torti V, Miaretsoa 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)

25. van der Steen MC, Keller PE. 2013 The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience,* **7**, 253. (doi:10.3389/fnhum.2013.00253)

26. van der Steen MC, Jacoby N, Fairhurst MT, Keller PE. 2015 Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Res*. **1626**, 66–87. (doi:10.1016/j.brainres.2015.01.053)

27. Mills PF, Harry B, Stevens CJ, Knoblich G, Keller PE. 2019 Intentionality of a co-actor influences sensorimotor synchronisation with a virtual partner. *Quarterly Journal of Experimental Psychology,* **72(6)**, 1478–1492. (doi:10.1177/1747021818796183)

28. Ravignani A, Kello CT, de Reus K, Kotz SA, Dalla Bella S, Méndez-Aróstegui M, Rapado-Tamarit B, Rubio-Garcia A, de Boer B. 2019 Ontogeny of vocal rhythms in harbour seal pups: an exploratory study. *Curr Zool.* **65(1)**, 107–120. (doi: 10.1093/cz/zoy055)

29. Patel AD. 2006 *Musical rhythm, linguistic rhythm, and human evolution.* Music Percept **24**,99–104.

30. Patel AD. 2014 The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol* **12**, 1001821. (doi:10.1371/journal.pbio.1001821)

31. Keehn RJJ, Iversen JR, Schulz I, Patel AD. 2019 Spontaneity and diversity of movement to music are not uniquely human. *Curr. Biol.* **29(13)**, R621-R622. (doi: 10.1016/j.cub.2019.05.035)

32. Schusterman RJ. 1977 Temporal patterning in sea lion barking (Zalophus californianus). *Behavioral biology*. **20(3)**, 404-408. (doi:10.1016/S0091-6773(77)90964-6)

33. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*. **63(5)**, 1050-1056. (doi:10.1139/z85-157)

34. Cook P, Rouse A, Wilson M, Reichmuth C. 2013 A California sea lion (Zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* **127(4)**, 412-27. (doi:10.1037/a0032345)

35. Torres Borda L, Jadoul Y, Rasilo H, Salazar Casals A, Ravignani A. 2021 Vocal plasticity in harbour seal pups. *Phil.Trans. R. Soc. B* **376**, 20200456. (doi:10.1098/rstb.2020.0456)

36. Sauvé CC, Beauplet G, Hammill MO, Charrier I. 2015 Acoustic analysis of airborne, underwater, and amphibious mother attraction calls by wild harbor seal pups (Phoca vitulina). *Journal of Mammalogy* **96(3)**, 591–602. (doi:10.1093/jmammal/gyv064)

37. Perry EA, Renouf D. 1988 Further-studies of the role of harbor seal (Phoca vitulina) pup vocalizations in preventing separation of mother pup pairs. *Can. J. Zool*. **66**, 934–938. (doi:10.1139/z88-138)

38. Khan CB, Markowitz H, McCowan B. 2006 Vocal development in captive harbor seal pups, Phoca vitulina richardii: age, sex, and individual differences. *J. Acoust. Soc. Am.* **120(3)**, 1684-94. (doi:10.1121/1.2226530)

39. Chow CP, Mitchell JF, Miller CT. 2015 Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. R. Soc. B* **282**:20150069. (doi:10.1098/rspb.2015.0069)

40. Landler L, Ruxton GD, Malkemper EP. 2018 Circular data in biology: advice for effectively implementing statistical procedures. *Behav. Ecol. Sociobiol.* **72(8)**, 1–10. (doi:10.1007/s00265-018-2538-y)

41. Ruxton GD. 2017 Testing for departure from uniformity and estimating mean direction for circular data. *Biol. Lett.* **13**, 20160756. (doi:10.1098/rsbl.2016.0756)

42. Pewsey A, Neuhäuser M, Ruxton GD. 2013 *Circular statistics in R*. Oxford University Press.

43. Zar JH. 1999 Circular distributions: hypothesis testing. *Biostatistical analysis*, 624–668.
44. Landler L, Ruxton GD, Malkemper EP. 2021 Advice on comparing two independent samples of circular data in biology. *Scientific reports*. **11(1)**, 1-10. (doi:10.1038/s41598-021-99299-5)
45. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica.* **37**, 424–438. (doi: 10.2307/1912791)
46. Ravignani A, Norton P. 2017 Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution* 4–19. (doi:10.1093/jole/lzx002)
47. Zeileis A, Hothorn T. 2002 Diagnostic Checking in Regression Relationships. *R News*, **2(3)**, 7–10. (https://CRAN.R-project.org/doc/Rnews/)
48. Keller PE, Novembre G, Hove MJ. 2014 Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369(1658)**, 20130394. (doi:10.1098/rstb.2013.0394)
49. Keller PE, Novembre G, Loehr J. 2016 Musical ensemble performance: Representing self, other and joint action outcomes. In S. S. Obhi & E. S. Cross (Eds.), *Shared representations: Sensorimotor foundations of social life*. pp. 280–310. Cambridge: Cambridge University Press. (doi:10.1017/CBO9781107279353.015)
50. Mills PF, van der Steen MC, Schultz BG, Keller PE. 2015 Individual differences in temporal anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception* **3(1-2)**, 13–31. (doi:10.1163/22134468-03002040)
51. van der Steen MC, Schwartze M, Kotz SA, Keller PE. 2015 Modeling effects of cerebellar and basal ganglia lesions on adaptation and anticipation during sensorimotor synchronization. *Annals of the New York Academy of Sciences,* **1337**, 101–110. (doi:10.1111/nyas.12628)
52. Lemasson A, Pereira H, Levréro F. 2018 Social basis of vocal interactions in western lowland gorillas (Gorilla g. gorilla). *J. Comp. Psychol.* **132(2)**, 141–151. (doi: 10.1037/com0000105)
53. Yang XJ, Ma XR, Slabbekoorn H. 2014 Timing vocal behaviour: experimental evidence for song overlap avoidance in Eurasian wrens. *Behav Processes* **103**, 84–90. (doi:10.1016/j.beproc.2013.11.011)
54. Song J, Sun R, Fang K, Zhang B, Fang G. 2020 Flexibility as a strategy for avoiding call overlap in male Anhui Treefrogs. *Asian Herpetological Research* **11(3)**, 230–239. (doi:10.16373/j.cnki.ahr.1900331)
55. Serrano A, Terhune JM. 2002 Antimasking aspects of harp seal (Pagophilus groenlandicus) underwater vocalizations. *J. Acoust. Soc. Am*. **112(6)**, 3083–3090. (doi:10.1121/1.1518987)
56. Verga L, Sroka MG, Varola M, Villanueva S, Ravignani, A. 2022 Spontaneous rhythm discrimination in a mammalian vocal learner. Biology Letters, **18(10)**, 20220316.
57. Deecke VB, Slater PJ, Ford JK. 2002 Selective habituation shapes acoustic predator recognition in harbour seals. *Nature* **6912**, 171–3. (doi:10.1038/nature01030)
58. Bigelow AE, DeCoste C. 2003 Sensitivity to Social Contingency From Mothers and Strangers in 2-, 4-, and 6-Month-Old Infants. *Infancy,* **4(1)**, 111–140. (doi:10.1207/S15327078IN0401_6)
59. Henning A, Striano T. 2011 Infant and Maternal Sensitivity to Interpersonal Timing. *Child Development,* **82(3)**, 916–931. (doi:10.1111/j.1467-8624.2010.01574.x)
60. Malloch S, Trevarthen C. 2018 The Human Nature of Music. *Frontiers in Psychology,* **9(1680)**. (doi:10.3389/fpsyg.2018.01680)
61. Soussignan R, Nadel J, Canet P, Gerardin P. 2006 Sensitivity to Social Contingency and Positive Emotion in 2-Month-Olds. *Infancy,* **10(2)**, 123–144. (doi:10.1207/s15327078in1002_2)
62. Stormark KM, Braarud HC. 2004 Infants' sensitivity to social contingency: a "double video" study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development,* **27(2)**, 195–203. (doi:10.1016/j.infbeh.2003.09.004)
63. Fairhurst MT, Janata P, Keller PE. 2013 Being and feeling in sync with an adaptive virtual partner: Brain mechanisms underlying dynamic cooperativity. *Cerebral Cortex,* **23**, 2592–2600. (doi:10.1093/cercor/bhs243)

64. Fairhurst MT, Janata P, Keller PE. 2014 Leading the follower: an fMRI investigation of dynamic cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner. *Neuroimage,* **84**, 688–697. (doi:10.1016/j.neuroimage.2013.09.027)

65. Repp BH, Keller PE. 2008 Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science,* **27**, 423–456. (doi:10.1016/j.humov.2008.02.016)

66. Harry B, Keller PE. 2019 Tutorial and simulations with ADAM: an adaptation and anticipation model of sensorimotor synchronization. *Biological Cybernetics,* **113(4)**, 397–421. (doi:10.1007/s00422-019-00798-6)

67. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. (doi:10.1016/j.cub.2009.03.038)

Measuring rhythms of vocal interactions: A proof of principle in harbour seal pups

Marianna Anichini[1,2,3,4], Koen de Reus[1,5,6], Taylor A. Hersh[1], Daria Valente[7], Anna Salazar-Casals[1,8], Caroline Berry[1], Peter E. Keller[9,10#] and Andrea Ravignani[1,10#]

[1] Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
[2] Department of Biological Sciences, Faculty of Natural Sciences, Norwegian University of Science and Technology Ålesund, Norway.
[3] Hanse-Wissenschaftskolleg Institute for Advanced Study, 'Brain' Research Area, Delmenhorst, Germany.
[4] Division Animal Physiology & Behaviour, Department for Neuroscience, Carl von Ossietzky University Oldenburg, Oldenburg, Germany.
[5] Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium.
[6] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
[7] Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.
[8] Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands.
[9] MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Australia.
[10] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark.
[#]Joint senior authors

Correspondence:
Marianna Anichini
e-mail: marianna.anichini@mpi.nl
Andrea Ravignani
e-mail: andrea.ravignani@mpi.nl

ORCID ID: MA, 0000-0002-2697-0297; KdR, 0000-0002-6129-9726; TAH, 0000-0002-7891-596X; DV, 0000-0001-6086-5135; ASC, 0000-0002-3722-5436; CB, 0000-0003-4918-5846; PK, 0000-0001-7579-6515; AR, 0000-0002-1058-0024

Abstract

Rhythmic patterns in interactive contexts characterise human behaviours such as conversational turn-taking. These timed patterns are also present in other animals, and often described as rhythm. Understanding fine-grained temporal adjustments in interaction requires complementary quantitative methodologies. Here, we showcase how vocal interactive rhythmicity in a non-human animal can be quantified using a multi-method approach. We record vocal interactions in harbour seal pups (*Phoca vitulina*) under controlled conditions. We analyse these data by combining analytical approaches, namely categorical rhythm analysis, circular statistics, and time series analyses. We test whether pups' vocal rhythmicity varies across behavioural contexts depending on the absence or presence of a calling partner. Four research questions illustrate which analytical approaches are complementary vs. orthogonal. For our data, circular statistics and categorical rhythms suggest that a calling partner affects a pup's call timing. Granger causality suggests that pups predictively adjust their call timing when interacting with a real partner. Lastly, the ADaptation and Anticipation Model (ADAM) estimates statistical parameters for a potential mechanism of temporal adaptation and anticipation. Our analytical complementary approach constitutes a proof of concept; it shows feasibility in applying typically unrelated techniques to seals to quantify vocal rhythmic interactivity across behavioural contexts.

