



AperTO - Archivio Istituzionale Open Access dell'Università di Torino

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

This is the author's manuscript
Original Citation:
Availability:
This version is available http://hdl.handle.net/2318/1894773 since 2023-03-20T08:37:17Z
Published version:
DOI:10.1098/rstb.2021.0477
Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

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

1 2	
3 4	SCHOLAR ONE [™]
5	Manuscripts
7	
8 9	
10	
12	
13 14	
15	
17	
18 19	
20	
21 22	
23 24	
25	
26 27	
28 29	
30	
31 32	
33 34	
35	
36 37	
38	
40	
41 42	
43	
45	
46 47	
48	
50	
51 52	
53	
54 55	
56 57	
58	
59 60	http://mc.manuscriptcentral.com/issue-ptrsb

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.

1 2	1	Measuring rhythms of vocal interactions: A proof of principle in harbour seal
3 4	2	pups
5 6 7	3 4	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#}
8	5	
9	6	
10	7	¹ Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
11	8	² Department of Biological Sciences, Faculty of Natural Sciences, Norwegian University of Science and Technology
12	9	Alesund, Norway.
17	10	⁴ Division Animal Physiology & Behaviour Department for Neuroscience, Carl von Ossietzky University Oldenburg
15	11	Oldenburg Germany
16	13	⁵ Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium.
17	14	⁶ Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
18	15	⁷ Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.
19	16	⁸ Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands.
20	17	⁹ MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Australia.
21	18	¹⁰ Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music
22	19	#Joint senior authors
23	20	
24	$\frac{21}{22}$	
25	22	Correspondence
20	$\frac{23}{24}$	Marianna Anichini
28	24 25	e-mail: marianna anichini@mni nl
29	$\frac{25}{26}$	Andrea Ravignani
30	$\frac{20}{27}$	e-mail: andrea ravignani@mni nl
31	$\frac{27}{28}$	
32	20	
33	30	ORCID ID: MA_0000_0002_2697_0297; KdR_0000_0002_6129_9726; TAH_0000_0002_7891_
34	21	506Y: DV 0000 0001 6086 5125: ASC 0000 0002 2722 5426: CP 0000 0002 4018 5846: PK
35	22	590Λ , DV, 0000-0001-0080-5155, ASC, 0000-0002-5722-5450, CD, 0000-0005-4918-5840, FK,
30	32 22	0000-0001-7379-0313, AK, 0000-0002-1038-0024
38	23 24	
39	34 25	Abstract
40	35	Rhythmic patterns in interactive contexts characterise numan benaviours such as conversational turn-
41	36	taking. These timed patterns are also present in other animals, and often described as rhythm.
42	37	Understanding fine-grained temporal adjustments in interaction requires complementary quantitative
43	38	methodologies. Here, we showcase how vocal interactive rhythmicity in a non-human animal can be
44	39	quantified using a multi-method approach. We record vocal interactions in harbour seal pups (<i>Phoca</i>
45	40	vitulina) under controlled conditions. We analyse these data by combining analytical approaches,
40 17	41	namely categorical rhythm analysis, circular statistics, and time series analyses. We test whether pups'
47 48	42	vocal rhythmicity varies across behavioural contexts depending on the absence or presence of a
49	43	calling partner. Four research questions illustrate which analytical approaches are complementary vs.
50	44	orthogonal. For our data, circular statistics and categorical rhythms suggest that a calling partner

orthogonal. For our data, circular statistics and categorical rhythms suggest that a calling partner 44 affects a pup's call timing. Granger causality suggests that pups predictively adjust their call timing 51 45 52 46 when interacting with a real partner. Lastly, the ADaptation and Anticipation Model (ADAM) ⁵³ 47 estimates statistical parameters for a potential mechanism of temporal adaptation and anticipation. 54 Our analytical complementary approach constitutes a proof of concept; it shows feasibility in 48 55 49 applying typically unrelated techniques to seals to quantify vocal rhythmic interactivity across 56 behavioural contexts. 50 57

- 58 59 52 Keywords: Behavioural interaction, asynchrony, circular statistics, categorical rhythms, time series 60 53 analysis, interactive vocal rhythm
 - 54

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 malefemale 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

2

3

4 5

6 7

8 9

10

11 12

13

15

17

18 19

20

22

27

29 ²₃₀ 106

34

39

41

46

48

51

53

music, which is often produced in an interactive context [23]. Similar rhythms are also present in non-90 human animal songs (e.g., thrush nightingales [23], indris [24], but whether such rhythms characterise 91 92 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 93 94 conspecific attention, by making sequences of vocalisations more predictable to listeners. Categorical 95 rhythm analyses could therefore be an interesting method to test the predictability of vocal sequences 96 in non-human animal interactions. Lastly, the ADaptation and Anticipation Model (ADAM), 14 97 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 98 16 99 interaction [26,27]. Although ADAM is designed for 'simultaneous chorusing', it could also be used 100 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 21 101 ₂₃ 102 causality, circular statistics, categorical rhythm analysis, and ADAM—can be applied to a new animal ²⁴ 103 model: the harbour seal (Phoca vitulina). To illustrate the utility and compatibility of these different 26 104 analytical methods, we show how they can be used to better understand the rhythmic communication 28 105 of a small sample of harbour seal pups in different behavioural contexts.

³¹ 107 32 1.3. Our animal model

33 108 The "vocal learning-beat perception and synchronisation" (VL-BPS) hypothesis states that only 35 109 vocal learning species—those capable of producing new vocalisations or modifying existing ones ³⁶ 37 110 based on auditory experience—may possess advanced rhythmic abilities [29,30]. This hypothesis is ³⁸ 111 inherently cross-modal: it suggests a strong link between audition and timed movement. For example, 40 112 Snowball, a sulphur-crested cockatoo (Cacatua galerita eleonora), was shown to perceive auditory 42 113 rhythms at different tempi and to predictively synchronise his body movements to them [31]. Parrots 43 44 114 are phylogenetically distant from humans and, among mammals, pinnipeds (seals, sea lions, and 45 115 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 47 116 49 117 extract a beat from periodic acoustic stimuli and entrain to it in a predictive and adaptive manner-⁵⁰ 118 since some species showed vocal mimicry and plasticity [32,33] and others can keep a beat [34]. 52 1 1 9 These characteristics, paralleling human abilities, make pinnipeds an ideal animal clade for 54 120 comparative research on the origins of rhythmic communicative behaviour.

⁵⁵ 121 Harbour seals exhibit both vocal flexibility [33,35] and rhythmic interactivity [21], and are 57 122 particularly vocal in the first few weeks of life [36]. During the lactation period, harbour seal pups 58 59 1 2 3 emit 'mother attraction calls' (hereafter 'calls') to draw their mothers' attention [37]. Mothers are 60 124 silent and use the individual vocal signatures in these calls to recognize their pups [36,38]. Against

125 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 126 127 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. 128 129 Unlike cooperative types of turn-taking (e.g., in humans and in common marmosets (Callithrix 11 130 *jacchus*) [39]) harbour seal pups' interactions are a by-product of neighbouring pups vocalising to 12 131 attract their silent mothers, and are thus probably competitive.

14 1 3 2 To date, only two papers studied vocal rhythms in harbour seals, crucially both focusing on 16 133 single individuals [21,28]. The first study was a playback experiment in which a pup vocally 17 18 134 interacted with sounds broadcasted from a loudspeaker [21]. The pup adjusted the timing of its calls ¹⁹ 135 in an asynchronous manner by responding to the broadcasted conspecific calls with a non-uniformly distributed response phase whose mean approximated 90° [21]. The second study looked at the 21 1 36 23¹³⁷ presence and development of vocal rhythms in three harbour seal pups [28]. Complementary ²⁴ 138 analytical approaches showed how the pups' individual calling patterns gained more rhythmic 26 139 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 28 140 30141 response of the tested animals).

1.4. Aims & research questions

1

2 3

4 5

6 7

8 9

10

13

15

20

22

27

29

34

39

41

46

48

51

53

58

60

35 144 In this work, we show how vocal interactive rhythmicity in non-human animals can be ³⁶ 37 145 quantified using a multi-method approach spanning various research domains (e.g., temporal, social, ³⁸ 146 cognitive) (Table 1). We illustrate this approach through four research questions, all of which relate 40 147 back to whether harbour seal vocal interactive rhythmicity varies in different behavioural contexts 42 148 (Table 1). While our sample sizes are too small to enable species-wide inferences, they are sufficient 43 44 149 to illustrate how methods typically used to study human communication can be adopted to study 45 150 interactivity in animal communication. The goal of this paper is thus to outline a quantitative roadmap 47 151 that future research can follow. Circular statistics and categorical rhythm analysis are used to address 49 152 the first question about temporal adjustment in interaction: "does the presence of a calling partner ⁵⁰ 153 affect the call timing of individual pups?" (Q1). The next two questions consider the effect of 52 1 54 behavioural context on temporal adjustment: "does the type of calling partner (real or broadcasted) 54 155 affect the call timing of individual pups?" (Q2) and "when the focal pup is vocalising, does the ⁵⁵ 156 presence of a silent partner (vs. no partner) affect call timing?" (Q3). We answer these questions ⁵⁷ 157 using circular statistics (Q2 and Q3) and Granger causality tests (Q2). Lastly, ADAM is used to 59 1 58 investigate the fourth question about the cognitive processes involved in temporal adjustment: "which 159 timing mechanisms are used by pups during vocal interactions?" (Q4).

58

60

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) 1way 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, interonset 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 57 192 between the call onset of the partner and the call onset of the pup's response) and the previous IOI, 59 193 multiplied by 360, resulting in a unit vector with an angle on a circle. A vocal interaction was defined 194 as a group of three calls within the same bout, which includes two calls from the broadcasted/real 195 partner followed by the response of the focal individual (Figure 1). Following previous methodology

196 [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

1 2

3

4 5

17 18

₃₅ 202

₃₆ 203

37 204

38 205

39 206

43

46

48

53

55

197

43 44 210 Circular statistics, wherein periodic measures are converted to angles on a circle and compared ⁴⁵ 211 to distributions of interest [40], were used to investigate rhythmic periodicities in pup call response 47 212 phases (Method S4). Following [21], we considered the values of the response phases as circular data 49 213 falling between 0° and 360°. We obtained the circular mean (μ) (i.e., the average direction of the ⁵⁰ 214 response phases calculated from the pup calls; Table S3). Then, we ran Rayleigh z-tests to investigate ⁵²215 whether the distribution of response phases was uniform (e.g., arousal hypothesis) or showed a 54 2 1 6 unimodal peak (Table S4) [21]. Subsequently, we tested for uniformity against a specified mean 55 56 217 direction for the unimodal peak using a V-test [40–43].

⁵⁷218 Our data met the assumptions for circular statistics. We tested if the response phases in all four 58 59219 behavioural contexts followed a von Mises distribution using one-sample Watson tests (Table S5). 60 220 With deviations from uniformity (null hypothesis von Mises distribution rejected), we used Kuiper's

2 221 test, Watson's test, and Rao's spacing test, to confirm the p-value obtained from the Rayleigh test 222 (Table S6) as suggested by Landler et al. [40]. More details are shown in Method S4.

223 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 224 225 (1-way interaction) or real (2-way interaction) conspecific calls to avoid overlap and, hence, their 11 226 response phases would show a unimodal distribution. Following previous work [21] and applying the ¹² 227 V-test, we tested the null hypothesis of call response phase uniformity against two alternative 14 228 unimodal departures: 0° (i.e., synchrony) and 90° (i.e., asynchrony). Using Watson's two-sample U² 16 229 test [44], we also compared the call phase distributions of 1) a pup calling alone vs. when responding $^{17}_{18}230$ to a broadcasted partner (pup I) and 2) a pup calling in the presence of a silent partner vs. when their ¹⁹231 partner was also calling (pups A-H). For interacting (1-way or 2-way) pups, we applied Watson's two-21 2 3 2 sample U² test to assess whether the type of partner (i.e., real or broadcasted) differentially affected 23 233 the pups' response timing. We predicted that pups interacting with a real partner would show more ²⁴ 234 adaptive call timing, thanks to potential communicative cues from other modalities. Lastly, we 26 235 compared the distributions of call phases of the single pup vocalising alone to those of the paired pups 27 28 2 3 6 when their partner was silent to test whether the simple presence of a silent partner affects individual ²⁹ 30 237 call timing. In both behavioural contexts, we predicted that calling patterns for pups without a ³¹ 238 responsive partner would show a different rhythmic structure to those observed in interaction.