Keywords: Behavioural interaction, asynchrony, circular statistics, categorical rhythms, time series analysis, interactive vocal rhythm

1. Introduction & methodological approaches

1.1. Rhythmic interaction and turn-taking

Humans adopt precise signalling behaviours to exchange information [1,2]. No matter the signal modality (e.g., acoustic, visual), an interactive event between sender and receiver is governed by a timed structure [3–6]. The structured exchange of communicative turns (i.e., turn taking) characterises our capacity for social interaction, enabling us to communicate with others [7,8]. The study of interactive rhythms—how two (or more) individuals coordinate their signalling in time—is an emerging field of research, with more recent work extending structural analyses of communication signals, once restricted to human spoken conversation, to other species [3,9]. Turn-taking in communication has been documented in primates [10] and in other non-primate species [6,11–13]. For example, call exchanges in adult common marmoset monkeys (*Callithrix jacchus*) exhibit coupled oscillator dynamics, like those observed in human turn-taking [14]. Two key components of turn-taking are the flexible organisation and distribution of turns, and the temporal relationship between adjacent turns [11]. In humans, face-to-face interactions require participants to be, among others, both socially and temporally sensitive [15]. Does behavioural context also affect signal timing in non-human animals? Comparative investigations on interactive vocal timing in mammals may help uncover shared turn-taking mechanisms, potentially providing more insights into their evolution.

1. 2. Methodological approaches and challenges for studying rhythmic interaction

Expanding the human turn-taking framework to other species is currently hindered by, among other things, lack of suitable methodological approaches [11,16]. Cross-species frameworks exist [17], but open questions still remain in animal face-to-face interaction, such as: Which analytical methodologies used to investigate turn-taking in humans may reveal temporal adjustments in other species? Can turn-taking arise from non-cooperative behavioural interactions? Can methods developed for individual rhythm analyses be used to study rhythmic interaction? Can parametric models for human rhythmic prediction and reaction detect similar features in other species? These open questions require suitable animal models and quantitative methods.

Different forms of vocal rhythmic interaction, such as synchronous chorusing and turn-taking, have been mainly studied in mammals within a cooperative dynamic, like parent-infant and male-female dyads. Time series analyses like Granger causality have shown temporal interdependence between vocalisations in male-female pairs [18,19] and movements [20] of non-human primates. Circular statistics is another method to study timing adjustments in interactions, and has been used in previous animal work, including a seal pup playback experiment [21] and a study on parent-infant monkey interaction [22]. Categorical rhythms—those for which the temporal intervals between signal onsets are distributed categorically rather than uniformly—are a universal characteristic of human

music, which is often produced in an interactive context [23]. Similar rhythms are also present in non-human animal songs (e.g., thrush nightingales [23], indris [24], but whether such rhythms characterise other interactive non-song vocalisations, such as animal calls, is unknown [23,24]. Roeske and colleagues [23] hypothesised that categorical rhythms play a role in calls produced to attract and hold conspecific attention, by making sequences of vocalisations more predictable to listeners. Categorical rhythm analyses could therefore be an interesting method to test the predictability of vocal sequences in non-human animal interactions. Lastly, the ADaptation and Anticipation Model (ADAM), originally developed to model the mechanisms for interpersonal coordination in humans [25], has been adopted to probe sensorimotor and cognitive mechanisms underlying temporal dynamics in interaction [26,27]. Although ADAM is designed for 'simultaneous chorusing', it could also be used for a mixture of bouts of synchrony, turn-taking, and other regimes [21,28].

In this proof of concept study, we showcase how these methodological tools— Granger causality, circular statistics, categorical rhythm analysis, and ADAM—can be applied to a new animal model: the harbour seal (*Phoca vitulina*). To illustrate the utility and compatibility of these different analytical methods, we show how they can be used to better understand the rhythmic communication of a small sample of harbour seal pups in different behavioural contexts.

1.3. Our animal model

The "vocal learning-beat perception and synchronisation" (VL-BPS) hypothesis states that only vocal learning species—those capable of producing new vocalisations or modifying existing ones based on auditory experience—may possess advanced rhythmic abilities [29,30]. This hypothesis is inherently cross-modal: it suggests a strong link between audition and timed movement. For example, Snowball, a sulphur-crested cockatoo (*Cacatua galerita eleonora*), was shown to perceive auditory rhythms at different tempi and to predictively synchronise his body movements to them [31]. Parrots are phylogenetically distant from humans and, among mammals, pinnipeds (seals, sea lions, and walruses) are one of the vocal learning groups (besides humans, bats, elephants, and cetaceans). Pinnipeds may well be the best mammalian model for testing the VL-BPS hypothesis—the ability to extract a beat from periodic acoustic stimuli and entrain to it in a predictive and adaptive manner—since some species showed vocal mimicry and plasticity [32,33] and others can keep a beat [34]. These characteristics, paralleling human abilities, make pinnipeds an ideal animal clade for comparative research on the origins of rhythmic communicative behaviour.

Harbour seals exhibit both vocal flexibility [33,35] and rhythmic interactivity [21], and are particularly vocal in the first few weeks of life [36]. During the lactation period, harbour seal pups emit 'mother attraction calls' (hereafter 'calls') to draw their mothers' attention [37]. Mothers are silent and use the individual vocal signatures in these calls to recognize their pups [36,38]. Against

the acoustically complex backdrop of large mother-pup rookeries, rhythmically tuned pup calls could constitute a socio-ecologically selected trait that allows individual pups to avoid conspecific call overlap by adjusting the timing of their own call onsets. Such timing plasticity could allow a pup to be more acoustically conspicuous and increase its chances of successful reunions with its mother. Unlike cooperative types of turn-taking (e.g., in humans and in common marmosets (*Callithrix jacchus*) [39]) harbour seal pups' interactions are a by-product of neighbouring pups vocalising to attract their silent mothers, and are thus probably competitive.

To date, only two papers studied vocal rhythms in harbour seals, crucially both focusing on single individuals [21,28]. The first study was a playback experiment in which a pup vocally interacted with sounds broadcasted from a loudspeaker [21]. The pup adjusted the timing of its calls in an asynchronous manner by responding to the broadcasted conspecific calls with a non-uniformly distributed response phase whose mean approximated 90º [21]. The second study looked at the presence and development of vocal rhythms in three harbour seal pups [28]. Complementary analytical approaches showed how the pups' *individual* calling patterns gained more rhythmic structure over time [28]. However, a major limitation of both studies was the lack of sociality (i.e., individuals were tested alone) and, by extension, interactivity (i.e., the stimuli did not adapt to the response of the tested animals).

1.4. Aims & research questions

In this work, we show how vocal interactive rhythmicity in non-human animals can be quantified using a multi-method approach spanning various research domains (e.g., temporal, social, cognitive) (Table 1). We illustrate this approach through four research questions, all of which relate back to whether harbour seal vocal interactive rhythmicity varies in different behavioural contexts (Table 1). While our sample sizes are too small to enable species-wide inferences, they are sufficient to illustrate how methods typically used to study human communication can be adopted to study interactivity in animal communication. The goal of this paper is thus to outline a quantitative roadmap that future research can follow. Circular statistics and categorical rhythm analysis are used to address the first question about temporal adjustment in interaction: *"does the presence of a calling partner affect the call timing of individual pups?"* (Q1). The next two questions consider the effect of behavioural context on temporal adjustment: "*does the type of calling partner (real or broadcasted) affect the call timing of individual pups?*" (Q2) and "*when the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing?*" (Q3). We answer these questions using circular statistics (Q2 and Q3) and Granger causality tests (Q2). Lastly, ADAM is used to investigate the fourth question about the cognitive processes involved in temporal adjustment: *"which timing mechanisms are used by pups during vocal interactions?"* (Q4).

1.5. Subjects, housing conditions, and behavioural contexts

We recorded nine wild-born pups (A-I) calling in different behavioural contexts while housed at Sealcentre Pieterburen (the Netherlands) (Method S1/S2). During the recordings, each pup was housed in an enclosure with a swimming pool and a resting platform (Figure S1). One pup was housed alone (I) while the others were housed in pairs (A/B, C/D, E/F, G/H). Note that the enclosures were physically but not acoustically isolated from each other, meaning that pups could hear other pups in neighbouring enclosures.

We analysed focal pup vocalisations during four different behavioural contexts (Figures 1, Table S1): (i) when the focal pup was alone (pup I), (ii) when the focal pup heard a playback of conspecific calls (pup I), (iii) when the focal pup's partner was silent (pups A-H), and (iv) when the focal pup's partner was also vocalising (pups A-H). Hereafter, we refer to these conditions as: (i) alone, (ii) 1-way interaction with a broadcasted partner, (iii) silent partner, and (iv) 2-way interaction with a real partner. Notice that only some pups entered each condition and vice-versa (Table S1).

Our sample size was affected by the unpredictable arrival of animals at the Sealcentre, which varies seasonally. Only medically stable and healthy pups were included in this study. The number of daily recording sessions per pup pair varied based on veterinary staff recommendations at the rehabilitation centre (Table S1); we did not record pups with signs of disease. Data from pup I (i.e., the alone and 1-way contexts both with and without playback stimuli) have been re-analysed from previous studies [28,21, respectively] and combined with unpublished data from pups A-H (the silent partner and 2-way contexts) (Table S1). The contribution of each pup to each analysis is shown in Table S2.

1.6. Extraction of temporal variables, definition of call bout and vocal interactions

We first extracted the onsets and offsets of each pup call recorded in each behavioural context (Figure 1, Method S3). From these values, we calculated rhythmic metrics such as call duration, inter-onset intervals (IOIs), ratios of adjacent IOIs, and inter-call intervals (ICIs; i.e., silent gaps). Each IOI was obtained by subtracting the *onset* of call *n* from the *onset* of call *n+1*, while the ICI was calculated by subtracting the *offset* of call *n* from the *onset* of call *n+1* (i.e., IOI minus duration of call *n*). Calls were organised into bouts, defined here as a series of at least three subsequent calls that were separated from adjacent bouts by a period greater than 1.5 times the median ICI of the recorded individual(s) calls (Figure 1). The response phase was computed as the ratio of the 'response IOI' (i.e., time interval between the call onset of the partner and the call onset of the pup's response) and the previous IOI, multiplied by 360, resulting in a unit vector with an angle on a circle. A vocal interaction was defined as a group of three calls within the same bout, which includes two calls from the broadcasted/real

partner followed by the response of the focal individual (Figure 1). Following previous methodology [28,38], we calculated the IOI ratio, r_k, for each pair of adjacent IOIs, t_k and t_{k+1}, in a bout as:

$$r_k = \frac{t_k}{t_k + t_{k+1}}$$

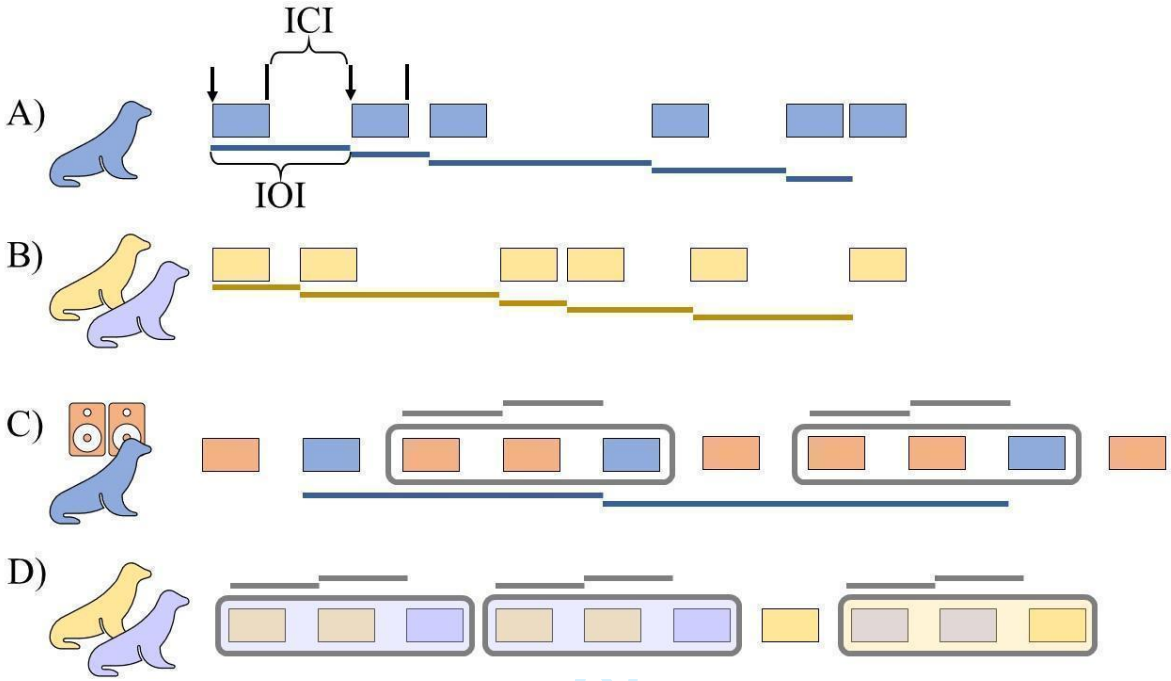

Figure 1. Schematic of experimental conditions and key measurements. Each of the four rows represents one bout and indicates a different behavioural context: (A) alone, (B) silent partner, (C) 1-way interaction, and (D) 2-way interaction. Boxes are coloured to represent the call source and grey rectangles denote vocal interactions. For bouts (C) and (D), calls are isochronously spaced for ease of visualisation, but empirical patterns were not necessarily isochronous. Horizontal bars indicate how IOIs were calculated for different analyses. The call onsets (black arrows) and offsets (black lines) are shown for the first two calls in (A).

1.7. Descriptions of analytical approaches

Circular statistics, wherein periodic measures are converted to angles on a circle and compared to distributions of interest [40], were used to investigate rhythmic periodicities in pup call response phases (Method S4). Following [21], we considered the values of the response phases as circular data falling between 0° and 360°. We obtained the circular mean (μ) (i.e., the average direction of the response phases calculated from the pup calls; Table S3). Then, we ran Rayleigh z-tests to investigate whether the distribution of response phases was uniform (e.g., arousal hypothesis) or showed a unimodal peak (Table S4) [21]. Subsequently, we tested for uniformity against a specified mean direction for the unimodal peak using a V-test [40–43].

Our data met the assumptions for circular statistics. We tested if the response phases in all four behavioural contexts followed a von Mises distribution using one-sample Watson tests (Table S5). With deviations from uniformity (null hypothesis von Mises distribution rejected), we used Kuiper's

test, Watson's test, and Rao's spacing test, to confirm the p-value obtained from the Rayleigh test (Table S6) as suggested by Landler et al. [40]. More details are shown in Method S4.

We then tested whether response phase distributions varied depending on the presence of the calling partner. We expected that vocally interacting pups would adjust their responses to broadcasted (1-way interaction) or real (2-way interaction) conspecific calls to avoid overlap and, hence, their response phases would show a unimodal distribution. Following previous work [21] and applying the V-test, we tested the null hypothesis of call response phase uniformity against two alternative unimodal departures: 0° (i.e., synchrony) and 90° (i.e., asynchrony). Using Watson's two-sample U^2 test [44], we also compared the call phase distributions of 1) a pup calling alone vs. when responding to a broadcasted partner (pup I) and 2) a pup calling in the presence of a silent partner vs. when their partner was also calling (pups A-H). For interacting (1-way or 2-way) pups, we applied Watson's two-sample U^2 test to assess whether the type of partner (i.e., real or broadcasted) differentially affected the pups' response timing. We predicted that pups interacting with a real partner would show more adaptive call timing, thanks to potential communicative cues from other modalities. Lastly, we compared the distributions of call phases of the single pup vocalising alone to those of the paired pups when their partner was silent to test whether the simple presence of a silent partner affects individual call timing. In both behavioural contexts, we predicted that calling patterns for pups without a responsive partner would show a different rhythmic structure to those observed in interaction.

Categorical rhythm analysis tests whether the temporal intervals between signal onsets, as inferred from IOI ratios, are distributed categorically rather than uniformly. We predicted that empirical and simulated null ratio distributions (i.e., the expected distribution if no rhythmic categories exist) will not differ when a pup is alone or with a silent partner but will differ when a pup is vocally interacting. For vocally interacting pups, we predicted a significant peak in empirical ratio distributions at the 4:1 small integer ratio based on the lone seal in [21], which called at approximately one-quarter of the playback's period. All categorical rhythm analyses were done within bouts following previous methodology (section Method S5) [23,24], with IOIs calculated in various ways depending on the behavioural context (Table S8, Figure 1). We used one-sample Kolmogorov-Smirnov (KS) tests to determine whether empirical IOI ratio distributions significantly differed from simulated null IOI ratio distributions. Our data met the one-sample KS test assumptions, namely that the sample is random and the theoretical distribution is continuous and fully defined. When the empirical and simulated distributions were significantly different, we also looked for evidence of small integer ratio (SIR) categorical rhythms—specifically at the 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, and 4:1 ratios—which have been found in other species' vocalisations [23,24]. In these analyses, the empirical ratio distributions were divided into "on-integer" and "off-integer" ratio bins (Table S7). On- and off-integer bin counts for each SIR were normalised by bin size and compared using a paired Wilcoxon

signed-rank test (a non-parametric test that allows for non-normality in the population data and assumes paired differences are continuous, symmetrically distributed, and mutually independent). When sample sizes allowed, we used two-sample KS tests (having met the assumption of mutual independence of measurements within samples) to determine whether the ratio distributions of individual pups differed across behavioural contexts.

The **Granger causality** test investigates whether the values of a time series A are better predicted when considering the values from a second time series B, as opposed to only using values from time series A [45]. Here, we assessed whether the call timing of a pup partaking in a 1-way or 2-way vocal interaction can be predicted using the call timing of its partner. More specifically, to investigate whether the call timing of a pup differed in relation to the type of partner, we tested whether there is a difference in predicting the time series of the pup interacting with a broadcasted partner vs. time series of the pups interacting with a real partner. Previous work showed that individuals respond to conspecific calls with a non-random pattern [18–20,46]. We therefore expected that the time series of a pup can be better predicted considering the time series of a vocalising partner rather than those of a broadcasted signal. In both the 1-way and 2-way interactive scenarios, we considered Granger causality at two levels: (1) the entire recording, regardless of the length of the pauses between consecutive calls, and (2) different bouts within each recorded session. We restricted the analysis on the different bouts to call sequences that were long enough to generate accurate estimates (i.e., a minimum of five paired calls [47]). The bouts included in this analysis range from 5 to 20 calls. We conducted the Granger causality test using call onsets and different *lag* measures, from one to five (Method S6), testing whether the previous one to five onsets in the first time series can be used to better predict the second time series (Figures S5/S6, Table S11). For the 1-way interaction, we performed a one-way analysis, considering whether the pup's timing could be predicted using the playback timing. For the 2-way interaction, we performed a two-way analysis to assess whether the two interacting pups influenced the timing of each other's calls.

We used **ADAM** to test for evidence of reactive error correction and predictive processes in the 1-way and 2-way interactive scenarios. ADAM consists of three computational modules that interact via internal models of 'self' and 'other' that support one's own action planning and external event prediction, respectively (Figure S2). The *adaptation module* compensates for synchronisation errors by implementing error correction processes that alter the phase and/or period of an internal timekeeper controlling for action (here, call) timing. These error correction processes determine the provisional timing of the next planned action by providing input to an internal model of the 'self'. The *anticipation module* computes the expected timing of upcoming events based on the weighted sum of two processes: the linear extrapolation of previous IOIs in the sequence and the copying (or 'tracking') of the previous IOI, with the output informing temporal predictions generated by the

'other' internal model. Finally, a *joint module* integrates and compares the output of the adaptation and anticipation modules and compensates for discrepancies by implementing an anticipatory error correction process before the next motor command is issued. The joint module thus reduces potential temporal mismatches between action plans in 'self' internal models and temporal predictions in 'other' internal models, thereby regulating the balance between the integration (merging) and segregation (distinction) of information about 'self' and 'other' [48,49]. Each process instantiated in ADAM is controlled by an independent parameter, and the value of these parameters can be estimated for a particular individual by fitting the model to behavioural time series data [26,27,50,51]. Parameter estimates were obtained for both the adaptation-only version of ADAM—which includes phase correction and period correction—and the full ('joint') version—including period correction, temporal prediction/tracking, and anticipatory error correction. Both versions of ADAM were applied to each interactive context because it is not possible to know a priori whether the pups' call sequences (real or broadcasted) have a steady base tempo (for which adaptation is sufficient) or a systematically changing tempo (which benefits from both anticipation and adaptation) (Method S7).

Table 1. Summary table showing in order: research questions, analyses, contributing pups, predictions, whether the data supports each prediction, statistical test(s) used, and result(s) obtained. The column "Supported by data?" has three possible answers: results support the prediction (Y), results only partially support the prediction (Partial), and results do not support the prediction (N). Due to sample size and/or analytical requirements, not all pup data could be used in each analysis. The acoustic variables of interest for the different analytical approaches were: response phases (circular statistics), IOI ratios (categorical rhythms), IOIs (Granger causality, ADAM), and asynchronies (ADAM).

Research question	Analytical approaches	Pups	Prediction	Supported by data?	Statistical test(s) and result(s)
Temporal domain, Q1: *Does the presence of a calling partner affect the call timing of individual pups?*	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
			Pups will call in asynchrony to avoid overlap during vocal interactions.	Y	V-test: Pup calls start at one-quarter of the partner's period
			Response phases will be affected by the presence of a vocalising partner.	Y	Watson's U^2 test: Response phase distributions differ between non-interactive and interactive contexts.

	Categorical rhythms	A, B, C, E, H, I	Empirical and chance ratio distributions will only significantly differ when pups are vocally interacting.	Partial	1-sample KS tests: Simulated and empirical ratio distributions are rarely significantly different (exceptions: pup I alone, pup I 1-way, pup B 2-way).
		A, B, I	Vocally interacting pups will have a significant peak at the 4:1 SIR.	N	Paired Wilcoxon signed-rank test: No significant peaks at any of the tested small integer ratios.
			Ratio distributions of individual pups will differ across behavioural contexts.	Y	2-sample KS tests: Ratio distributions significantly differed for the same individuals in different behavioural contexts.
Social domain, Q2: *Does the type of calling partner (real or broadcasted) affect the call timing of individual pups?*	Circular statistics	A-I	Pups interacting with a real partner will show more adaptive call timing than the pup interacting with a broadcasted partner.	N	Watson's U^2 test: No difference in response phases between 1-way and 2-way interactions.
	Granger causality	A, B, C, D, E, F, I	The time series of a pup will be better predicted considering the time series of a vocalising partner rather than the time series of a broadcasted signal.	Y	Bidirectional & unidirectional causality: Interaction with a real partner impacted the pup's vocal behaviour more than the playback. Mutual temporal adaptation among pairs of vocally interacting pups.
Social domain, Q3: *When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing?*	Circular statistics	A-I	The calling pattern of pups will show similar rhythmic structure in both the alone and silent partner conditions, as no vocal interaction is taking place in both cases.	Y	Watson's U^2 test: No difference between response phase distributions of pup calling alone and pups calling with a silent partner.