33 2 3 9 Categorical rhythm analysis tests whether the temporal intervals between signal onsets, as 34 35 240 inferred from IOI ratios, are distributed categorically rather than uniformly. We predicted that ³⁶ 37 241 empirical and simulated null ratio distributions (i.e., the expected distribution if no rhythmic ³⁸ 242 categories exist) will not differ when a pup is alone or with a silent partner but will differ when a pup 39 40 2 4 3 is vocally interacting. For vocally interacting pups, we predicted a significant peak in empirical ratio 41 42⁴⁴ distributions at the 4:1 small integer ratio based on the lone seal in [21], which called at approximately 43 44 245 one-quarter of the playback's period. All categorical rhythm analyses were done within bouts 45 246 following previous methodology (section Method S5) [23,24], with IOIs calculated in various ways 46 depending on the behavioural context (Table S8, Figure 1). We used one-sample Kolmogorov-47 247 48 49 248 Smirnov (KS) tests to determine whether empirical IOI ratio distributions significantly differed from ⁵⁰ 249 simulated null IOI ratio distributions. Our data met the one-sample KS test assumptions, namely that 51 52 2 50 the sample is random and the theoretical distribution is continuous and fully defined. When the 53 54 251 empirical and simulated distributions were significantly different, we also looked for evidence of ⁵⁵₅₆ 252 small integer ratio (SIR) categorical rhythms—specifically at the 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, and 4:1 57 253 ratios—which have been found in other species' vocalisations [23,24]. In these analyses, the empirical 58 59 254 ratio distributions were divided into "on-integer" and "off-integer" ratio bins (Table S7). On- and off-60 255 integer bin counts for each SIR were normalised by bin size and compared using a paired Wilcoxon

256 signed-rank test (a non-parametric test that allows for non-normality in the population data and 2 257 assumes paired differences are continuous, symmetrically distributed, and mutually independent). 258 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 259 260 individual pups differed across behavioural contexts.

1

3

4 5

6 7

8 9

10 11²⁶¹ The Granger causality test investigates whether the values of a time series A are better ¹²/₁₂ 262 predicted when considering the values from a second time series B, as opposed to only using values 13 14 263 from time series A [45]. Here, we assessed whether the call timing of a pup partaking in a 1-way or 15 16 264 2-way vocal interaction can be predicted using the call timing of its partner. More specifically, to $^{17}_{18}265$ investigate whether the call timing of a pup differed in relation to the type of partner, we tested ¹⁹266 whether there is a difference in predicting the time series of the pup interacting with a broadcasted 20 21 267 partner vs. time series of the pups interacting with a real partner. Previous work showed that 22 23 268 individuals respond to conspecific calls with a non-random pattern [18-20,46]. We therefore expected ²⁴ 269 that the time series of a pup can be better predicted considering the time series of a vocalising partner 26 270 rather than those of a broadcasted signal. In both the 1-way and 2-way interactive scenarios, we 27 28 271 considered Granger causality at two levels: (1) the entire recording, regardless of the length of the ²⁹ 30 272 pauses between consecutive calls, and (2) different bouts within each recorded session. We restricted ³¹ 273 the analysis on the different bouts to call sequences that were long enough to generate accurate 33 274 estimates (i.e., a minimum of five paired calls [47]). The bouts included in this analysis range from 5 34 35 275 to 20 calls. We conducted the Granger causality test using call onsets and different lag measures, from ³⁶ 37 276 one to five (Method S6), testing whether the previous one to five onsets in the first time series can be ³⁸ 277 used to better predict the second time series (Figures S5/S6, Table S11). For the 1-way interaction, 39 40 278 we performed a one-way analysis, considering whether the pup's timing could be predicted using the 41 41 42 279 playback timing. For the 2-way interaction, we performed a two-way analysis to assess whether the 43 44 280 two interacting pups influenced the timing of each other's calls.

45 281 We used **ADAM** to test for evidence of reactive error correction and predictive processes in the 46 47 282 1-way and 2-way interactive scenarios. ADAM consists of three computational modules that interact 48 49 283 via internal models of 'self' and 'other' that support one's own action planning and external event 50 284 prediction, respectively (Figure S2). The adaptation module compensates for synchronisation errors 51 52 285 by implementing error correction processes that alter the phase and/or period of an internal timekeeper 53 54 286 controlling for action (here, call) timing. These error correction processes determine the provisional ⁵⁵ 287 timing of the next planned action by providing input to an internal model of the 'self'. The 57 288 anticipation module computes the expected timing of upcoming events based on the weighted sum 58 59 289 of two processes: the linear extrapolation of previous IOIs in the sequence and the copying (or ⁶⁰ 290 'tracking') of the previous IOI, with the output informing temporal predictions generated by the

26 305

'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 11 296 segregation (distinction) of information about 'self' and 'other' [48,49]. Each process instantiated in ¹² 297 ADAM is controlled by an independent parameter, and the value of these parameters can be estimated 14 298 for a particular individual by fitting the model to behavioural time series data [26,27,50,51]. 16 299 Parameter estimates were obtained for both the adaptation-only version of ADAM—which includes 18 300 phase correction and period correction-and the full ('joint') version-including period correction, ¹⁹ 301 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 21 3 0 2 ₂₃ 303 (real or broadcasted) have a steady base tempo (for which adaptation is sufficient) or a systematically ²⁴ 304 changing tempo (which benefits from both anticipation and adaptation) (Method S7).

28 306 Table 1. Summary table showing in order: research questions, analyses, contributing pups, predictions, 29 307 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	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
calling partner affect the call timing of individual pups?			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.

			1		
	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.	Ν	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	Circular statistics	A-I	Pups interacting with a real partner will show more adaptive call timing than the pup interacting with a broadcasted partner.	Ν	Watson's U ² test: No difference in response phases between 1-way and 2-way interactions.
pups?	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.

23

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

²⁰ 21 315 3. Results

²² 316 3.1. Does the presence of a calling partner affect the call timing of individual pups? (Q1)

24 3 17 All pups' data entered the circular statistics analysis. Running the Rayleigh test, we found that 25 26 318 the response phase distribution was uniform for pup I which was recorded alone (z = 0.04, p = 0.254; 27 28 319 Figure 2A), whereas it was non-uniform for pups A-H which were recorded with a silent partner (z =29 320 0.11, p < 0.001; Figure 2A; Table 1). This non-uniformity may have been driven by the individual 30 31 321 contributions of pups B and C, which had non-uniformly distributed response phases (Table S4), ³² 33 322 whereas the other six pups had a uniform distribution. The Rayleigh tests run anew in the interactive ³⁴ 323 contexts, showed that the response phase distributions of pup calls were unimodal in both the 1-way 36 324 (pup I: z = 0.39, p < 0.001) and 2-way (pups A - H: z = 0.41, p < 0.001) interactions (Figure 2B; Table 37 1). Applying the V-test in both contexts, the direction of the response phases did not statistically match 38 325 ³⁹ 326 0° (1-way: z = -0.02, p = 0.587; 2-way: z = 0.06, p = 0.110), suggesting that pups did not synchronise 41 327 with their partner (real or broadcasted). However, the response phase direction did match 90° (1-way: 42 43 328 z = 0.38, 2-way: z = 0.41, p < 0.001), supporting the previously reported evidence of asynchronous 44 45 329 calling behaviour [21]. A Watson's two-sample U² test confirmed that the response phase distributions 46 47 330 significantly differed between the alone vs. 1-way interaction context for pup I (U² 1.76, p < 0.001), 48 331 and between the silent partner vs. 2-way interaction context for pups A-H ($U^2 = 0.78$, p < 0.001, 49 50 332 Figure S3; Table 1). Finally, the circular standard deviation values were higher for the alone and silent 51 52 333 partner contexts compared to both interactive vocal contexts, indicating a larger dispersion of the ⁵³ 334 54 response phases for the former conditions. This outcome is also confirmed by the values for the mean 55 335 resultant length. 56

- 57
- 58
- 59 60



Figure 2. Circular histogram plots (bin width $= 20^{\circ}$) 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 (1way/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).

²⁶ 343 In the categorical rhythm analyses, the empirical ratio distribution was significantly different 28 3 4 4 from chance when pup I was recorded alone (Figure 3A) and during the 1-way interaction (Figure 29 30 3 4 5 3C). In both contexts, there was no evidence of significant peaks at any of the tested ratios (Table 1). 31 31 346 When the playback calls were disregarded from IOI calculations (Figure 3B), there was no significant ³³.347 difference in empirical and simulated ratio distributions for pup I. However, pairwise KS tests showed 34 35 3 48 that the ratio distributions significantly differed when comparing each of the three behavioural 36 37 349 contexts (alone vs. 1-way interaction disregarding playback vs. 1-way interaction when pup I 38 39 350 responds) to each other (Table S10; Table 1). For the five pups that were well-sampled in the silent 40 351 partner context (pups A, B, C, E, and H; Table S8), the empirical ratio distributions did not 41 42 3 5 2 significantly differ from chance (Table S9, Figures 3/S4, Method S4). Finally, when considering 2-43 44 353 way interactions, only pups A and B were well-sampled enough (i.e., had at least ten ratios for both 45 354 the silent partner and 2-way interaction contexts) to compare, but the empirical and simulated ratio 46 47 355 distributions were significantly different only when pup B was the responder (Figure 3; Table S8; 48 49 356 Table S9). Once again, there were no significant peaks at any of the tested ratios for pup B. For both 50 50 51 357 pups A and B, the empirical ratio distributions significantly differed when comparing different ⁵² 358 behavioural contexts (Table S10). Collectively, there was thus little evidence of SIR rhythmic 53 54 3 59 categories in pup calls, but IOI ratios did significantly differ when looking at the same individuals in 55 ₅₆ 360 different behavioural contexts (Table 1). 57

57 58 361 Thus, the results from both analyses suggest that the presence of a calling partner does affect ⁵⁹ 362 the call timing of the focal pup, in terms of both call response phases and IOI ratios. 60

363

1 2

3 4

5

6

7

8

9

10

11

16

17 18

20 338

21 339

3

4 5

6

7

8 9

35

37

40

42

45



23 365 Figure 3. IOI ratio density plots for pups in different behavioural contexts. Pup I: (A) alone, (B) 1-way 24 366 interaction (disregarding playback), (C) 1-way interaction (responding to playback). Pup A: (D) silent partner, 25 367 (E) 2-way interaction (responding to partner). Pup B: (F) silent partner, (G) 2-way interaction (responding to ²⁶ 368 partner). For each plot, the dashed vertical lines indicate, from left to right, 1:4, 1:3, 1:2, 1:1 (i.e., isochrony), 20 368 27 369 28 370 29 371 30 372 31 372 32 373 33 374 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.

34 375 3.2. Does the type of calling partner (real or broadcasted) affect the call timing of individual 36 376 **pups? (Q2)**

37 38 377 To address this research question, we once again used circular statistics and the response phase 39 378 distributions of all nine pups. Specifically, we compared calls from pup I during the playback (1-way 41 379 interaction) with calls of pups A-H when their partner was also calling (2-way interaction). A Watson's 43 380 two-sample U² test statistically confirmed that the response phase distributions did not differ between ⁴⁴_381 the 1-way and 2-way vocal interactions ($U^2 = 0.07$, p > 0.10; Table 1).

46 382 Interestingly, however, the Granger causality results from seven pups (A, B, C, D, E, F, I) 47 48 383 showed that call timing behaviour differed depending on the type of partner (Table 1). For the 1-way ⁴⁹ 384 interaction, five different playback sessions featuring pup I were considered (ranging from 34 to 121 ⁵¹ 385 paired calls) and the timing of the pup's calls were never significantly predicted by the timing of the 52 53 386 playback (Table S11, Figure S5). For the 2-way interaction, five different recording sessions were 54 55 387 considered (ranging from 8 to 71 paired onsets). We found that in two sessions, the timing of the first ⁵⁶ 388 57 pup did not significantly predict the timing of the second, in either direction (pair A>B and B>A; pair 58 389 E>F and F>E; Table S11, Figure S6B/S6D). However, in two other sessions, the timing of a pup was 59 significantly predicted by the calling partner in both directions, across different lag values (A>B: lag-60 390

2 391 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 392 S6B/S6C). In the last session, the timing of pup A could be predicted by that of pup B, while we found 393 no indication of temporal adjustment for pup B. At the bout level, we restricted our analyses to four 394 series of paired onsets featuring pups A and B (as they were the only pup pair that met the sample 395 size requirements). In this last scenario, only the timing of pup B was significantly influenced by pup 10 11 396 A (Figure S6A).