Cognitive domain, Q4: *Which timing mechanisms are used by pups during vocal interactions?*	ADAM	A, B, I	In the 1-way interaction, pups may show sensitivity to (non-)interactivity which would be reflected by parameter changes over repeated sessions. The 2-way interaction may be mediated by basic temporal adaptation and possibly higher-level anticipatory timing.	Partial	1-way interaction: Temporal anticipation, and to a lesser extent also adaptation, decreased across the playback sessions, and was absent in the final session. 2-way interaction: Clear evidence for temporal adaptation, with differing parameter estimates for each seal pup suggesting the emergence of different interactive roles.

3. Results

3.1. Does the presence of a calling partner affect the call timing of individual pups? (Q1)

All pups' data entered the circular statistics analysis. Running the Rayleigh test, we found that the response phase distribution was uniform for pup I which was recorded alone ($z = 0.04$, $p = 0.254$; Figure 2A), whereas it was non-uniform for pups A-H which were recorded with a silent partner ($z = 0.11$, $p < 0.001$; Figure 2A; Table 1). This non-uniformity may have been driven by the individual contributions of pups B and C, which had non-uniformly distributed response phases (Table S4), whereas the other six pups had a uniform distribution. The Rayleigh tests run anew in the interactive contexts, showed that the response phase distributions of pup calls were unimodal in both the 1-way (pup I: $z = 0.39$, $p < 0.001$) and 2-way (pups A - H: $z = 0.41$, $p < 0.001$) interactions (Figure 2B; Table 1). Applying the V-test in both contexts, the direction of the response phases did not statistically match 0º (1-way: $z = -0.02$, $p = 0.587$; 2-way: $z = 0.06$, $p = 0.110$), suggesting that pups did not synchronise with their partner (real or broadcasted). However, the response phase direction did match 90º (1-way: $z = 0.38$, 2-way: $z = 0.41$, $p < 0.001$), supporting the previously reported evidence of asynchronous calling behaviour [21]. A Watson's two-sample U^2 test confirmed that the response phase distributions significantly differed between the alone vs. 1-way interaction context for pup I (U^2 1.76, $p < 0.001$), and between the silent partner vs. 2-way interaction context for pups A-H ($U^2 = 0.78$, $p < 0.001$, Figure S3; Table 1). Finally, the circular standard deviation values were higher for the alone and silent partner contexts compared to both interactive vocal contexts, indicating a larger dispersion of the response phases for the former conditions. This outcome is also confirmed by the values for the mean resultant length.

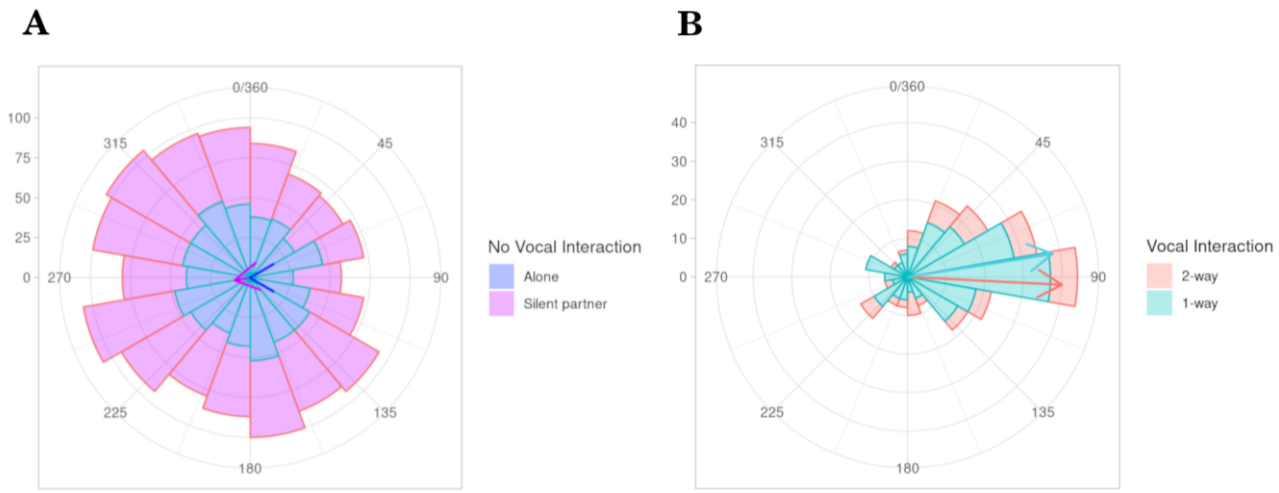

Figure 2. Circular histogram plots (bin width = 20˚) showing response phases in (A) both types of behavioural contexts with no vocal interaction (alone/silent partner), and (B) in both types of vocal interaction contexts (1-way/2-way). Angles are measured in degrees starting from 0˚ and going clockwise to 360˚. The arrows indicate the circular mean (μ) and colours correspond to the different behavioural contexts. The length of the arrow corresponds to the value of the mean resultant length (ρ).

In the categorical rhythm analyses, the empirical ratio distribution was significantly different from chance when pup I was recorded alone (Figure 3A) and during the 1-way interaction (Figure 3C). In both contexts, there was no evidence of significant peaks at any of the tested ratios (Table 1). When the playback calls were disregarded from IOI calculations (Figure 3B), there was no significant difference in empirical and simulated ratio distributions for pup I. However, pairwise KS tests showed that the ratio distributions significantly differed when comparing each of the three behavioural contexts (alone vs. 1-way interaction disregarding playback vs. 1-way interaction when pup I responds) to each other (Table S10; Table 1). For the five pups that were well-sampled in the silent partner context (pups A, B, C, E, and H; Table S8), the empirical ratio distributions did not significantly differ from chance (Table S9, Figures 3/S4, Method S4). Finally, when considering 2-way interactions, only pups A and B were well-sampled enough (i.e., had at least ten ratios for both the silent partner and 2-way interaction contexts) to compare, but the empirical and simulated ratio distributions were significantly different only when pup B was the responder (Figure 3; Table S8; Table S9). Once again, there were no significant peaks at any of the tested ratios for pup B. For both pups A and B, the empirical ratio distributions significantly differed when comparing different behavioural contexts (Table S10). Collectively, there was thus little evidence of SIR rhythmic categories in pup calls, but IOI ratios did significantly differ when looking at the same individuals in different behavioural contexts (Table 1).

Thus, the results from both analyses suggest that the *presence* of a calling partner does affect the call timing of the focal pup, in terms of both call response phases and IOI ratios.

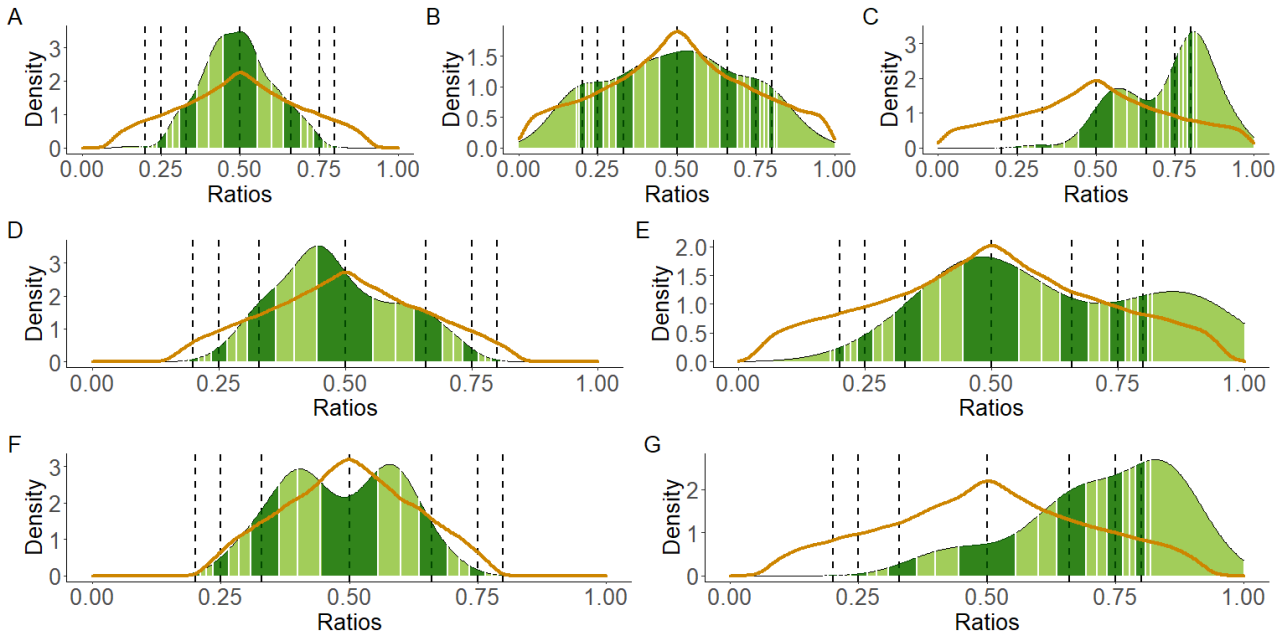

Figure 3. IOI ratio density plots for pups in different behavioural contexts. Pup I: (A) alone, (B) 1-way interaction (disregarding playback), (C) 1-way interaction (responding to playback). Pup A: (D) silent partner, (E) 2-way interaction (responding to partner). Pup B: (F) silent partner, (G) 2-way interaction (responding to partner). For each plot, the dashed vertical lines indicate, from left to right, 1:4, 1:3, 1:2, 1:1 (i.e., isochrony), 2:1, 3:1, and 4:1 small integer ratios. On-integer ratio ranges are in dark green and off-integer ratio ranges are in light green, with white lines and black dashed lines denoting bin boundaries. The orange curves indicate the ratio distribution expected under a uniform distribution if no rhythm categories exist. The empirical ratio distribution significantly differed from the simulated ratio distribution for panels (A), (C), and (G) only. Note that the scale of the y-axes differ.

3.2. Does the type of calling partner (real or broadcasted) affect the call timing of individual pups? (Q2)

To address this research question, we once again used circular statistics and the response phase distributions of all nine pups. Specifically, we compared calls from pup I during the playback (1-way interaction) with calls of pups A-H when their partner was also calling (2-way interaction). A Watson's two-sample U^2 test statistically confirmed that the response phase distributions did not differ between the 1-way and 2-way vocal interactions ($U^2 = 0.07$, $p > 0.10$; Table 1).

Interestingly, however, the Granger causality results from seven pups (A, B, C, D, E, F, I) showed that call timing behaviour differed depending on the type of partner (Table 1). For the 1-way interaction, five different playback sessions featuring pup I were considered (ranging from 34 to 121 paired calls) and the timing of the pup's calls were never significantly predicted by the timing of the playback (Table S11, Figure S5). For the 2-way interaction, five different recording sessions were considered (ranging from 8 to 71 paired onsets). We found that in two sessions, the timing of the first pup did not significantly predict the timing of the second, in either direction (pair A>B and B>A; pair E>F and F>E; Table S11, Figure S6B/S6D). However, in two other sessions, the timing of a pup was significantly predicted by the calling partner in both directions, across different lag values (A>B: lag-

1, lag-3, lag-4, lag-5; B>A: lag-3, lag-4, lag-5; C>D and D>C: lag-2 and lag-3; Table S11, Figure S6B/S6C). In the last session, the timing of pup A could be predicted by that of pup B, while we found no indication of temporal adjustment for pup B. At the bout level, we restricted our analyses to four series of paired onsets featuring pups A and B (as they were the only pup pair that met the sample size requirements). In this last scenario, only the timing of pup B was significantly influenced by pup A (Figure S6A).

While the circular statistics results thus suggest that the type of calling partner (real vs. broadcasted) does not affect focal pup call timing, the Granger causality results suggest that in certain 2-way (but not 1-way) interactions, focal pup call timing can be predicted by the partner's call timing.

3.3. When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing? (Q3)

Using circular statistics, we compared the calls of pup I recorded alone with those of the 8 other pups (A-H) recorded with a silent partner present. Watson's two-sample U^2 test results show that the response phase distributions did not statistically differ between the two contexts ($U^2 = 0.10$, $p > 0.05$). In other words, having a silent partner was essentially the same as having no partner—in terms of the effect on focal pup response phase distributions—for the pups in our study. This is intuitive, given that in both behavioural contexts, there is no acoustic stimuli to "respond" to.

3.4. Which timing mechanisms are used by pups during vocal interactions? (Q4)

ADAM parameter estimation was conducted on call data from all five playback sessions featuring pup I (1-way, Figure 4 A/B), but was restricted to pups A and B for the 2-way interactions due to sample size limitations (Figure 4 C/D). Interactive vocal bouts were concatenated to obtain a time series length which would provide reliable ADAM parameter estimates. A simulation test then ensured that the estimates were not compromised by differing sequence lengths or by the concatenation procedure (Method S7). The reliability of observed parameter estimates was tested by comparison against corresponding values for randomly permuted data (Method S7). The quality of the fits to the data did not differ significantly between versions of ADAM (Method S7). Results for the 1-way interaction featuring pup I were remarkable with regards to typical human data (e.g. [26,27]) as most significant parameters were negative in sign (Figure 4A/B, Tables S13/S14). Negative phase and period correction estimates indicate that calling earlier will lead to a shortening of the next IOI, while calling later will lead to a lengthening of the next IOI. Negative prediction/tracking estimates mean that when the playback's IOIs increase (i.e., deceleration), the pup's IOIs will decrease (i.e., acceleration), and vice versa. This systematically enhances the timing distinction between calls, possibly testing the responsiveness of the partner (i.e., playback) by

introducing timing asynchronies and gauging their effects. It is worth noting that evidence of such behaviour generally decreased across the five playback sessions, with no significant parameter estimates emerging in the final session.

For the 2-way interaction (Figure 4C/D, Tables S13/S14), the parameter estimation procedure was run twice, each time with one of the two pups serving as the focal individual and the other as the external reference. With this procedure, similar parameter estimates for each pup would indicate a symmetrical pattern of influence, whereas different estimates would indicate asymmetrical influence. Results for this type of vocal interaction showed evidence for temporal adaptation. In particular, parameter estimates were consistent with pup A engaging in period correction while pup B engaged in both phase and period correction. Period correction estimates were positive in sign for pup A and negative for pup B. Pup A's positive period correction estimates suggest that calling earlier will lead to a lengthening of the next IOI, whereas calling later will lead to a shortening of the next interval. Pup B exhibited the opposite effect but to a lesser degree. Pup A thus implemented a timing mechanism that increased the similarity in their calling rates, while pup B implemented a timing mechanism that increased the distinction in calling rates. It should be noted, however, that the observed parameter estimates do not necessarily reflect individual call styles per se, but rather the roles that dynamically emerge within the context of this specific pairing of pups.

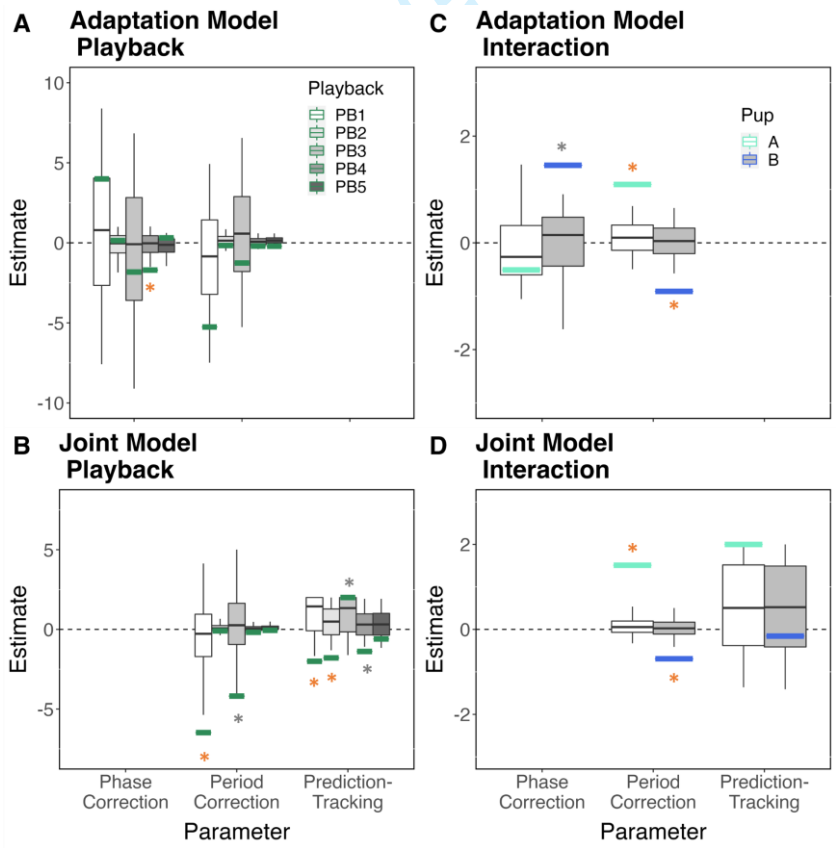

Figure 4. ADAM parameter estimates for seal pups A, B, and I. Left panels (A & B) show parameter estimates for pup I in five playback sessions (1-way interaction); right panels (C & D) show estimates for pups A and B (2-way interaction). Top panels (A & C) show phase correction and period estimates obtained with the adaptation-only version of ADAM; bottom panels (B & D) show period correction and prediction-tracking estimates for the full 'joint' model (anticipatory error correction estimates are not shown). Parameter estimates are shown separately for pup A relative to pup B (aquamarine bars) and pup B relative to pup A (blue bars) in the recording session, and for pup I (green bars) relative to playback calls in separate playback sessions (PB1-PB5). Corresponding parameter estimates for randomly permuted data from each session are represented by box plots, with the central horizontal line indicating the median, the bottom and top edges of the box indicating the 25th and 75th percentiles, and the whiskers indicating the 5th and 95th percentiles. Real estimates with asterisks are significantly different from permuted data estimates at the 2-tailed (orange) or 1-tailed (grey) level.

4. Discussion

This study sought to provide a methodological proof of concept for quantifying vocal rhythmicity in non-human animal interactive communication. Particularly, we showed how complementary methodologies can be used to test whether the (Q1) presence and (Q2) type of a vocalising partner, or the (Q3) presence of a silent partner affect patterns in animal communication (Table 1). We also investigated which (Q4) underlying cognitive mechanisms potentially play a role in vocal interactions (Table 1). The different analytical techniques proved fruitful; combining and contrasting their results could highlight nuances in rhythmic behaviour. Nonetheless, our sample size is undeniably small, and our opportunistic design (e.g., mixing within- and between-individual comparisons) cannot provide strong inference. We believe that our experimental setup, combined with the approaches we present, can be adapted and expanded in future work to better understand the temporal, social, and cognitive processes underlying interactive communication in animals.

Q1. The presence of a vocalising partner influences rhythm in vocal interactions

The prediction of overlap avoidance was supported by circular statistics, which showed that the distributions of response phase angles were unimodal. When vocally interacting, pups responded at approximately one-quarter of the playback/real partner calling period (90◦) and showed phase angles significantly different from 0◦, supporting previous results from one individual (pup I) [21]. Pups responding to conspecifics seem to time their calls to avoid overlap, consistently de-synchronizing their call onsets. When pups were recorded alone or with a silent partner, the distribution of phase angles was uniform, with no significant rhythmic pattern. Our findings confirm and extend previous outcomes [21]; in our limited sample, pups not only react to playbacks with asynchrony, but also respond to conspecific partner calls in an asynchronous manner.

Regarding categorical rhythms, the empirical ratio distribution did not significantly differ from the distribution expected by chance if IOIs were uniformly distributed for most pup/behavioural context combinations (7/10). For the remaining three combinations (pup I alone; pup I, 1-way interaction; pup B, 2-way interaction), there was no significant evidence that call timing exhibited

categorical rhythms at the seven tested small integer ratios. This includes the 4:1 ratio, which we hypothesised would frequently occur when pups were vocally interacting with playback stimuli or partners based on previous work [21]. Importantly, however, the empirical ratio distributions significantly differed across behavioural context conditions for the three pups (A, B, I) for whom such comparisons were possible. When alone or with a silent partner, pup calls generally showed unimodal ratio distributions centred around isochrony, whereas calls emitted by vocally interacting pups showed a clear right-shifted peak, or a second peak to the right of isochrony, indicating that the pup's IOIs during vocal interactions were generally shorter than the playback's/partner's IOIs. This context-dependent asymmetry bolsters results related to Q4, namely that the presence of vocalising partner can significantly impact pup vocal behaviour, with interacting individuals trying to prevent call overlap. Roeske and colleagues [23] hypothesised that categorical rhythms may be prevalent in calls meant to attract and hold conspecific attention, such as the calls pups use to attract the attention of their mothers. We did not find evidence of rhythm categories at seven tested ratios; this negative result has methodological value, highlighting that not all species that produce attention-seeking vocalisations show integer ratio categories.

A vocal interaction with non-adaptive playback stimuli (1-way) may represent a limitation to studying spontaneous vocal production. The extent of this limitation can be gauged by comparing results to data from vocal interactions with a real partner (2-way), and with the use of complementary analytical methodologies, such as categorical rhythm analysis and circular statistics.

Q2. Type of calling partner (real or broadcasted) partly affects rhythm in interaction.

Findings from Q2 show contrasting results. Call phases were statistically the same no matter if pups could interactively adjust their call timing to each other (2-way) or not (in 1-way). Partly in contrast with this, the Granger causality analysis showed how vocally interacting with a real individual impacted the pup's vocal behaviour more than interacting with a broadcasted partner Indeed, we found evidence of mutual temporal adaptation among pairs of vocally interacting pups and, consistent with a recent hypothesis [6], conspecific interactions can be facilitated by the reciprocal adjustment of timing behaviour. In contrast, we found little evidence that a pup adjusts the time series of its calls to a playback series. Indeed, in most cases, the pup's call timing could not be predicted by the playback's call timing.