1

3

4 5

6 7

8 9

13

15

20

22

27

29

 $^{17}_{18}400$

¹² 397 While the circular statistics results thus suggest that the type of calling partner (real vs. 14 3 98 broadcasted) does not affect focal pup call timing, the Granger causality results suggest that in certain 16 399 2-way (but not 1-way) interactions, focal pup call timing can be predicted by the partner's call timing.

¹⁹ 401 3.3. When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) 21 4 0 2 affect call timing? (Q3)

23 403 Using circular statistics, we compared the calls of pup I recorded alone with those of the 8 other ²⁴ 404 pups (A-H) recorded with a silent partner present. Watson's two-sample U² test results show that the 26 4 0 5 response phase distributions did not statistically differ between the two contexts ($U^2 = 0.10, p > 0.05$). 28 4 06 In other words, having a silent partner was essentially the same as having no partner—in terms of the ²₃₀407 effect on focal pup response phase distributions-for the pups in our study. This is intuitive, given ³¹ 408 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? (O4)

36 37 411 ADAM parameter estimation was conducted on call data from all five playback sessions ³⁸ 412 featuring pup I (1-way, Figure 4 A/B), but was restricted to pups A and B for the 2-way interactions 39 40 4 1 3 due to sample size limitations (Figure 4 C/D). Interactive vocal bouts were concatenated to obtain a 41 42 414 time series length which would provide reliable ADAM parameter estimates. A simulation test then 43 44 415 ensured that the estimates were not compromised by differing sequence lengths or by the 45 416 concatenation procedure (Method S7). The reliability of observed parameter estimates was tested by 46 47 417 comparison against corresponding values for randomly permuted data (Method S7). The quality of 48 49</sub>418 the fits to the data did not differ significantly between versions of ADAM (Method S7). Results for ⁵⁰ 419 the 1-way interaction featuring pup I were remarkable with regards to typical human data (e.g. 51 [26,27]) as most significant parameters were negative in sign (Figure 4A/B, Tables S13/S14). 52 4 2 0 53 54 421 Negative phase and period correction estimates indicate that calling earlier will lead to a shortening ⁵⁵ 422 of the next IOI, while calling later will lead to a lengthening of the next IOI. Negative 57 423 prediction/tracking estimates mean that when the playback's IOIs increase (i.e., deceleration), the 58 pup's IOIs will decrease (i.e., acceleration), and vice versa. This systematically enhances the timing 59 4 2 4 60 ′425 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 11⁴³¹ 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. 14 4 3 3 Results for this type of vocal interaction showed evidence for temporal adaptation. In particular, 16 4 3 4 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 21 4 37 to a lengthening of the next IOI, whereas calling later will lead to a shortening of the next interval. 23 438 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 26 4 40 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 28 4 4 1 ₃₀ 442 roles that dynamically emerge within the context of this specific pairing of pups.



60 4 4 5

Figure 4. ADAM parameter estimates for seal pups A, B, and I. Left panels (A & B) show parameter estimates 2 446 3 447 for pup I in five playback sessions (1-way interaction); right panels (C & D) show estimates for pups A and B 4 448 (2-way interaction). Top panels (A & C) show phase correction and period estimates obtained with the 449 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 450 451 are shown separately for pup A relative to pup B (aquamarine bars) and pup B relative to pup A (blue bars) in 452 the recording session, and for pup I (green bars) relative to playback calls in separate playback sessions (PB1-453 PB5). Corresponding parameter estimates for randomly permuted data from each session are represented by 10 10 454 11 454 12 455 13 456 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) 14 457 level. 15 458

¹⁶ 4. Discussion

1

5

6

7

8

9

17

19

21

26

31

33

45

50

18 460 This study sought to provide a methodological proof of concept for quantifying vocal 20 461 rhythmicity in non-human animal interactive communication. Particularly, we showed how 22 462 complementary methodologies can be used to test whether the (Q1) presence and (Q2) type of a ²³ 463 24 vocalising partner, or the (Q3) presence of a silent partner affect patterns in animal communication 25 464 (Table 1). We also investigated which (Q4) underlying cognitive mechanisms potentially play a role 27 465 in vocal interactions (Table 1). The different analytical techniques proved fruitful; combining and 28 29 466 contrasting their results could highlight nuances in rhythmic behaviour. Nonetheless, our sample size ³⁰ 467 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 32 468 33 34 469 with the approaches we present, can be adapted and expanded in future work to better understand the ³⁵ 470 temporal, social, and cognitive processes underlying interactive communication in animals.

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

40 41 473 The prediction of overlap avoidance was supported by circular statistics, which showed that the 42 474 43 distributions of response phase angles were unimodal. When vocally interacting, pups responded at 44 475 approximately one-quarter of the playback/real partner calling period (90°) and showed phase angles 46 476 significantly different from 0°, supporting previous results from one individual (pup I) [21]. Pups 47 48</sub>477 responding to conspecifics seem to time their calls to avoid overlap, consistently de-synchronizing ⁴⁹ 478 their call onsets. When pups were recorded alone or with a silent partner, the distribution of phase 51 479 angles was uniform, with no significant rhythmic pattern. Our findings confirm and extend previous 52 53 480 outcomes [21]; in our limited sample, pups not only react to playbacks with asynchrony, but also ⁵⁴ 481 55 respond to conspecific partner calls in an asynchronous manner.

56 4 8 2 Regarding categorical rhythms, the empirical ratio distribution did not significantly differ from 57 58 483 the distribution expected by chance if IOIs were uniformly distributed for most pup/behavioural ⁵⁹ 484 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 485

3

4 5

6 7

8 9

41

46

2 486 categorical rhythms at the seven tested small integer ratios. This includes the 4:1 ratio, which we 487 hypothesised would frequently occur when pups were vocally interacting with playback stimuli or 488 partners based on previous work [21]. Importantly, however, the empirical ratio distributions 489 significantly differed across behavioural context conditions for the three pups (A, B, I) for whom such 490 comparisons were possible. When alone or with a silent partner, pup calls generally showed unimodal 10 11 491 ratio distributions centred around isochrony, whereas calls emitted by vocally interacting pups ¹² 492 showed a clear right-shifted peak, or a second peak to the right of isochrony, indicating that the pup's 13 14 4 93 IOIs during vocal interactions were generally shorter than the playback's/partner's IOIs. This context-15 16 494 dependent asymmetry bolsters results related to Q4, namely that the presence of vocalising partner $^{17}_{18}495$ can significantly impact pup vocal behaviour, with interacting individuals trying to prevent call ¹⁹ 496 overlap. Roeske and colleagues [23] hypothesised that categorical rhythms may be prevalent in calls 20 21 4 97 meant to attract and hold conspecific attention, such as the calls pups use to attract the attention of 22 ²²₂₃ 498 their mothers. We did not find evidence of rhythm categories at seven tested ratios; this negative result ²⁴ 499 25 has methodological value, highlighting that not all species that produce attention-seeking 26 500 vocalisations show integer ratio categories. 27

28 501 A vocal interaction with non-adaptive playback stimuli (1-way) may represent a limitation to ²⁹ 30 502 studying spontaneous vocal production. The extent of this limitation can be gauged by comparing ³¹ 503 results to data from vocal interactions with a real partner (2-way), and with the use of complementary 33 504 analytical methodologies, such as categorical rhythm analysis and circular statistics.

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

³⁸ 507 39 Findings from Q2 show contrasting results. Call phases were statistically the same no matter if 40 508 pups could interactively adjust their call timing to each other (2-way) or not (in 1-way). Partly in 42 509 contrast with this, the Granger causality analysis showed how vocally interacting with a real 43 44 510 individual impacted the pup's vocal behaviour more than interacting with a broadcasted partner 45 511 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 47 512 48 49 513 reciprocal adjustment of timing behaviour. In contrast, we found little evidence that a pup adjusts the ⁵⁰ 514 51 time series of its calls to a playback series. Indeed, in most cases, the pup's call timing could not be 52 515 predicted by the playback's call timing.

53 54 516 Together, findings from circular statistics (Q1) and Granger causality (Q2) point towards a ⁵⁵ 517 directionality-overlap avoidance relationship, which has also been suggested for other non-human 57 518 species (e.g., primates [18,52]; birds [53]; amphibians [54]; seals [55]). Interestingly, when infant 58 marmosets interact with their parents, the probability that their vocalisations will overlap with those 59 519 60

2 520 of adults decreases over time [22], suggesting that turn-taking in some mammals is a learned vocal 521 behaviour scaffolded by active parental feedback [39].

523 Q3. A silent partner does not trigger call rhythmicity

1

3

4 5 522

6 7

8 9

10

15

21 5 3 1 22 23 532

34

41

46

524 Circular statistics indicated that pups did not show any periodic pattern both, either when calling 11 525 alone or with a silent partner present (Q1). This is consistent with the fact that wild pups produce calls 12 526 13 to attract a silent mother [36]: if no other calling conspecifics are present, there is no need to adjust 14 527 one's call timing and vocalisations are produced with a random onset. Moreover, given that the 16 528 response phase distributions did not differ between pups recorded with a silent partner or alone, it $^{17}_{18}529$ suggests that the presence of a silent partner did not trigger variable calling behaviour in our study ¹⁹ 530 20 population.

Q4. Purported timing mechanisms for vocal interactions

²⁴ 533 The ADAM analysis suggests that seal pups may perceive temporal patterns [56], which arise 26 534 27 between their calls and those of others, and react to them by adopting different mechanisms for 28 5 35 temporal adaptation. The negative parameter estimates that we observed in the 1-way interaction ²⁹ 30 536 sessions (consistent with enhancing the distinction between calls) could reflect attempts to lead the ³¹ 537 32 temporal interaction or even to test the responsiveness of the (broadcasted) partner by introducing 33 538 timing perturbations and gauging their effects. The decrease of temporal adaptation and/or 35 539 anticipation observed across playback sessions is consistent with a gradual process of habituation, ³⁶ 37 540 with the pup possibly learning that the playback is non-interactive. Harbour seals are capable of ³⁸ 541 39 acoustic recognition based on habituation paradigms; they can discriminate among different stimuli 40 542 and selectively habituate to them [57]. The lack of temporal adaptation to the playback stimulus we 42 543 observed in a seal pup may entail similar habituation processes. From a comparative perspective, 43 44 544 these findings also suggest a sensitivity to social contingency that may be analogous to capacities in 45 545 human infants, who become disinterested and display fewer signs of positive affect during vocal 47 546 interactions with non-responsive or delayed video recordings of their mothers [58-62]. Future studies 48 49 547 could address the role of temporal contingency by using interactive playback sequences [63–65].

⁵⁰ 548 51 Ours constitutes the first attempt to apply ADAM to non-human animals. On the technical side, 52 549 this necessitated the validation of an approach where brief interactive vocal bouts were concatenated 53 54 550 to produce longer time series and thereby reduce the risk of model overfitting (Discussion S1). ⁵⁵ 551 Demonstrating the utility of this procedure opens the door to applying the model in a wider range of 57 552 behavioural contexts. Nevertheless, caveats are necessary when interpreting the seal data in light of 58 previous work with ADAM in humans, where individuals intentionally produce movements, whose 59 553 ⁶⁰ 554 sensory effects occur simultaneously with rhythmically regular sounds (e.g. [26,27]). Assumptions

Page 21 of 55

1

3

4 5

6 7

8 9 559

2 555 about intentionality and simultaneity may not apply to seal pup vocal interactions or rhythmic 556 interactions in other animals [6,9]. Points of convergence in the main outcomes of the complementary 557 analysis techniques suggest that ADAM, like categorical rhythm analysis, circular statistics, and 558 Granger causality, may also be robust and informative under such conditions.

10 11 560 **5.** Conclusions

12 561 13 Crucially, our work highlights the efficacy of combining multiple methods to study rhythmic 14 562 vocal behaviour. Our approaches vary in the degree to which they capture global temporal 15 16 563 characteristics across events vs. local temporal dependencies between events. Global measures (e.g., $^{17}_{18}564$ from circular statistics or categorical rhythm analysis) reveal predominant rhythmic features of a ¹⁹ 565 20 vocal interaction while local time series measures (e.g., from Granger causality or ADAM) provide 21 566 information about how these features might arise. Through this approach we could tease apart rhythm 22 23 567 nuances in our dataset, further develop harbour seals as a model species, and demonstrate how certain ²⁴ 568 analyses often restricted to humans, such as categorical rhythms (but see [23,24]) and ADAM, can be 26 569 27 applied to non-human animals. The categorical rhythm and circular statistical analyses tackled similar 28 570 questions from different angles, namely whether the distributions of IOI ratios (the former) or call 29 30 571 response phases (the latter) significantly differed across behavioural contexts. The categorical rhythm ³¹ 572 analyses also sought to determine whether the rhythmicity of pup calling behaviour conforms to small 33 573 integer ratios. Our negative result is, to our knowledge, the first published case of a species for which 34 35 574 categorical rhythms are clearly absent from vocalisations, which adds to understanding of how, why, ³⁶ 37 575 and when such rhythms evolve in communication systems. Time series analyses such as Granger ³⁸ 576 39 causality allowed testing for timing adjustment. In the case of ADAM, mechanisms of temporal 40 577 adaptation and anticipation that have previously been used to describe rhythmic behaviour in humans 41 42 578 [66] were used to describe rhythmic behaviour in seals. Interestingly, the ADAM model provides an 43 44 579 empirical warning about potential seal pup habituation effects when vocally interacting with a 45 580 recorded partner. 46

Though our sample size is small, studies on single individuals are not unusual in comparative 47 581 48 49 582 research [20,67]. Nevertheless, it is possible that the lack of adjustment to a playback, the adaptation ⁵⁰ 583 51 to a real individual, or both, reflect a peculiar vocal behaviour of the individuals we tested and cannot 52 584 be generalised to the species as a whole. The pups in this study were in a temporarily captive setting, 53 54 585 albeit in acoustic proximity to other individuals, similar to conditions they would experience in ⁵⁵ 586 nature. Unfortunately, vocal development in harbour seal pups has not been studied in wild colonies, 57 587 hence we do not know whether captivity affects their vocal development. We do know, however, that 58 pups vocally interact with neighbouring pups in the colony and not with their silent mothers; hence 59 588 60

2 589 the turn-taking behaviour observed in our captive conditions might extend to the same behaviour in 590 wild conspecifics.

591 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 592 9 593 triggered by individual-specific behaviours or by cues from other modalities. Multimodal 10 11 594 communication should be the target of future studies since we cannot assume that such interactions 12 595 13 are limited to acoustic cues. Despite these limitations, our study shows that adopting multiple 14 596 complementary approaches can be a fruitful way to study rhythmic interactivity in non-human animal 15 16 597 communication. 17 18 598

¹⁹ 599 Acknowledgements

1

3

4 5

6 7

8

20

22

27

29

We are grateful to the editor and the two anonymous reviewers who provided insightful 21 600 23 601 comments and suggestions on an earlier version of this manuscript. We thank all the volunteers, seal ²⁴₂₅ 602 care members and veterinarians of Sealcentre Pieterburen who helped during the experiment and in 26 603 the rehabilitation process of the seal pups. We are thankful to Chiara De Gregorio for sharing the R 28 604 code used for the categorical rhythm analyses and to Laura Verga for providing helpful comments on ____605 a draft of this manuscript.

rez oniz

- 2 606 References
- 3 607 1. Dusenbery DB. 1992 Sensory ecology: How organisms acquire and respond to information. 4 608 Freeman, New York. 5
- 609 6 2. Stegmann UE. 2013 Animal Communication Theory: Information and Influence. Cambridge, UK:
- 7 610 Cambridge University Press. (doi:10.1017/CBO9781139003551)
- 8 611 3. Anichini M, de Heer Kloots M, Ravignani A. 2020 Interactive rhythms in the wild, in the brain, 9 and in silico. Can. J. Exp. Psychol. 74(3), 170-175. (doi: 10.1037/cep0000224)
- j₁₀ 612
- 11 613 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) 12614
- 13 615 5. Greenfield MD, Aihara I, Amichay G, Anichini M, Nityananda V. 2021 Rhythm interaction in
- 14616 animal groups: selective attention in communication networks. Phil. Trans. R. Soc. B 376(1835),
- ¹⁵ 617 20200338. (doi:10.1098/rstb.2020.0338)
- 16 618 17 619 18 619 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) 19 620
- 20 621 7. Levinson SC. 2006 Cognition at the heart of human interaction. *Discourse studies*. 8(1), 85-93.
- 21 6 2 2 8. Levinson SC. 2016 The Interaction Engine hypothesis. The Language in Interaction Summerschool 22 623 on Human Language: From Genes and Brains to Behavior. Berg en Dal, The Netherlands.
- ²³ 624 9. Ravignani A, Verga L, Greenfield MD. 2019 Interactive rhythms across species: the evolutionary
- ²⁴ 625 biology of animal chorusing and turn-taking. Annals of the New York Academy of Sciences, 1453(1), 26⁶²⁶ 12-21. (doi:10.1111/nyas.14230)
- 27 627 10. Levinson SC. 2016 Turn-taking in human communication-origins and implications for language 28 6 28 processing. Trends in cognitive sciences. 20(1), 6-14. (doi.org/10.1016/j.tics.2015.10.010)
- 29 6 29 11. Pika S, Wilkinson R, Kendrick KH, Vernes SC. 2018 Taking turns: bridging the gap between 30 630 animal communication. Proc. *R*. 285(1880), human and Soc. B 20180598. $31 \\ 32 \\ 33 \\ 632 \\ 63$ (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) 34 633
- 35 634 13. Schulz TM, Whitehead H, Gero S, Rendell L. 2008 Overlapping and matching of codas in vocal 36 6 35 interactions between sperm whales: insights into communication function. Anim. Behav. 76(6), 1977-
- 37 636 1988. (doi:10.1016/j.anbehav.2008.07.032)
- ³⁸ 637 14. Takahashi DY, Narayanan DZ, Ghazanfar AA. 2013 Coupled Oscillator Dynamics of Vocal Turn- $\begin{array}{r}
 39 \\
 40 \\
 41 \\
 639 \\
 41
 \end{array}$ 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. 42 640 Routledge.
- 43 641 16. Yoshida S, Okanoya K. 2005 Evolution of turn-taking: a bio-cognitive perspective. Cognitive 44 6 4 2 Bulletin of Japanese Cognitive Science Studies: the Society. 12(3). 153-165.
- 45 643 (doi:10.11225/jcss.12.153)
- ⁴⁶ 644 17. Henry MJ, Cook PF, de Reus K, Nityananda V, Rouse AA, Kotz SA. 2021 An ecological approach
- 47 48 645 to measuring synchronization abilities across the animal kingdom. Phil. Trans. R. Soc. B. 376(1835), .0 49 646 20200336. (doi:10.1098/rstb.2020.0336)
- 50 647 18. Gamba M, Torti V, Estienne V, Randrianarison RM, Valente D, Rovara P, Bonadonna G, Friard 51 648 O, Giacoma C. 2016. The Indris Have Got Rhythm! Timing and Pitch Variation of a Primate Song
- 52 649 Examined between Sexes and Age Classes. Front. Neurosci. 10, 249. (doi:10.3389/fnins.2016.00249)
- ⁵³650 19. Clink DJ, Tasirin JS, Klinck H. 2020 Vocal individuality and rhythm in male and female duet
- ⁵⁴ 651 ⁵⁵ 652 contributions of a nonhuman primate. Curr. Zool. 66(2), 173–186. (doi:10.1093/cz/zoz035)
- 56 652 20. Fuhrmann D, Ravignani A, Marshall-Pescini S, Whiten A. 2014 Synchrony and motor mimicking ₅₇ 653 in chimpanzee observational learning. Sci. Rep. 4, 1–7. (doi:10.1038/srep05283)
- 58 654 21. Ravignani A. 2019 Timing of antisynchronous calling: A case study in a harbour seal pup (Phoca 59 655 vitulina). J Comp Psychol. 133(2), 272–277. (doi:10.1037/com0000160)
- 60

- 2 656 22. Takahashi DY, Fenley AR, Ghazanfar AA. 2016 Early development of turn-taking with parents
- 3 657 shapes vocal acoustics in infant marmoset monkeys. Phil. Trans. R. Soc. B: Biological Sciences.
- 4 658 **371(1693)**, 20150370. (doi:10.1098/rstb.2015.0370) 5
- 23. Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical rhythms are shared 659 6 between songbirds and humans. Curr. Biol. 30, 3544–3555. (doi:10.1016/j.cub.2020.06.072) 660 7
- 24. De Gregorio C, Valente D, Raimondi T, Torti V, Miaretsoa L, Friard O, Giacoma C, Ravignani A, 661 8
- 9 662 Gamba M. 2021 Categorical rhythms in a singing primate. Curr. Biol. 31, R1379-R1380. 10 663 (doi:10.1016/j.cub.2021.09.032)
- 11 664 25. van der Steen MC, Keller PE. 2013 The ADaptation and Anticipation Model (ADAM) of ¹² 665 sensorimotor synchronization. **Frontiers** in Human Neuroscience, 7, 253.
- 14⁶⁶⁶ (doi:10.3389/fnhum.2013.00253)

- 26. van der Steen MC, Jacoby N, Fairhurst MT, Keller PE. 2015 Sensorimotor synchronization with 15 667 tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. Brain Res. 16 668 17 669 1626, 66-87. (doi:10.1016/j.brainres.2015.01.053)
- 18 670 27. Mills PF, Harry B, Stevens CJ, Knoblich G, Keller PE. 2019 Intentionality of a co-actor influences ¹⁹671 sensorimotor synchronisation with a virtual partner. *Quarterly Journal of Experimental Psychology*, ²⁰ 672 21 672 72(6), 1478–1492. (doi:10.1177/1747021818796183)
- 22⁶⁷³ 28. Ravignani A, Kello CT, de Reus K, Kotz SA, Dalla Bella S, Méndez-Aróstegui M, Rapado-23 674 Tamarit B, Rubio-Garcia A, de Boer B. 2019 Ontogeny of vocal rhythms in harbour seal pups: an
- 24 675 exploratory study. Curr Zool. 65(1), 107-120. (doi: 10.1093/cz/zoy055)
- 25 676 29. Patel AD. 2006 Musical rhythm, linguistic rhythm, and human evolution. Music Percept 24,99-26 677 104.
- ²⁷ 678 ²⁸ 679 ²⁹ 679 30. Patel AD. 2014 The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol* 12, 1001821. (doi:10.1371/journal.pbio.1001821)
- ₃₀ 680 31. Keehn RJJ, Iversen JR, Schulz I, Patel AD. 2019 Spontaneity and diversity of movement to music 31 681 are not uniquely human. Curr. Biol. 29(13), R621-R622. (doi: 10.1016/j.cub.2019.05.035)
- 32 682 32. Schusterman RJ. 1977 Temporal patterning in sea lion barking (Zalophus californianus). 33 683 Behavioral biology. 20(3), 404-408. (doi:10.1016/S0091-6773(77)90964-6)
- ³⁴ 684 33. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, Phoca ³⁵ 685 36 685 vitulina. Canadian Journal of Zoology. 63(5), 1050-1056. (doi:10.1139/z85-157)
- 37 686 34. Cook P, Rouse A, Wilson M, Reichmuth C. 2013 A California sea lion (Zalophus californianus)
- ₃₈ 687 can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. J. Comp. 39 688 *Psychol.* **127(4)**, 412-27. (doi:10.1037/a0032345)
- 40 689 35. Torres Borda L, Jadoul Y, Rasilo H, Salazar Casals A, Ravignani A. 2021 Vocal plasticity in 41 690 harbour seal pups. Phil. Trans. R. Soc. B 376, 20200456. (doi:10.1098/rstb.2020.0456)
- ⁴² 691 36. Sauvé CC, Beauplet G, Hammill MO, Charrier I. 2015 Acoustic analysis of airborne, underwater,
- $43_{44}^{692}_{692}$ and amphibious mother attraction calls by wild harbor seal pups (Phoca vitulina). Journal of 45⁴⁹³ Mammalogy 96(3), 591–602. (doi:10.1093/jmammal/gyv064)
- 46 694 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. 47 695 48 696 (doi:10.1139/z88-138)
- 49 697 38. Khan CB, Markowitz H, McCowan B. 2006 Vocal development in captive harbor seal pups, Phoca ⁵⁰ 698 vitulina richardii: age, sex, and individual differences. J. Acoust. Soc. Am. 120(3), 1684-94.
- (doi:10.1121/1.2226530)
- 51 699 52 699 53 700 39. Chow CP, Mitchell JF, Miller CT. 2015 Vocal turn-taking in a non-human primate is learned 54 701 during ontogeny. Proc. R. Soc. B 282:20150069. (doi:10.1098/rspb.2015.0069)
- 55 702 40. Landler L, Ruxton GD, Malkemper EP. 2018 Circular data in biology: advice for effectively
- 56 703 implementing statistical procedures. Behav. Ecol. Sociobiol. 72(8), 1-10. (doi:10.1007/s00265-018-57 704 2538-y)
- ⁵⁸ 705 41. Ruxton GD. 2017 Testing for departure from uniformity and estimating mean direction for circular ⁵⁹ 706 60 707 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.