Together, findings from circular statistics (Q1) and Granger causality (Q2) point towards a directionality-overlap avoidance relationship, which has also been suggested for other non-human species (e.g., primates [18,52]; birds [53]; amphibians [54]; seals [55]). Interestingly, when infant marmosets interact with their parents, the probability that their vocalisations will overlap with those

of adults decreases over time [22], suggesting that turn-taking in some mammals is a learned vocal behaviour scaffolded by active parental feedback [39].

Q3. A silent partner does not trigger call rhythmicity

Circular statistics indicated that pups did not show any periodic pattern both, either when calling alone or with a silent partner present (Q1). This is consistent with the fact that wild pups produce calls to attract a silent mother [36]: if no other calling conspecifics are present, there is no need to adjust one's call timing and vocalisations are produced with a random onset. Moreover, given that the response phase distributions did not differ between pups recorded with a silent partner or alone, it suggests that the presence of a silent partner did not trigger variable calling behaviour in our study population.

Q4. Purported timing mechanisms for vocal interactions

The ADAM analysis suggests that seal pups may perceive temporal patterns [56], which arise between their calls and those of others, and react to them by adopting different mechanisms for temporal adaptation. The negative parameter estimates that we observed in the 1-way interaction sessions (consistent with enhancing the distinction between calls) could reflect attempts to lead the temporal interaction or even to test the responsiveness of the (broadcasted) partner by introducing timing perturbations and gauging their effects. The decrease of temporal adaptation and/or anticipation observed across playback sessions is consistent with a gradual process of habituation, with the pup possibly learning that the playback is non-interactive. Harbour seals are capable of acoustic recognition based on habituation paradigms; they can discriminate among different stimuli and selectively habituate to them [57]. The lack of temporal adaptation to the playback stimulus we observed in a seal pup may entail similar habituation processes. From a comparative perspective, these findings also suggest a sensitivity to social contingency that may be analogous to capacities in human infants, who become disinterested and display fewer signs of positive affect during vocal interactions with non-responsive or delayed video recordings of their mothers [58–62]. Future studies could address the role of temporal contingency by using interactive playback sequences [63–65].

Ours constitutes the first attempt to apply ADAM to non-human animals. On the technical side, this necessitated the validation of an approach where brief interactive vocal bouts were concatenated to produce longer time series and thereby reduce the risk of model overfitting (Discussion S1). Demonstrating the utility of this procedure opens the door to applying the model in a wider range of behavioural contexts. Nevertheless, caveats are necessary when interpreting the seal data in light of previous work with ADAM in humans, where individuals intentionally produce movements, whose sensory effects occur simultaneously with rhythmically regular sounds (e.g. [26,27]). Assumptions

about intentionality and simultaneity may not apply to seal pup vocal interactions or rhythmic interactions in other animals [6,9]. Points of convergence in the main outcomes of the complementary analysis techniques suggest that ADAM, like categorical rhythm analysis, circular statistics, and Granger causality, may also be robust and informative under such conditions.

5. Conclusions

Crucially, our work highlights the efficacy of combining multiple methods to study rhythmic vocal behaviour. Our approaches vary in the degree to which they capture global temporal characteristics across events vs. local temporal dependencies between events. Global measures (e.g., from circular statistics or categorical rhythm analysis) reveal predominant rhythmic features of a vocal interaction while local time series measures (e.g., from Granger causality or ADAM) provide information about how these features might arise. Through this approach we could tease apart rhythm nuances in our dataset, further develop harbour seals as a model species, and demonstrate how certain analyses often restricted to humans, such as categorical rhythms (but see [23,24]) and ADAM, can be applied to non-human animals. The categorical rhythm and circular statistical analyses tackled similar questions from different angles, namely whether the distributions of IOI ratios (the former) or call response phases (the latter) significantly differed across behavioural contexts. The categorical rhythm analyses also sought to determine whether the rhythmicity of pup calling behaviour conforms to small integer ratios. Our negative result is, to our knowledge, the first published case of a species for which categorical rhythms are clearly absent from vocalisations, which adds to understanding of how, why, and when such rhythms evolve in communication systems. Time series analyses such as Granger causality allowed testing for timing adjustment. In the case of ADAM, mechanisms of temporal adaptation and anticipation that have previously been used to describe rhythmic behaviour in humans [66] were used to describe rhythmic behaviour in seals. Interestingly, the ADAM model provides an empirical warning about potential seal pup habituation effects when vocally interacting with a recorded partner.

Though our sample size is small, studies on single individuals are not unusual in comparative research [20,67]. Nevertheless, it is possible that the lack of adjustment to a playback, the adaptation to a real individual, or both, reflect a peculiar vocal behaviour of the individuals we tested and cannot be generalised to the species as a whole. The pups in this study were in a temporarily captive setting, albeit in acoustic proximity to other individuals, similar to conditions they would experience in nature. Unfortunately, vocal development in harbour seal pups has not been studied in wild colonies, hence we do not know whether captivity affects their vocal development. We do know, however, that pups vocally interact with neighbouring pups in the colony and not with their silent mothers; hence

the turn-taking behaviour observed in our captive conditions might extend to the same behaviour in wild conspecifics.

Motivation for an individual to respond and engage in an interaction, with the closest partner in the colony, may depend on the degree of participation signalled by the partner. This, in turn, may be triggered by individual-specific behaviours or by cues from other modalities. Multimodal communication should be the target of future studies since we cannot assume that such interactions are limited to acoustic cues. Despite these limitations, our study shows that adopting multiple complementary approaches can be a fruitful way to study rhythmic interactivity in non-human animal communication.

Acknowledgements

We are grateful to the editor and the two anonymous reviewers who provided insightful comments and suggestions on an earlier version of this manuscript. We thank all the volunteers, seal care members and veterinarians of Sealcentre Pieterburen who helped during the experiment and in the rehabilitation process of the seal pups. We are thankful to Chiara De Gregorio for sharing the R code used for the categorical rhythm analyses and to Laura Verga for providing helpful comments on a draft of this manuscript.

References

1. Dusenbery DB. 1992 *Sensory ecology: How organisms acquire and respond to information.* Freeman, New York.
2. Stegmann UE. 2013 *Animal Communication Theory: Information and Influence.* Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9781139003551)
3. Anichini M, de Heer Kloots M, Ravignani A. 2020 Interactive rhythms in the wild, in the brain, and in silico. *Can. J. Exp. Psychol.* **74(3)**, 170-175. (doi: 10.1037/cep0000224)
4. Ravignani A, Bowling DL, Fitch WT. 2014 Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology.* **5**,1118. (doi:10.3389/fpsyg.2014.01118)
5. Greenfield MD, Aihara I, Amichay G, Anichini M, Nityananda V. 2021 Rhythm interaction in animal groups: selective attention in communication networks. *Phil. Trans. R. Soc. B* **376(1835)**, 20200338. (doi:10.1098/rstb.2020.0338)
6. 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)
7. Levinson SC. 2006 Cognition at the heart of human interaction. *Discourse studies*. **8(1)**, 85-93.
8. Levinson SC. 2016 *The Interaction Engine hypothesis.* The Language in Interaction Summerschool on Human Language: From Genes and Brains to Behavior. Berg en Dal, The Netherlands.
9. Ravignani A, Verga L, Greenfield MD. 2019 Interactive rhythms across species: the evolutionary biology of animal chorusing and turn-taking. *Annals of the New York Academy of Sciences,* **1453(1)**, 12-21. (doi:10.1111/nyas.14230)
10. Levinson SC. 2016 Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences*. **20(1)**, 6-14. (doi.org/10.1016/j.tics.2015.10.010)
11. Pika S, Wilkinson R, Kendrick KH, Vernes SC. 2018 Taking turns: bridging the gap between human and animal communication. *Proc. R. Soc. B* **285(1880)**, 20180598. (doi:10.1098/rspb.2018.0598)
12. Demartsev V, Strandburg-Peshkin A, Ruffner M, Manser M. 2018 Vocal turn-taking in meerkat group calling sessions. *Curr. Biol*. **28(22)**, 3661-3666. (doi: 10.1016/j.cub.2018.09.065)
13. Schulz TM, Whitehead H, Gero S, Rendell L. 2008 Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Anim. Behav.* **76(6)**, 1977-1988. (doi:10.1016/j.anbehav.2008.07.032)
14. Takahashi DY, Narayanan DZ, Ghazanfar AA. 2013 Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys, *Curr. Biol.* (doi:10.1016/j.cub.2013.09.005)
15. Levinson SC. 2020 On the human" interaction engine". In *Roots of human sociality*, pp. 39-69. Routledge.
16. Yoshida S, Okanoya K. 2005 Evolution of turn-taking: a bio-cognitive perspective. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, **12(3)**, 153-165. (doi:10.11225/jcss.12.153)
17. Henry MJ, Cook PF, de Reus K, Nityananda V, Rouse AA, Kotz SA. 2021 An ecological approach to measuring synchronization abilities across the animal kingdom. *Phil. Trans. R. Soc. B*. **376(1835)**, 20200336. (doi:10.1098/rstb.2020.0336)
18. 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)
19. Clink DJ, Tasirin JS, Klinck H. 2020 Vocal individuality and rhythm in male and female duet contributions of a nonhuman primate. *Curr. Zool.* **66(2)**, 173–186. (doi:10.1093/cz/zoz035)
20. Fuhrmann D, Ravignani A, Marshall-Pescini S, Whiten A. 2014 Synchrony and motor mimicking in chimpanzee observational learning. *Sci. Rep.* **4**, 1–7. (doi:10.1038/srep05283)
21. Ravignani A. 2019 Timing of antisynchronous calling: A case study in a harbour seal pup (Phoca vitulina). *J Comp Psychol.* **133(2)**, 272–277. (doi:10.1037/com0000160)

22. Takahashi DY, Fenley AR, Ghazanfar AA. 2016 Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Phil. Trans. R. Soc. B: Biological Sciences*. **371(1693)**, 20150370. (doi:10.1098/rstb.2015.0370)

23. Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical rhythms are shared between songbirds and humans. *Curr. Biol.* **30**, 3544–3555. (doi:10.1016/j.cub.2020.06.072)

24. De Gregorio C, Valente D, Raimondi T, Torti V, Miaretsoa 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)

25. van der Steen MC, Keller PE. 2013 The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience,* **7**, 253. (doi:10.3389/fnhum.2013.00253)

26. van der Steen MC, Jacoby N, Fairhurst MT, Keller PE. 2015 Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Res*. **1626**, 66–87. (doi:10.1016/j.brainres.2015.01.053)

27. Mills PF, Harry B, Stevens CJ, Knoblich G, Keller PE. 2019 Intentionality of a co-actor influences sensorimotor synchronisation with a virtual partner. *Quarterly Journal of Experimental Psychology,* **72(6)**, 1478–1492. (doi:10.1177/1747021818796183)

28. Ravignani A, Kello CT, de Reus K, Kotz SA, Dalla Bella S, Méndez-Aróstegui M, Rapado-Tamarit B, Rubio-Garcia A, de Boer B. 2019 Ontogeny of vocal rhythms in harbour seal pups: an exploratory study. *Curr Zool.* **65(1)**, 107–120. (doi: 10.1093/cz/zoy055)

29. Patel AD. 2006 *Musical rhythm, linguistic rhythm, and human evolution.* Music Percept **24**,99–104.

30. Patel AD. 2014 The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol* **12**, 1001821. (doi:10.1371/journal.pbio.1001821)