- 2 708 43. Zar JH. 1999 Circular distributions: hypothesis testing. *Biostatistical analysis*, 624–668.
- 3 709 44. Landler L, Ruxton GD, Malkemper EP. 2021 Advice on comparing two independent samples of
- 4 710 circular data in biology. *Scientific reports*. **11(1)**, 1-10. (doi:10.1038/s41598-021-99299-5) 5
- 711 45. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral 6 methods. Econometrica. 37, 424-438. (doi: 10.2307/1912791) 712 7
- 8 713 46. Ravignani A, Norton P. 2017 Measuring rhythmic complexity: A primer to quantify and compare
- 9 714 temporal structure in speech, movement, and animal vocalizations. Journal of Language Evolution 10715 4–19. (doi:10.1093/jole/lzx002)
- 11716 47. Zeileis A, Hothorn T. 2002 Diagnostic Checking in Regression Relationships. R News, 2(3), 7-
- 10. (https://CRAN.R-project.org/doc/Rnews/)
- $\begin{array}{r}
 12 \\
 12 \\
 717 \\
 13 \\
 718 \\
 14 \\
 718 \\
 718 \\
 14 \\
 718 \\
 718 \\
 14 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\
 718 \\$ 48. Keller PE, Novembre G, Hove MJ. 2014 Rhythm in joint action: Psychological and 15⁷719 neurophysiological mechanisms for real-time interpersonal coordination. Philos. Trans. R. Soc. Lond. *B Biol. Sci.* **369(1658)**, 20130394. (doi:10.1098/rstb.2013.0394) 16720
- 49. Keller PE, Novembre G, Loehr J. 2016 Musical ensemble performance: Representing self, other 17721 18 722 and joint action outcomes. In S. S. Obhi & E. S. Cross (Eds.), Shared representations: Sensorimotor ¹⁹723 foundations of social life. pp. 280–310.
- ²⁰ 724 Cambridge: Cambridge University Press. (doi:10.1017/CBO9781107279353.015) 21
- 22 725 50. Mills PF, van der Steen MC, Schultz BG, Keller PE. 2015 Individual differences in temporal
- 23 726 anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception* 3(1-2),
- 24 7 27 13-31. (doi:10.1163/22134468-03002040)
- 25 7 28 51. van der Steen MC, Schwartze M, Kotz SA, Keller PE. 2015 Modeling effects of cerebellar and 26 7 29 basal ganglia lesions on adaptation and anticipation during sensorimotor synchronization. Annals of ²⁷730 the New York Academy of Sciences, **1337**, 101–110. (doi:10.1111/nvas.12628)
- ²⁸731 29731 52. Lemasson A, Pereira H, Levréro F. 2018 Social basis of vocal interactions in western lowland ²₃₀732 gorillas (Gorilla g. gorilla). J. Comp. Psychol. 132(2), 141–151. (doi: 10.1037/com0000105)
- 31 733 53. Yang XJ, Ma XR, Slabbekoorn H. 2014 Timing vocal behaviour: experimental evidence for song 32 7 34 avoidance in Eurasian wrens. Behav Processes 84-90. overlap 103. 33 7 3 5 (doi:10.1016/j.beproc.2013.11.011)
- ³⁴736 54. Song J, Sun R, Fang K, Zhang B, Fang G. 2020 Flexibility as a strategy for avoiding call overlap ³⁵737 Treefrogs. male Anhui Asian Herpetological Research 11(3), 230-239. in 36 737 37 738 (doi:10.16373/j.cnki.ahr.1900331)
- ₃₈739 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) 39 740
- 40 7 4 1 56. Verga L, Sroka MG, Varola M, Villanueva S, Ravignani, A. 2022 Spontaneous rhythm 41 7 4 2 discrimination in a mammalian vocal learner. Biology Letters, 18(10), 20220316.
- ⁴²743 57. Deecke VB, Slater PJ, Ford JK. 2002 Selective habituation shapes acoustic predator recognition in 43 44 44 45 745 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
- 46 746 2-, 4-, and 6-Month-Old Infants. *Infancy*, **4(1)**, 111–140. (doi:10.1207/S15327078IN0401 6)
- 47 747 59. Henning A, Striano T. 2011 Infant and Maternal Sensitivity to Interpersonal Timing. Child 48 7 48 Development, 82(3), 916–931. (doi:10.1111/j.1467-8624.2010.01574.x)
- 49749 60. Malloch S, Trevarthen C. 2018 The Human Nature of Music. Frontiers in Psychology, 9(1680). ⁵⁰750 (doi:10.3389/fpsyg.2018.01680)
- 51 52 53 751 53 752 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)
- 54 753 62. Stormark KM, Braarud HC. 2004 Infants' sensitivity to social contingency: a "double video"
- 55754 study of face-to-face communication between 2- and 4-month-olds and their mothers. Infant Behavior 56755 and Development, 27(2), 195-203. (doi:10.1016/j.infbeh.2003.09.004)
- 57 7 56 63. Fairhurst MT, Janata P, Keller PE. 2013 Being and feeling in sync with an adaptive virtual partner:
- ⁵⁸757 Brain mechanisms underlying dynamic cooperativity. Cerebral Cortex, 23, 2592-2600.
- ⁵⁹758 (doi:10.1093/cercor/bhs243)

- 2 759 64. Fairhurst MT, Janata P, Keller PE. 2014 Leading the follower: an fMRI investigation of dynamic
- ³ 760 cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner.
- ⁴ 761 *Neuroimage*, **84**, 688–697. (doi:10.1016/j.neuroimage.2013.09.027)
- ⁵ 762
 ⁶ 763
 ⁷ 763
 ⁶ 763
- 8 764 66. Harry B, Keller PE. 2019 Tutorial and simulations with ADAM: an adaptation and anticipation 9 765 model of sensorimotor synchronization. *Biological Cybernetics*, **113(4)**, 397–421.
- ¹⁰766 (doi:10.1007/s00422-019-00798-6)
- ¹¹767 67. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization

for Review Only

http://mc.manuscriptcentral.com/issue-ptrsb

- $\frac{12}{13}$ 768 to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. (doi:10.1016/j.cub.2009.03.038)

1 2 3 4	1 2	Measuring rhythms of vocal interactions: A proof of principle in harbour seal pups
5 6 7 8 9	3 4 5 6	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#}
10 11 12	7 8	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
12 13 14	9 10 11	Alesund, 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,
15	12	Oldenburg, Germany.
16 17	13	⁶ Donders Institute for Brain Cognition and Behaviour Radboud University Niimegen Netherlands
18	15	⁷ Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.
19	16	⁸ Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands.
20	17	⁹ MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Australia.
21	18	Aarhus/Aalborg Denmark
22 23	20	[#] Joint senior authors
24	21	
25	22	
26	23	Correspondence:
27	24	Marianna Anichini
29	25	e-mail: marianna.anichini@mpi.ni
30	20 27	e-mail: andrea ravignani@mni nl
31	28	
32	29	
33 34	30	ORCID ID: MA, 0000-0002-2697-0297; KdR, 0000-0002-6129-9726; TAH, 0000-0002-7891-
35	31	596X; DV, 0000-0001-6086-5135; ASC, 0000-0002-3722-5436; CB, 0000-0003-4918-5846; PK,
36	32	0000-0001-7579-6515; AR, 0000-0002-1058-0024
37	33	
38 39	34	Abstract
40	35	Rhythmic patterns in interactive contexts characterise human behaviours such as conversational turn-
41	36	taking. These timed patterns are also present in other animals, and often described as rhythm.
42	31 20	Understanding line-grained temporal adjustments in interaction requires complementary quantitative
43 44	30 30	quantified using a multimethod approach. We record vocal interactions in harbour seal pupe (<i>Phoca</i>
45	<i>4</i> 0	vituling) under controlled conditions. We analyse these data by combining analytical approaches
46	40 41	namely categorical rhythm analysis circular statistics and time series analyses. We test whether pups'
47	42	vocal rhythmicity varies across behavioural contexts depending on the absence or presence of a
48	43	calling partner. Four research questions illustrate which analytical approaches are complementary vs.
49 50	44	orthogonal. For our data, circular statistics and categorical rhythms suggest that a calling partner
51	45	affects a pup's call timing. Granger causality suggests that pups predictively adjust their call timing
52	46	when interacting with a real partner. Lastly, the ADaptation and Anticipation Model (ADAM)
53	47	estimates statistical parameters for a potential mechanism of temporal adaptation and anticipation.
54	48	Our analytical complementary approach constitutes a proof of concept; it shows feasibility in
55 56	49	applying typically unrelated techniques to seals to quantify vocal rhythmic interactivity across
50	50	1. how see 1. so where the

- 57 50 behavioural contexts.
- 58 51

59 52 Keywords: Behavioural interaction, asynchrony, circular statistics, categorical rhythms, time series
 60 53 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 malefemale 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

2

3

4 5

6 7

8 9

10

11 12

13 14

15

16 17

18 19

20

22

27

29 ²⁰₃₀106

34

41

46

48

51

53

music, which is often produced in an interactive context [23]. Similar rhythms are also present in non-90 human animal songs (e.g., thrush nightingales [23], indris [24], but whether such rhythms characterise 91 92 other interactive non-song vocalisations, such as animal calls, is unknown [23,24]. Roeske and 93 colleagues [23] hypothesised that categorical rhythms play a role in calls produced to attract and hold 94 conspecific attention, by making sequences of vocalisations more predictable to listeners. Categorical 95 rhythm analyses could therefore be an interesting method to test the predictability of vocal sequences 96 in non-human animal interactions. Lastly, the ADaptation and Anticipation Model (ADAM), 97 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 98 99 interaction [26,27]. Although ADAM is designed for 'simultaneous chorusing', it could also be used 100 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 21 101 23 102 causality, circular statistics, categorical rhythm analysis, and ADAM—can be applied to a new animal ²⁴ 103 model: the harbour seal (Phoca vitulina). To illustrate the utility and compatibility of these different 26 104 analytical methods, we show how they can be used to better understand the rhythmic communication 28 105 of a small sample of harbour seal pups in different behavioural contexts.

³¹ 107 32 1.3. Our animal model

33 108 The "vocal learning-beat perception and synchronisation" (VL-BPS) hypothesis states that only 35 109 vocal learning species—those capable of producing new vocalisations or modifying existing ones ³⁶ 37 110 based on auditory experience-may possess advanced rhythmic abilities [29,30]. This hypothesis is ³⁸ 111 39 inherently cross-modal: it suggests a strong link between audition and timed movement. For example, 40 112 Snowball, a sulphur-crested cockatoo (Cacatua galerita eleonora), was shown to perceive auditory 42 113 rhythms at different tempi and to predictively synchronise his body movements to them [31]. Parrots 43 44 114 are phylogenetically distant from humans and, among mammals, pinnipeds (seals, sea lions, and 45 115 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 47 116 49 117 extract a beat from periodic acoustic stimuli and entrain to it in a predictive and adaptive manner-⁵⁰ 118 since some species showed vocal mimicry and plasticity [32,33] and others can keep a beat [34]. 52 1 1 9 These characteristics, paralleling human abilities, make pinnipeds an ideal animal clade for 54 120 comparative research on the origins of rhythmic communicative behaviour.

⁵⁵ 121 Harbour seals exhibit both vocal flexibility [33,35] and rhythmic interactivity [21], and are 57 122 particularly vocal in the first few weeks of life [36]. During the lactation period, harbour seal pups 58 59 1 2 3 emit 'mother attraction calls' (hereafter 'calls') to draw their mothers' attention [37]. Mothers are 60 124 silent and use the individual vocal signatures in these calls to recognize their pups [36,38]. Against

125 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 126 127 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. 128 129 Unlike cooperative types of turn-taking (e.g., in humans and in common marmosets (Callithrix 11 130 *jacchus*) [39]) harbour seal pups' interactions are a by-product of neighbouring pups vocalising to 12 131 attract their silent mothers, and are thus probably competitive.

14 1 3 2 To date, only two papers studied vocal rhythms in harbour seals, crucially both focusing on 16 133 single individuals [21,28]. The first study was a playback experiment in which a pup vocally ¦' 134 interacted with sounds broadcasted from a loudspeaker [21]. The pup adjusted the timing of its calls ¹⁹ 135 in an asynchronous manner by responding to the broadcasted conspecific calls with a non-uniformly distributed response phase whose mean approximated 90° [21]. The second study looked at the 21 1 36 23¹³⁷ presence and development of vocal rhythms in three harbour seal pups [28]. Complementary ²⁴ 138 analytical approaches showed how the pups' individual calling patterns gained more rhythmic 26 139 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 28 140 30141 response of the tested animals).