31. Keehn RJJ, Iversen JR, Schulz I, Patel AD. 2019 Spontaneity and diversity of movement to music are not uniquely human. *Curr. Biol.* **29(13)**, R621-R622. (doi: 10.1016/j.cub.2019.05.035)

32. Schusterman RJ. 1977 Temporal patterning in sea lion barking (Zalophus californianus). *Behavioral biology*. **20(3)**, 404-408. (doi:10.1016/S0091-6773(77)90964-6)

33. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*. **63(5)**, 1050-1056. (doi:10.1139/z85-157)

34. Cook P, Rouse A, Wilson M, Reichmuth C. 2013 A California sea lion (Zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* **127(4)**, 412-27. (doi:10.1037/a0032345)

35. Torres Borda L, Jadoul Y, Rasilo H, Salazar Casals A, Ravignani A. 2021 Vocal plasticity in harbour seal pups. *Phil.Trans. R. Soc. B* **376**, 20200456. (doi:10.1098/rstb.2020.0456)

36. Sauvé CC, Beauplet G, Hammill MO, Charrier I. 2015 Acoustic analysis of airborne, underwater, and amphibious mother attraction calls by wild harbor seal pups (Phoca vitulina). *Journal of Mammalogy* **96(3)**, 591–602. (doi:10.1093/jmammal/gyv064)

37. Perry EA, Renouf D. 1988 Further-studies of the role of harbor seal (Phoca vitulina) pup vocalizations in preventing separation of mother pup pairs. *Can. J. Zool*. **66**, 934–938. (doi:10.1139/z88-138)

38. Khan CB, Markowitz H, McCowan B. 2006 Vocal development in captive harbor seal pups, Phoca vitulina richardii: age, sex, and individual differences. *J. Acoust. Soc. Am.* **120(3)**, 1684-94. (doi:10.1121/1.2226530)

39. Chow CP, Mitchell JF, Miller CT. 2015 Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. R. Soc. B* **282**:20150069. (doi:10.1098/rspb.2015.0069)

40. Landler L, Ruxton GD, Malkemper EP. 2018 Circular data in biology: advice for effectively implementing statistical procedures. *Behav. Ecol. Sociobiol.* **72(8)**, 1–10. (doi:10.1007/s00265-018-2538-y)

41. Ruxton GD. 2017 Testing for departure from uniformity and estimating mean direction for circular data. *Biol. Lett.* **13**, 20160756. (doi:10.1098/rsbl.2016.0756)

42. Pewsey A, Neuhäuser M, Ruxton GD. 2013 *Circular statistics in R*. Oxford University Press.

43. Zar JH. 1999 Circular distributions: hypothesis testing. *Biostatistical analysis*, 624–668.
44. Landler L, Ruxton GD, Malkemper EP. 2021 Advice on comparing two independent samples of circular data in biology. *Scientific reports*. **11(1)**, 1-10. (doi:10.1038/s41598-021-99299-5)
45. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica.* **37**, 424–438. (doi: 10.2307/1912791)
46. Ravignani A, Norton P. 2017 Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution* 4–19. (doi:10.1093/jole/lzx002)
47. Zeileis A, Hothorn T. 2002 Diagnostic Checking in Regression Relationships. *R News*, **2(3)**, 7–10. (https://CRAN.R-project.org/doc/Rnews/)
48. Keller PE, Novembre G, Hove MJ. 2014 Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369(1658)**, 20130394. (doi:10.1098/rstb.2013.0394)
49. Keller PE, Novembre G, Loehr J. 2016 Musical ensemble performance: Representing self, other and joint action outcomes. In S. S. Obhi & E. S. Cross (Eds.), *Shared representations: Sensorimotor foundations of social life*. pp. 280–310. Cambridge: Cambridge University Press. (doi:10.1017/CBO9781107279353.015)
50. Mills PF, van der Steen MC, Schultz BG, Keller PE. 2015 Individual differences in temporal anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception* **3(1-2)**, 13–31. (doi:10.1163/22134468-03002040)
51. van der Steen MC, Schwartze M, Kotz SA, Keller PE. 2015 Modeling effects of cerebellar and basal ganglia lesions on adaptation and anticipation during sensorimotor synchronization. *Annals of the New York Academy of Sciences,* **1337**, 101–110. (doi:10.1111/nyas.12628)
52. Lemasson A, Pereira H, Levréro F. 2018 Social basis of vocal interactions in western lowland gorillas (Gorilla g. gorilla). *J. Comp. Psychol.* **132(2)**, 141–151. (doi: 10.1037/com0000105)
53. Yang XJ, Ma XR, Slabbekoorn H. 2014 Timing vocal behaviour: experimental evidence for song overlap avoidance in Eurasian wrens. *Behav Processes* **103**, 84–90. (doi:10.1016/j.beproc.2013.11.011)
54. Song J, Sun R, Fang K, Zhang B, Fang G. 2020 Flexibility as a strategy for avoiding call overlap in male Anhui Treefrogs. *Asian Herpetological Research* **11(3)**, 230–239. (doi:10.16373/j.cnki.ahr.1900331)
55. Serrano A, Terhune JM. 2002 Antimasking aspects of harp seal (Pagophilus groenlandicus) underwater vocalizations. *J. Acoust. Soc. Am*. **112(6)**, 3083–3090. (doi:10.1121/1.1518987)
56. Verga L, Sroka MG, Varola M, Villanueva S, Ravignani, A. 2022 Spontaneous rhythm discrimination in a mammalian vocal learner. Biology Letters, **18(10)**, 20220316.
57. Deecke VB, Slater PJ, Ford JK. 2002 Selective habituation shapes acoustic predator recognition in harbour seals. *Nature* **6912**, 171–3. (doi:10.1038/nature01030)
58. Bigelow AE, DeCoste C. 2003 Sensitivity to Social Contingency From Mothers and Strangers in 2-, 4-, and 6-Month-Old Infants. *Infancy,* **4(1)**, 111–140. (doi:10.1207/S15327078IN0401_6)
59. Henning A, Striano T. 2011 Infant and Maternal Sensitivity to Interpersonal Timing. *Child Development,* **82(3)**, 916–931. (doi:10.1111/j.1467-8624.2010.01574.x)
60. Malloch S, Trevarthen C. 2018 The Human Nature of Music. *Frontiers in Psychology,* **9(1680)**. (doi:10.3389/fpsyg.2018.01680)
61. Soussignan R, Nadel J, Canet P, Gerardin P. 2006 Sensitivity to Social Contingency and Positive Emotion in 2-Month-Olds. *Infancy,* **10(2)**, 123–144. (doi:10.1207/s15327078in1002_2)
62. Stormark KM, Braarud HC. 2004 Infants' sensitivity to social contingency: a "double video" study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development,* **27(2)**, 195–203. (doi:10.1016/j.infbeh.2003.09.004)
63. Fairhurst MT, Janata P, Keller PE. 2013 Being and feeling in sync with an adaptive virtual partner: Brain mechanisms underlying dynamic cooperativity. *Cerebral Cortex,* **23**, 2592–2600. (doi:10.1093/cercor/bhs243)

64. Fairhurst MT, Janata P, Keller PE. 2014 Leading the follower: an fMRI investigation of dynamic cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner. *Neuroimage,* **84**, 688–697. (doi:10.1016/j.neuroimage.2013.09.027)

65. Repp BH, Keller PE. 2008 Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science,* **27**, 423–456. (doi:10.1016/j.humov.2008.02.016)

66. Harry B, Keller PE. 2019 Tutorial and simulations with ADAM: an adaptation and anticipation model of sensorimotor synchronization. *Biological Cybernetics,* **113(4)**, 397–421. (doi:10.1007/s00422-019-00798-6)

67. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. (doi:10.1016/j.cub.2009.03.038)

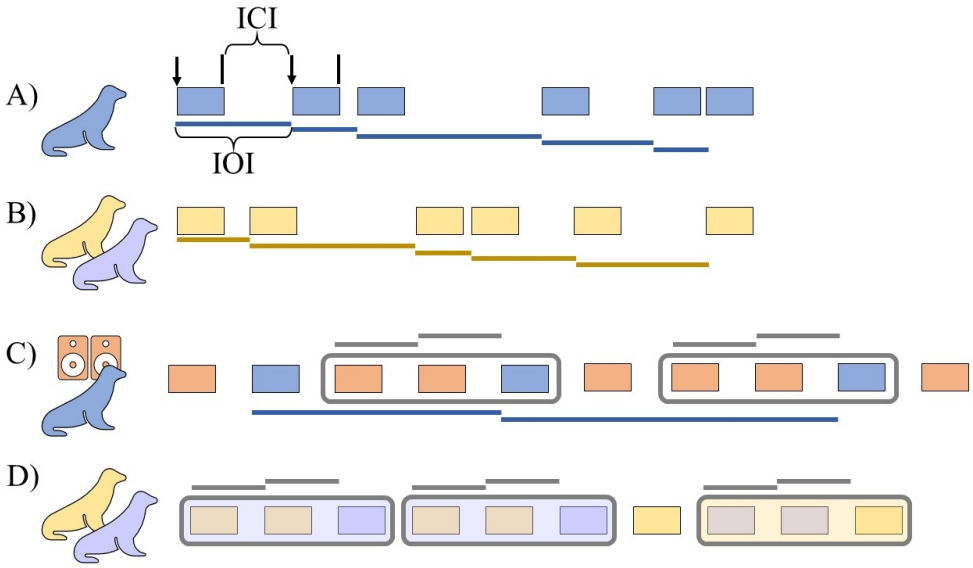

Figure 1. Schematic of experimental conditions and key measurements. Each of the four rows represents one bout and indicates a different behavioural context: (A) alone, (B) silent partner, (C) 1-way interaction, and (D) 2-way interaction. Boxes are coloured to represent the call source and grey rectangles denote vocal interactions. For bouts (C) and (D), calls are isochronously spaced for ease of visualisation, but empirical patterns were not necessarily isochronous. Horizontal bars indicate how IOIs were calculated for different analyses. The call onsets (black arrows) and offsets (black lines) are shown for the first two calls in (A).

338x190mm (96 x 96 DPI)

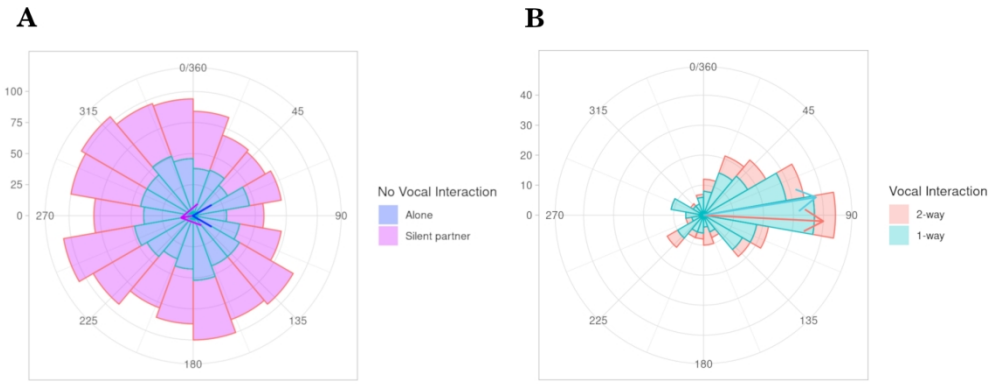

Figure 2. Circular histogram plots (bin width = 20◦) showing response phases in (A) both types of behavioural contexts with no vocal interaction (alone/silent partner), and (B) in both types of vocal interaction contexts (1-way/2-way). Angles are measured in degrees starting from 0◦ and going clockwise to 360◦. The arrows indicate the circular mean (μ) and colours correspond to the different behavioural contexts. The length of the arrow corresponds to the value of the mean resultant length (ρ).

413x232mm (118 x 118 DPI)

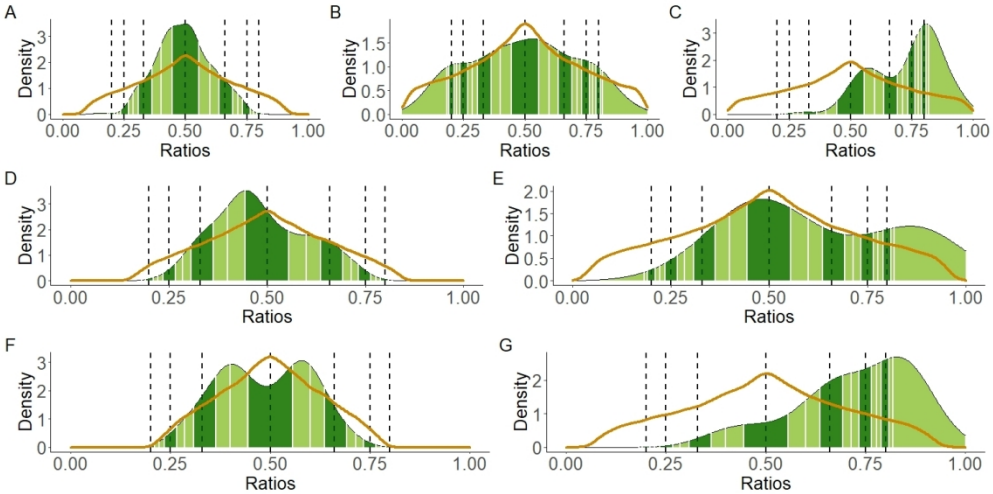

Figure 3. IOI ratio density plots for pups in different behavioural contexts. Pup I: (A) alone, (B) 1-way interaction (disregarding playback), (C) 1-way interaction (responding to playback). Pup A: (D) silent partner, (E) 2-way interaction (responding to partner). Pup B: (F) silent partner, (G) 2-way interaction (responding to partner). For each plot, the dashed vertical lines indicate, from left to right, 1:4, 1:3, 1:2, 1:1 (i.e., isochrony), 2:1, 3:1, and 4:1 small integer ratios. On-integer ratio ranges are in dark green and off-integer ratio ranges are in light green, with white lines and black dashed lines denoting bin boundaries. The orange curves indicate the ratio distribution expected under a uniform distribution if no rhythm categories exist. The empirical ratio distribution significantly differed from the simulated ratio distribution for panels (A), (C), and (G) only. Note that the scale of the y-axes differ.

541x275mm (72 x 72 DPI)

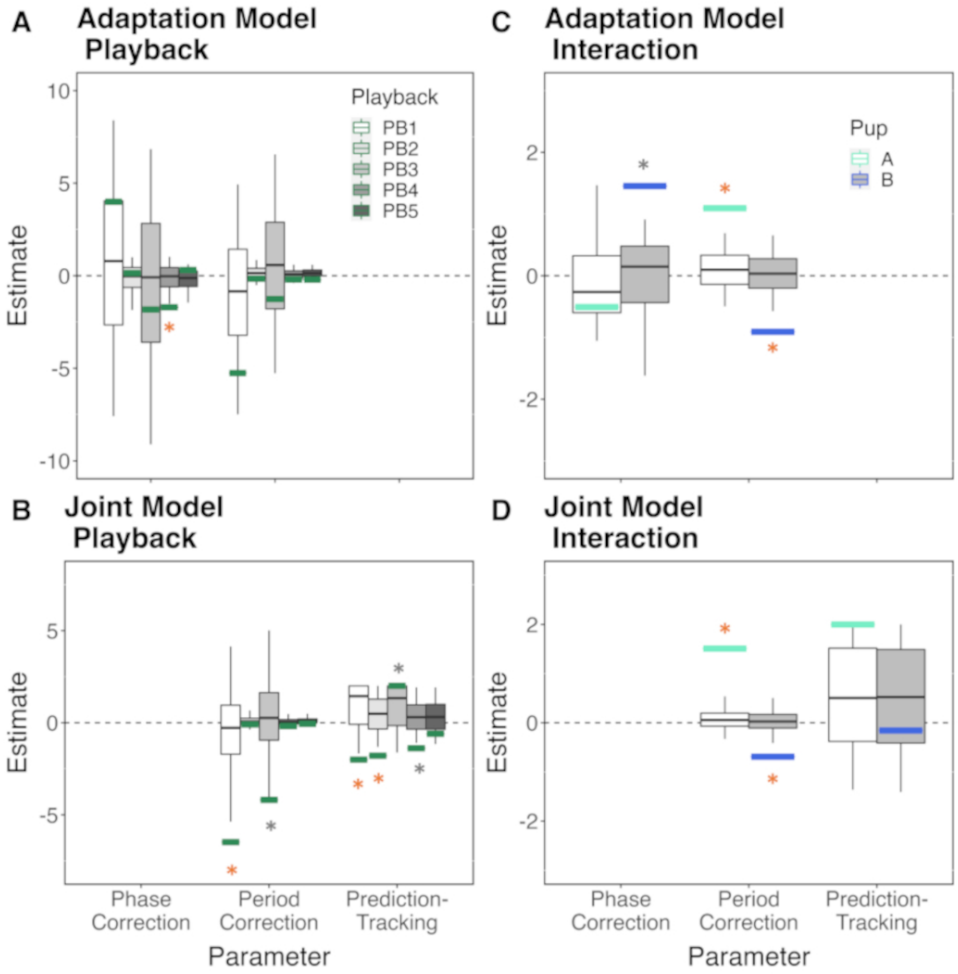

Figure 4. ADAM parameter estimates for seal pups A, B, and I. Left panels (A & B) show parameter estimates for pup I in five playback sessions (1-way interaction); right panels (C & D) show estimates for pups A and B (2-way interaction). Top panels (A & C) show phase correction and period estimates obtained with the adaptation-only version of ADAM; bottom panels (B & D) show period correction and prediction-tracking estimates for the full 'joint' model (anticipatory error correction estimates are not shown). Parameter estimates are shown separately for pup A relative to pup B (aquamarine bars) and pup B relative to pup A (blue bars) in the recording session, and for pup I (green bars) relative to playback calls in separate playback sessions (PB1-PB5). Corresponding parameter estimates for randomly permuted data from each session are represented by box plots, with the central horizontal line indicating the median, the bottom and top edges of the box indicating the 25th and 75th percentiles, and the whiskers indicating the 5th and 95th percentiles. Real estimates with asterisks are significantly different from permuted data estimates at the 2-tailed (orange) or 1-tailed (grey) level.

232x231mm (118 x 118 DPI)

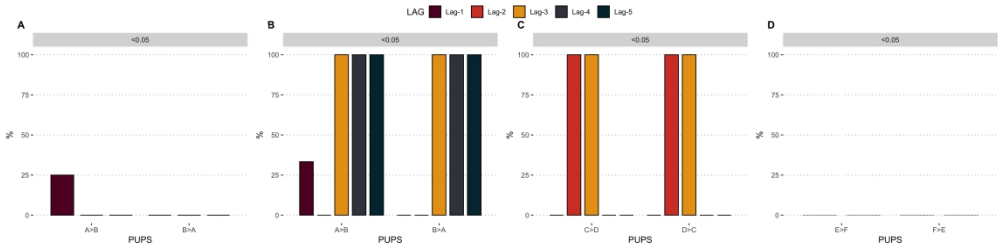

Figure S6. Granger causality results for the 2-way interaction context. The bar plots indicate the percentage of times (y-axis) that the timing of a calling pup was significantly influenced ($p < 0.05$) by another calling pup (i.e., the partner). We considered different pairs of pups bidirectionally and different lag values. We measured Granger Causality for pup pair A/B in two conditions: (A) different bouts within a recording and (B) whole recordings. For pup pairs C/D (C) and E/F (D) we measured Granger Causality only on whole recordings.

479x119mm (300 x 300 DPI)

Table 1. Summary table showing in order: research questions, analyses, contributing pups, predictions, whether the data supports each prediction, statistical test(s) used, and result(s) obtained. The column "Supported by data?" has three possible answers: results support the prediction (Y), results only partially support the prediction (Partial), and results do not support the prediction (N). Due to sample size and/or analytical requirements, not all pup data could be used in each analysis. The acoustic variables of interest for the different analytical approaches were: response phases (circular statistics), IOI ratios (categorical rhythms), IOIs (Granger causality, ADAM), and asynchronies (ADAM).

Research question	Analytical approaches	Pups	Prediction	Supported by data?	Statistical test(s) and result(s)
Temporal domain, Q1: *Does the presence of a calling partner affect the call timing of individual pups?*	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
			Pups will call in asynchrony to avoid overlap during vocal interactions.	Y	V-test: Pup calls start at one-quarter of the partner's period
			Response phases will be affected by the presence of a vocalising partner.	Y	Watson's U^2 test: Response phase distributions differ between non-interactive and interactive contexts.
	Categorical rhythms	A, B, C, E, H, I	Empirical and chance ratio distributions will only significantly differ when pups are vocally interacting.	Partial	1-sample KS tests: Simulated and empirical ratio distributions are rarely significantly different (exceptions: pup I alone, pup I 1-way, pup B 2-way).
		A, B, I	Vocally interacting pups will have a significant peak at the 4:1 SIR.	N	Paired Wilcoxon signed-rank test: No significant peaks at any of the tested small integer ratios.
			Ratio distributions of individual pups will differ across behavioural contexts.	Y	2-sample KS tests: Ratio distributions significantly differed for the same individuals in different behavioural contexts.

Social domain, Q2: *Does the type of calling partner (real or broadcasted) affect the call timing of individual pups?*	Circular statistics	A-I	Pups interacting with a real partner will show more adaptive call timing than the pup interacting with a broadcasted partner.	N	Watson's U^2 test: No difference in response phases between 1-way and 2-way interactions.
	Granger causality	A, B, C, D, E, F, I	The time series of a pup will be better predicted considering the time series of a vocalising partner rather than the time series of a broadcasted signal.	Y	Bidirectional & unidirectional causality: Interaction with a real partner impacted the pup's vocal behaviour more than the playback. Mutual temporal adaptation among pairs of vocally interacting pups.
Social domain, Q3: *When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) affect call timing?*	Circular statistics	A-I	The calling pattern of pups will show similar rhythmic structure in both the alone and silent partner conditions, as no vocal interaction is taking place in both cases.	Y	Watson's U^2 test: No difference between response phase distributions of pup calling alone and pups calling with a silent partner.
Cognitive domain, Q4: *Which timing mechanisms are used by pups during vocal interactions?*	ADAM	A, B, I	In the 1-way interaction, pups may show sensitivity to (non-)interactivity which would be reflected by parameter changes over repeated sessions. The 2-way interaction may be mediated by basic temporal adaptation and possibly higher-level anticipatory timing.	Partial	1-way interaction: Temporal anticipation, and to a lesser extent also adaptation, decreased across the playback sessions, and was absent in the final session. 2-way interaction: Clear evidence for temporal adaptation, with differing parameter estimates for each seal pup suggesting the emergence of different interactive roles.