1.4. Aims & research questions

1

2 3

4 5

6 7

8 9

10

13

15

17

20

22

27

29

34

39

41

46

48

51

53

58

60

35 144 In this work, we show how vocal interactive rhythmicity in non-human animals can be ³⁶ 37 145 quantified using a multi-method approach spanning various research domains (e.g., temporal, social, ³⁸ 146 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 40 147 42 148 (Table 1). While our sample sizes are too small to enable species-wide inferences, they are sufficient 43 44 149 to illustrate how methods typically used to study human communication can be adopted to study 45 150 interactivity in animal communication. The goal of this paper is thus to outline a quantitative roadmap 47 151 that future research can follow. Circular statistics and categorical rhythm analysis are used to address 49 152 the first question about temporal adjustment in interaction: "does the presence of a calling partner ⁵⁰ 153 affect the call timing of individual pups?" (Q1). The next two questions consider the effect of 52 1 54 behavioural context on temporal adjustment: "does the type of calling partner (real or broadcasted) 54 155 affect the call timing of individual pups?" (Q2) and "when the focal pup is vocalising, does the ⁵⁵ 156 presence of a silent partner (vs. no partner) affect call timing?" (Q3). We answer these questions ⁵⁷ 157 using circular statistics (Q2 and Q3) and Granger causality tests (Q2). Lastly, ADAM is used to 59 1 58 investigate the fourth question about the cognitive processes involved in temporal adjustment: "which 159 timing mechanisms are used by pups during vocal interactions?" (Q4).

60

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) 1way 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, interonset 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, 59 193 multiplied by 360, resulting in a unit vector with an angle on a circle. A vocal interaction was defined 194 as a group of three calls within the same bout, which includes two calls from the broadcasted/real 195 partner followed by the response of the focal individual (Figure 1). Following previous methodology

196 [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

1 2

3

4 5

6 7

8 9

17 18

₃₅ 202

₃₆ 203

37 204

38 205

39 206

43

46

48

53

55

197

43 44 210 Circular statistics, wherein periodic measures are converted to angles on a circle and compared ⁴⁵ 211 to distributions of interest [40], were used to investigate rhythmic periodicities in pup call response 47 212 phases (Method S4). Following [21], we considered the values of the response phases as circular data 49 213 falling between 0° and 360°. We obtained the circular mean (μ) (i.e., the average direction of the ⁵⁰ 214 response phases calculated from the pup calls; Table S3). Then, we ran Rayleigh z-tests to investigate ⁵²215 whether the distribution of response phases was uniform (e.g., arousal hypothesis) or showed a 54 216 unimodal peak (Table S4) [21]. Subsequently, we tested for uniformity against a specified mean 55 56 217 direction for the unimodal peak using a V-test [40–43].

⁵⁷218 Our data met the assumptions for circular statistics. We tested if the response phases in all four 58 59219 behavioural contexts followed a von Mises distribution using one-sample Watson tests (Table S5). 60 220 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 U² 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 twosample U² 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-255 integer bin counts for each SIR were normalised by bin size and compared using a paired Wilcoxon

256 signed-rank test (a non-parametric test that allows for non-normality in the population data and 2 257 assumes paired differences are continuous, symmetrically distributed, and mutually independent). 258 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 259 260 individual pups differed across behavioural contexts.

1

3

4 5

6 7

8 9

10 11 261 The Granger causality test investigates whether the values of a time series A are better ¹² 262 predicted when considering the values from a second time series B, as opposed to only using values 13 14 263 from time series A [45]. Here, we assessed whether the call timing of a pup partaking in a 1-way or 15 16 264 2-way vocal interaction can be predicted using the call timing of its partner. More specifically, to $^{17}_{18}265$ investigate whether the call timing of a pup differed in relation to the type of partner, we tested ¹⁹266 whether there is a difference in predicting the time series of the pup interacting with a broadcasted 20 21 267 partner vs. time series of the pups interacting with a real partner. Previous work showed that 22 ²²₂₃268 individuals respond to conspecific calls with a non-random pattern [18-20,46]. We therefore expected ²⁴ 269 that the time series of a pup can be better predicted considering the time series of a vocalising partner 26 270 rather than those of a broadcasted signal. In both the 1-way and 2-way interactive scenarios, we 27 28 271 considered Granger causality at two levels: (1) the entire recording, regardless of the length of the 29 30 272 pauses between consecutive calls, and (2) different bouts within each recorded session. We restricted ³¹ 273 the analysis on the different bouts to call sequences that were long enough to generate accurate 33 274 estimates (i.e., a minimum of five paired calls [47]). The bouts included in this analysis range from 5 34 35 275 to 20 calls. We conducted the Granger causality test using call onsets and different lag measures, from ³⁶ 37 276 one to five (Method S6), testing whether the previous one to five onsets in the first time series can be ³⁸ 277 used to better predict the second time series (Figures S5/S6, Table S11). For the 1-way interaction, 39 40 278 we performed a one-way analysis, considering whether the pup's timing could be predicted using the 41 41 42 279 playback timing. For the 2-way interaction, we performed a two-way analysis to assess whether the 43 44 280 two interacting pups influenced the timing of each other's calls.

45 281 We used **ADAM** to test for evidence of reactive error correction and predictive processes in the 46 47 282 1-way and 2-way interactive scenarios. ADAM consists of three computational modules that interact 48 49 283 via internal models of 'self' and 'other' that support one's own action planning and external event ⁵⁰ 284 prediction, respectively (Figure S2). The adaptation module compensates for synchronisation errors 51 52 285 by implementing error correction processes that alter the phase and/or period of an internal timekeeper 53 54 286 controlling for action (here, call) timing. These error correction processes determine the provisional ⁵⁵ 287 timing of the next planned action by providing input to an internal model of the 'self'. The 57 288 anticipation module computes the expected timing of upcoming events based on the weighted sum 58 59 289 of two processes: the linear extrapolation of previous IOIs in the sequence and the copying (or ⁶⁰ 290 'tracking') of the previous IOI, with the output informing temporal predictions generated by the

26 305 27

³⁶ 313 37 314

291 'other' internal model. Finally, a *joint module* integrates and compares the output of the adaptation 2 3 292 and anticipation modules and compensates for discrepancies by implementing an anticipatory error 4 5 293 correction process before the next motor command is issued. The joint module thus reduces potential 6 7 294 temporal mismatches between action plans in 'self' internal models and temporal predictions in 8 9 295 'other' internal models, thereby regulating the balance between the integration (merging) and 10 11 296 segregation (distinction) of information about 'self' and 'other' [48,49]. Each process instantiated in ¹² 297 ADAM is controlled by an independent parameter, and the value of these parameters can be estimated 13 14 298 for a particular individual by fitting the model to behavioural time series data [26,27,50,51]. 15 16 299 Parameter estimates were obtained for both the adaptation-only version of ADAM—which includes 18 300 17 phase correction and period correction-and the full ('joint') version-including period correction, ¹⁹ 301 temporal prediction/tracking, and anticipatory error correction. Both versions of ADAM were applied 20 to each interactive context because it is not possible to know a priori whether the pups' call sequences 21 302 22 ₂₃ 303 (real or broadcasted) have a steady base tempo (for which adaptation is sufficient) or a systematically ²⁴ 304 changing tempo (which benefits from both anticipation and adaptation) (Method S7).

28 306 Table 1. Summary table showing in order: research questions, analyses, contributing pups, predictions, 29 307
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	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
calling partner affect the call timing of individual			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
рирз.			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.

- 57 58
- 59 60

	1	r			1
	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.	Ν	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	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.
pups?	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.

http://mc.manuscriptcentral.com/issue-ptrsb

23

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

²⁰₂₁ 316 3. Results

²² 317 3.1. Does the presence of a calling partner affect the call timing of individual pups? (Q1)

24 3 18 All pups' data entered the circular statistics analysis. Running the Rayleigh test, we found that 25 26 319 the response phase distribution was uniform for pup I which was recorded alone (z = 0.04, p = 0.254; ²⁷₂₈ 320 Figure 2A), whereas it was non-uniform for pups A-H which were recorded with a silent partner (z =29 321 0.11, p < 0.001; Figure 2A; Table 1). This non-uniformity may have been driven by the individual 30 31 322 contributions of pups B and C, which had non-uniformly distributed response phases (Table S4), ³² 33 323 whereas the other six pups had a uniform distribution. The Rayleigh tests run anew in the interactive ³⁴ 324 35 contexts, showed that the response phase distributions of pup calls were unimodal in both the 1-way 36 3 2 5 (pup I: z = 0.39, p < 0.001) and 2-way (pups A - H: z = 0.41, p < 0.001) interactions (Figure 2B; Table 37 1). Applying the V-test in both contexts, the direction of the response phases did not statistically match 38 326 ³⁹ 327 0° (1-way: z = -0.02, p = 0.587; 2-way: z = 0.06, p = 0.110), suggesting that pups did not synchronise ⁴¹ 328 with their partner (real or broadcasted). However, the response phase direction did match 90° (1-way: 42 43 329 z = 0.38, 2-way: z = 0.41, p < 0.001), supporting the previously reported evidence of asynchronous 44 45 330 calling behaviour [21]. A Watson's two-sample U² test confirmed that the response phase distributions 46 47 331 significantly differed between the alone vs. 1-way interaction context for pup I (U² 1.76, p < 0.001), 48 332 and between the silent partner vs. 2-way interaction context for pups A-H (U² = 0.78, p < 0.001, 49 50 333 Figure S3; Table 1). Finally, the circular standard deviation values were higher for the alone and silent 51 52 334 partner contexts compared to both interactive vocal contexts, indicating a larger dispersion of the ⁵³ 335 54 response phases for the former conditions. This outcome is also confirmed by the values for the mean 55 336 resultant length. 56

- 57
- 58 59
- 60



Figure 2. Circular histogram plots (bin width $= 20^{\circ}$) 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 (1way/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).

26 27 344 In the categorical rhythm analyses, the empirical ratio distribution was significantly different 28 3 4 5 from chance when pup I was recorded alone (Figure 3A) and during the 1-way interaction (Figure 29 30 3 4 6 3C). In both contexts, there was no evidence of significant peaks at any of the tested ratios (Table 1). 31 31 32 347 When the playback calls were disregarded from IOI calculations (Figure 3B), there was no significant ³³.348 difference in empirical and simulated ratio distributions for pup I. However, pairwise KS tests showed 34 35 3 49 that the ratio distributions significantly differed when comparing each of the three behavioural 36 contexts (alone vs. 1-way interaction disregarding playback vs. 1-way interaction when pup I 37 350 38 39 351 responds) to each other (Table S10; Table 1). For the five pups that were well-sampled in the silent 40 352 partner context (pups A, B, C, E, and H; Table S8), the empirical ratio distributions did not 41 42 3 5 3 significantly differ from chance (Table S9, Figures 3/S4, Method S4). Finally, when considering 2-43 44 354 way interactions, only pups A and B were well-sampled enough (i.e., had at least ten ratios for both 45 355 the silent partner and 2-way interaction contexts) to compare, but the empirical and simulated ratio 46 47 356 distributions were significantly different only when pup B was the responder (Figure 3; Table S8; 48 49 357 Table S9). Once again, there were no significant peaks at any of the tested ratios for pup B. For both 50 50 51 358 pups A and B, the empirical ratio distributions significantly differed when comparing different ⁵² 359 behavioural contexts (Table S10). Collectively, there was thus little evidence of SIR rhythmic 53 54 360 categories in pup calls, but IOI ratios did significantly differ when looking at the same individuals in 55 56 361 different behavioural contexts (Table 1). 57

57 58 362 Thus, the results from both analyses suggest that the presence of a calling partner does affect ⁵⁹ 363 the call timing of the focal pup, in terms of both call response phases and IOI ratios.

364

60

1 2

3 4

5

6

7

8

9

10

11

16

17 18

20 339

21 340

3

4 5

6

7

8 9

11

35

37

40

42

45



23 366 Figure 3. IOI ratio density plots for pups in different behavioural contexts. Pup I: (A) alone, (B) 1-way 24 367 interaction (disregarding playback), (C) 1-way interaction (responding to playback). Pup A: (D) silent partner, 25 368 (E) 2-way interaction (responding to partner). Pup B: (F) silent partner, (G) 2-way interaction (responding to ²⁶ 369 partner). For each plot, the dashed vertical lines indicate, from left to right, 1:4, 1:3, 1:2, 1:1 (i.e., isochrony), 20 369 27 370 28 371 29 372 30 373 31 373 32 374 33 375 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.

34 376 3.2. Does the type of calling partner (real or broadcasted) affect the call timing of individual 36 377 **pups? (Q2)**

37 378 38 To address this research question, we once again used circular statistics and the response phase 39 379 distributions of all nine pups. Specifically, we compared calls from pup I during the playback (1-way 41 380 interaction) with calls of pups A-H when their partner was also calling (2-way interaction). A Watson's 43 381 two-sample U² test statistically confirmed that the response phase distributions did not differ between ⁴⁴ 382 the 1-way and 2-way vocal interactions ($U^2 = 0.07$, p > 0.10; Table 1).

46 383 Interestingly, however, the Granger causality results from seven pups (A, B, C, D, E, F, I) 47 48 384 showed that call timing behaviour differed depending on the type of partner (Table 1). For the 1-way ⁴⁹ 385 interaction, five different playback sessions featuring pup I were considered (ranging from 34 to 121 ⁵¹ 386 paired calls) and the timing of the pup's calls were never significantly predicted by the timing of the 52 53 387 playback (Table S11, Figure S5). For the 2-way interaction, five different recording sessions were 54 55 388 considered (ranging from 8 to 71 paired onsets). We found that in two sessions, the timing of the first ⁵⁶ 389 57 pup did not significantly predict the timing of the second, in either direction (pair A>B and B>A; pair 58 390 E>F and F>E; Table S11, Figure S6B/S6D). However, in two other sessions, the timing of a pup was 59 significantly predicted by the calling partner in both directions, across different lag values (A>B: lag-60 391

2 392 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 393 S6B/S6C). In the last session, the timing of pup A could be predicted by that of pup B, while we found 394 no indication of temporal adjustment for pup B. At the bout level, we restricted our analyses to four 395 series of paired onsets featuring pups A and B (as they were the only pup pair that met the sample 396 size requirements). In this last scenario, only the timing of pup B was significantly influenced by pup 10 11 397 A (Figure S6A).

1

3

4 5

6 7

8 9

13

15

20

22

27

29

 $^{17}_{18}401$

¹² 398 While the circular statistics results thus suggest that the type of calling partner (real vs. 14 399 broadcasted) does not affect focal pup call timing, the Granger causality results suggest that in certain 16 400 2-way (but not 1-way) interactions, focal pup call timing can be predicted by the partner's call timing.

¹⁹ 402 3.3. When the focal pup is vocalising, does the presence of a silent partner (vs. no partner) 21 4 0 3 affect call timing? (Q3)

23 404 Using circular statistics, we compared the calls of pup I recorded alone with those of the 8 other ²⁴ 405 pups (A-H) recorded with a silent partner present. Watson's two-sample U² test results show that the 26 4 0 6 response phase distributions did not statistically differ between the two contexts ($U^2 = 0.10, p > 0.05$). 28 407 In other words, having a silent partner was essentially the same as having no partner—in terms of the 29 30 408 effect on focal pup response phase distributions-for the pups in our study. This is intuitive, given ³¹ 409 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? (O4)

³⁶ 37 412 ADAM parameter estimation was conducted on call data from all five playback sessions ³⁸ 413 featuring pup I (1-way, Figure 4 A/B), but was restricted to pups A and B for the 2-way interactions 39 40 4 1 4 due to sample size limitations (Figure 4 C/D). Interactive vocal bouts were concatenated to obtain a 41 42 415 time series length which would provide reliable ADAM parameter estimates. A simulation test then 43 44 416 ensured that the estimates were not compromised by differing sequence lengths or by the 45 417 concatenation procedure (Method S7). The reliability of observed parameter estimates was tested by 46 47 418 comparison against corresponding values for randomly permuted data (Method S7). The quality of 48 49</sub>419 the fits to the data did not differ significantly between versions of ADAM (Method S7). Results for ⁵⁰ 420 the 1-way interaction featuring pup I were remarkable with regards to typical human data (e.g. 51 [26,27]) as most significant parameters were negative in sign (Figure 4A/B, Tables S13/S14). 52 4 2 1 53 54 422 Negative phase and period correction estimates indicate that calling earlier will lead to a shortening ⁵⁵ 423 of the next IOI, while calling later will lead to a lengthening of the next IOI. Negative 57 424 prediction/tracking estimates mean that when the playback's IOIs increase (i.e., deceleration), the 58 pup's IOIs will decrease (i.e., acceleration), and vice versa. This systematically enhances the timing 59 4 2 5 60 426 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 11 432 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. 14 4 3 4 Results for this type of vocal interaction showed evidence for temporal adaptation. In particular, 16 4 35 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 21 4 3 8 to a lengthening of the next IOI, whereas calling later will lead to a shortening of the next interval. 23 4 39 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 26 441 mechanism that increased the distinction in calling rates. It should be noted, however, that the 28 4 4 2 observed parameter estimates do not necessarily reflect individual call styles per se, but rather the 30 443 roles that dynamically emerge within the context of this specific pairing of pups.





2 447 Figure 4. ADAM parameter estimates for seal pups A, B, and I. Left panels (A & B) show parameter estimates 3 448 for pup I in five playback sessions (1-way interaction); right panels (C & D) show estimates for pups A and B 4 449 (2-way interaction). Top panels (A & C) show phase correction and period estimates obtained with the 450 adaptation-only version of ADAM; bottom panels (B & D) show period correction and prediction-tracking 451 estimates for the full 'joint' model (anticipatory error correction estimates are not shown). Parameter estimates 452 are shown separately for pup A relative to pup B (aquamarine bars) and pup B relative to pup A (blue bars) in 453 the recording session, and for pup I (green bars) relative to playback calls in separate playback sessions (PB1-454 PB5). Corresponding parameter estimates for randomly permuted data from each session are represented by 10 455 box plots, with the central horizontal line indicating the median, the bottom and top edges of the box indicating $11 \\ 12 \\ 456 \\ 13 \\ 457 \\ 13 \\ 450$ 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) 14 458 level. 15 459

$^{16}_{17}460$ 4. Discussion

1

5

6

7

8

9

10

19

21

24

26

31

33

45

50

18 461 This study sought to provide a methodological proof of concept for quantifying vocal 20 462 rhythmicity in non-human animal interactive communication. Particularly, we showed how 22 463 complementary methodologies can be used to test whether the (Q1) presence and (Q2) type of a ²³ 464 vocalising partner, or the (Q3) presence of a silent partner affect patterns in animal communication 25 465 (Table 1). We also investigated which (Q4) underlying cognitive mechanisms potentially play a role 27 466 in vocal interactions (Table 1). The different analytical techniques proved fruitful; combining and 28 29 467 contrasting their results could highlight nuances in rhythmic behaviour. Nonetheless, our sample size ³⁰ 468 is undeniably small, and our opportunistic design (e.g., mixing within- and between-individual 32 469 comparisons) cannot provide strong inference. We believe that our experimental setup, combined 33 34 470 with the approaches we present, can be adapted and expanded in future work to better understand the ³⁵ 471 36 temporal, social, and cognitive processes underlying interactive communication in animals.

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

40 41 474 The prediction of overlap avoidance was supported by circular statistics, which showed that the 42 475 43 distributions of response phase angles were unimodal. When vocally interacting, pups responded at 44 476 approximately one-quarter of the playback/real partner calling period (90°) and showed phase angles 46 477 significantly different from 0°, supporting previous results from one individual (pup I) [21]. Pups 47 48 478 responding to conspecifics seem to time their calls to avoid overlap, consistently de-synchronizing ⁴⁹ 479 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 51 480 52 53 481 outcomes [21]; in our limited sample, pups not only react to playbacks with asynchrony, but also ⁵⁴ 482 respond to conspecific partner calls in an asynchronous manner.

56 4 8 3 Regarding categorical rhythms, the empirical ratio distribution did not significantly differ from 57 58 484 the distribution expected by chance if IOIs were uniformly distributed for most pup/behavioural ⁵⁹ 60 485 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 486

3

4 5

6 7

8 9

41

46

2 487 categorical rhythms at the seven tested small integer ratios. This includes the 4:1 ratio, which we 488 hypothesised would frequently occur when pups were vocally interacting with playback stimuli or 489 partners based on previous work [21]. Importantly, however, the empirical ratio distributions 490 significantly differed across behavioural context conditions for the three pups (A, B, I) for whom such 491 comparisons were possible. When alone or with a silent partner, pup calls generally showed unimodal 10 11 492 ratio distributions centred around isochrony, whereas calls emitted by vocally interacting pups ¹²493 showed a clear right-shifted peak, or a second peak to the right of isochrony, indicating that the pup's 13 14 4 9 4 IOIs during vocal interactions were generally shorter than the playback's/partner's IOIs. This context-15 16 495 dependent asymmetry bolsters results related to Q4, namely that the presence of vocalising partner $^{17}_{18}496$ can significantly impact pup vocal behaviour, with interacting individuals trying to prevent call ¹⁹ 497 overlap. Roeske and colleagues [23] hypothesised that categorical rhythms may be prevalent in calls 20 21 4 98 meant to attract and hold conspecific attention, such as the calls pups use to attract the attention of 22 22 23 499 their mothers. We did not find evidence of rhythm categories at seven tested ratios; this negative result ²⁴ 500 has methodological value, highlighting that not all species that produce attention-seeking 26 501 vocalisations show integer ratio categories. 27

28 502 A vocal interaction with non-adaptive playback stimuli (1-way) may represent a limitation to ²⁹ 30 503 studying spontaneous vocal production. The extent of this limitation can be gauged by comparing ³¹ 504 results to data from vocal interactions with a real partner (2-way), and with the use of complementary 33 505 analytical methodologies, such as categorical rhythm analysis and circular statistics.

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

³⁸ 508 39 Findings from Q2 show contrasting results. Call phases were statistically the same no matter if 40 509 pups could interactively adjust their call timing to each other (2-way) or not (in 1-way). Partly in 42 510 contrast with this, the Granger causality analysis showed how vocally interacting with a real 43 44 511 individual impacted the pup's vocal behaviour more than interacting with a broadcasted partner 45 512 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 47 513 48 49 514 reciprocal adjustment of timing behaviour. In contrast, we found little evidence that a pup adjusts the ⁵⁰ 515 51 time series of its calls to a playback series. Indeed, in most cases, the pup's call timing could not be 52 516 predicted by the playback's call timing.

53 54 517 Together, findings from circular statistics (Q1) and Granger causality (Q2) point towards a ⁵⁵ 518 directionality-overlap avoidance relationship, which has also been suggested for other non-human ⁵⁷ 519 species (e.g., primates [18,52]; birds [53]; amphibians [54]; seals [55]). Interestingly, when infant 58 marmosets interact with their parents, the probability that their vocalisations will overlap with those 59 520 60

of adults decreases over time [22], suggesting that turn-taking in some mammals is a learned vocal 2 521 522 behaviour scaffolded by active parental feedback [39].

524 Q3. A silent partner does not trigger call rhythmicity

1

3

4 5 523

6 7

8 9

10

13

15

21 5 3 2 22 23 533

34

41

46

525 Circular statistics indicated that pups did not show any periodic pattern both, either when calling 11 526 alone or with a silent partner present (Q1). This is consistent with the fact that wild pups produce calls ¹² 527 to attract a silent mother [36]: if no other calling conspecifics are present, there is no need to adjust 14 528 one's call timing and vocalisations are produced with a random onset. Moreover, given that the 16 529 response phase distributions did not differ between pups recorded with a silent partner or alone, it $^{17}_{18}530$ suggests that the presence of a silent partner did not trigger variable calling behaviour in our study ¹⁹ 531 20 population.

Q4. Purported timing mechanisms for vocal interactions

²⁴ 534 The ADAM analysis suggests that seal pups may perceive temporal patterns [56], which arise 26 535 27 between their calls and those of others, and react to them by adopting different mechanisms for 28 536 temporal adaptation. The negative parameter estimates that we observed in the 1-way interaction 29 30 537 sessions (consistent with enhancing the distinction between calls) could reflect attempts to lead the ³¹ 538 temporal interaction or even to test the responsiveness of the (broadcasted) partner by introducing 33 539 timing perturbations and gauging their effects. The decrease of temporal adaptation and/or 35 540 anticipation observed across playback sessions is consistent with a gradual process of habituation, ³⁶ 37 541 with the pup possibly learning that the playback is non-interactive. Harbour seals are capable of ³⁸ 542 39 acoustic recognition based on habituation paradigms; they can discriminate among different stimuli 40 543 and selectively habituate to them [57]. The lack of temporal adaptation to the playback stimulus we 42 544 observed in a seal pup may entail similar habituation processes. From a comparative perspective, 43 545 44 these findings also suggest a sensitivity to social contingency that may be analogous to capacities in 45 546 human infants, who become disinterested and display fewer signs of positive affect during vocal 47 547 interactions with non-responsive or delayed video recordings of their mothers [58-62]. Future studies 48 49 548 could address the role of temporal contingency by using interactive playback sequences [63–65].

⁵⁰ 549 51 Ours constitutes the first attempt to apply ADAM to non-human animals. On the technical side, 52 550 this necessitated the validation of an approach where brief interactive vocal bouts were concatenated 53 54 551 to produce longer time series and thereby reduce the risk of model overfitting (Discussion S1). ⁵⁵₅₆ 552 Demonstrating the utility of this procedure opens the door to applying the model in a wider range of 57 553 behavioural contexts. Nevertheless, caveats are necessary when interpreting the seal data in light of 58 previous work with ADAM in humans, where individuals intentionally produce movements, whose 59 554 60 555 sensory effects occur simultaneously with rhythmically regular sounds (e.g. [26,27]). Assumptions

Page 45 of 55

1

3

4 5

6 7

8 9 560

2 556 about intentionality and simultaneity may not apply to seal pup vocal interactions or rhythmic 557 interactions in other animals [6,9]. Points of convergence in the main outcomes of the complementary 558 analysis techniques suggest that ADAM, like categorical rhythm analysis, circular statistics, and 559 Granger causality, may also be robust and informative under such conditions.

10 11 561 **5.** Conclusions

12 562 13 Crucially, our work highlights the efficacy of combining multiple methods to study rhythmic 14 563 vocal behaviour. Our approaches vary in the degree to which they capture global temporal 15 16 564 characteristics across events vs. local temporal dependencies between events. Global measures (e.g., $^{17}_{18}565$ from circular statistics or categorical rhythm analysis) reveal predominant rhythmic features of a ¹⁹ 566 20 vocal interaction while local time series measures (e.g., from Granger causality or ADAM) provide 21 567 information about how these features might arise. Through this approach we could tease apart rhythm 22 23 568 nuances in our dataset, further develop harbour seals as a model species, and demonstrate how certain ²⁴ 569 analyses often restricted to humans, such as categorical rhythms (but see [23,24]) and ADAM, can be 26 570 27 applied to non-human animals. The categorical rhythm and circular statistical analyses tackled similar 28 571 questions from different angles, namely whether the distributions of IOI ratios (the former) or call ²⁹ 30 572 response phases (the latter) significantly differed across behavioural contexts. The categorical rhythm ³¹ 573 analyses also sought to determine whether the rhythmicity of pup calling behaviour conforms to small 33 574 integer ratios. Our negative result is, to our knowledge, the first published case of a species for which 34 35 575 categorical rhythms are clearly absent from vocalisations, which adds to understanding of how, why, ³⁶ 37 576 and when such rhythms evolve in communication systems. Time series analyses such as Granger ³⁸ 577 39 causality allowed testing for timing adjustment. In the case of ADAM, mechanisms of temporal 40 578 adaptation and anticipation that have previously been used to describe rhythmic behaviour in humans 41 41 42 579 [66] were used to describe rhythmic behaviour in seals. Interestingly, the ADAM model provides an 43 44 580 empirical warning about potential seal pup habituation effects when vocally interacting with a 45 581 recorded partner. 46

47 582 Though our sample size is small, studies on single individuals are not unusual in comparative 48 49 583 research [20,67]. Nevertheless, it is possible that the lack of adjustment to a playback, the adaptation ⁵⁰ 584 51 to a real individual, or both, reflect a peculiar vocal behaviour of the individuals we tested and cannot 52 585 be generalised to the species as a whole. The pups in this study were in a temporarily captive setting, 53 54 586 albeit in acoustic proximity to other individuals, similar to conditions they would experience in ⁵⁵ 587 nature. Unfortunately, vocal development in harbour seal pups has not been studied in wild colonies, 57 588 hence we do not know whether captivity affects their vocal development. We do know, however, that 58 pups vocally interact with neighbouring pups in the colony and not with their silent mothers; hence 59 589 60

2 590 the turn-taking behaviour observed in our captive conditions might extend to the same behaviour in 591 wild conspecifics.

592 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 593 9 594 triggered by individual-specific behaviours or by cues from other modalities. Multimodal 11 595 communication should be the target of future studies since we cannot assume that such interactions ¹² 596 are limited to acoustic cues. Despite these limitations, our study shows that adopting multiple 14 597 complementary approaches can be a fruitful way to study rhythmic interactivity in non-human animal 16 598 communication. 17 18 599

¹⁹ 600 Acknowledgements

1

3

4 5

6 7

8

10

13

15

20

22

27

29

We are grateful to the editor and the two anonymous reviewers who provided insightful 21 601 23 602 comments and suggestions on an earlier version of this manuscript. We thank all the volunteers, seal ²⁴ 603 care members and veterinarians of Sealcentre Pieterburen who helped during the experiment and in 26 604 the rehabilitation process of the seal pups. We are thankful to Chiara De Gregorio for sharing the R 28 605 code used for the categorical rhythm analyses and to Laura Verga for providing helpful comments on 20 606 a draft of this manuscript.

rez oniz

- 2 607 References
- 3 608 1. Dusenbery DB. 1992 Sensory ecology: How organisms acquire and respond to information. 4 609 Freeman, New York. 5
- 610 6 2. Stegmann UE. 2013 Animal Communication Theory: Information and Influence. Cambridge, UK:
- 7 611 Cambridge University Press. (doi:10.1017/CBO9781139003551)
- 8 612 3. Anichini M, de Heer Kloots M, Ravignani A. 2020 Interactive rhythms in the wild, in the brain, 9 and in silico. Can. J. Exp. Psychol. 74(3), 170-175. (doi: 10.1037/cep0000224)
- -10⁶¹³ 4. Ravignani A, Bowling DL, Fitch WT. 2014 Chorusing, synchrony, and the evolutionary functions
- 11 614 12 615 of rhythm. Frontiers in Psychology. 5,1118. (doi:10.3389/fpsyg.2014.01118)
- 13616 5. Greenfield MD, Aihara I, Amichay G, Anichini M, Nityananda V. 2021 Rhythm interaction in
- 14617 animal groups: selective attention in communication networks. Phil. Trans. R. Soc. B 376(1835), ¹⁵ 618 20200338. (doi:10.1098/rstb.2020.0338)
- $16 \\ 619 \\ 17 \\ 620 \\ 18 \\ 620$ 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) 19 621
- 20 622 7. Levinson SC. 2006 Cognition at the heart of human interaction. *Discourse studies*. 8(1), 85-93.
- 21 6 2 3 8. Levinson SC. 2016 The Interaction Engine hypothesis. The Language in Interaction Summerschool 22 624 on Human Language: From Genes and Brains to Behavior. Berg en Dal, The Netherlands.
- ²³ 625 9. Ravignani A, Verga L, Greenfield MD. 2019 Interactive rhythms across species: the evolutionary
- ²⁴ 626 25 627 biology of animal chorusing and turn-taking. Annals of the New York Academy of Sciences, 1453(1),
- ₂₆ 627 12-21. (doi:10.1111/nyas.14230)
- 10. Levinson SC. 2016 Turn-taking in human communication-origins and implications for language 27 628 28 6 29 processing. Trends in cognitive sciences. 20(1), 6-14. (doi.org/10.1016/j.tics.2015.10.010)
- 29 630 11. Pika S, Wilkinson R, Kendrick KH, Vernes SC. 2018 Taking turns: bridging the gap between ³⁰ 631 animal communication. Proc. *R*. Soc. 285(1880), human and B 20180598. $31 \\ 32 \\ 33 \\ 633 \\ 33 \\ 633 \\ 633 \\ 634$ (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) 34 634
- 35 635 13. Schulz TM, Whitehead H, Gero S, Rendell L. 2008 Overlapping and matching of codas in vocal 36 6 36 interactions between sperm whales: insights into communication function. Anim. Behav. 76(6), 1977-
- 37 637 1988. (doi:10.1016/j.anbehav.2008.07.032)
- ³⁸ 638 14. Takahashi DY, Narayanan DZ, Ghazanfar AA. 2013 Coupled Oscillator Dynamics of Vocal Turn-³⁹ 639 40 41 640 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. 42 641 Routledge.
- 43 642 16. Yoshida S, Okanoya K. 2005 Evolution of turn-taking: a bio-cognitive perspective. Cognitive 44 643 Bulletin of Japanese Cognitive Science Studies: the Society. 12(3). 153-165.
- 45 644 (doi:10.11225/jcss.12.153)
- ⁴⁶ 645 17. Henry MJ, Cook PF, de Reus K, Nityananda V, Rouse AA, Kotz SA. 2021 An ecological approach
- 47 48 646 to measuring synchronization abilities across the animal kingdom. Phil. Trans. R. Soc. B. 376(1835), 49⁶⁴⁷ 20200336. (doi:10.1098/rstb.2020.0336)
- 50 648 18. Gamba M, Torti V, Estienne V, Randrianarison RM, Valente D, Rovara P, Bonadonna G, Friard 51 649 O, Giacoma C. 2016. The Indris Have Got Rhythm! Timing and Pitch Variation of a Primate Song
- 52 650 Examined between Sexes and Age Classes. Front. Neurosci. 10, 249. (doi:10.3389/fnins.2016.00249)
- ⁵³ 651 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)
- ⁵⁴ 652 ⁵⁵ 653 56 653 20. Fuhrmann D, Ravignani A, Marshall-Pescini S, Whiten A. 2014 Synchrony and motor mimicking
- ₅₇ 654 in chimpanzee observational learning. Sci. Rep. 4, 1–7. (doi:10.1038/srep05283)
- 58 655 21. Ravignani A. 2019 Timing of antisynchronous calling: A case study in a harbour seal pup (Phoca 59 6 56 vitulina). J Comp Psychol. 133(2), 272–277. (doi:10.1037/com0000160)
- 60

- 2 657 22. Takahashi DY, Fenley AR, Ghazanfar AA. 2016 Early development of turn-taking with parents
- 3 658 shapes vocal acoustics in infant marmoset monkeys. Phil. Trans. R. Soc. B: Biological Sciences.
- 4 659 **371(1693)**, 20150370. (doi:10.1098/rstb.2015.0370) 5
- 23. Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical rhythms are shared 660 6 between songbirds and humans. Curr. Biol. 30, 3544–3555. (doi:10.1016/j.cub.2020.06.072) 661 7
- 24. De Gregorio C, Valente D, Raimondi T, Torti V, Miaretsoa L, Friard O, Giacoma C, Ravignani A, 662 8
- 9 663 Gamba M. 2021 Categorical rhythms in a singing primate. Curr. Biol. 31, R1379-R1380. 10 664 (doi:10.1016/j.cub.2021.09.032)
- 11 665 25. van der Steen MC, Keller PE. 2013 The ADaptation and Anticipation Model (ADAM) of $\frac{12}{13}$ 666 sensorimotor synchronization. **Frontiers** in Human Neuroscience, 7, 253.
- 14 667 (doi:10.3389/fnhum.2013.00253)

- 26. van der Steen MC, Jacoby N, Fairhurst MT, Keller PE. 2015 Sensorimotor synchronization with 15 668 tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. Brain Res. 16 6 6 9 17 670 **1626**, 66–87. (doi:10.1016/j.brainres.2015.01.053)
- 18 671 27. Mills PF, Harry B, Stevens CJ, Knoblich G, Keller PE. 2019 Intentionality of a co-actor influences ¹⁹ 672 sensorimotor synchronisation with a virtual partner. *Quarterly Journal of Experimental Psychology*, ²⁰ 673 21 673 72(6), 1478–1492. (doi:10.1177/1747021818796183)
- 22 674 28. Ravignani A, Kello CT, de Reus K, Kotz SA, Dalla Bella S, Méndez-Aróstegui M, Rapado-23 675 Tamarit B, Rubio-Garcia A, de Boer B. 2019 Ontogeny of vocal rhythms in harbour seal pups: an
- 24 676 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 24,99-25 677 26 678 104.
- ²⁷ 679 ²⁸ 680 ²⁹ 680 30. Patel AD. 2014 The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol* 12, 1001821. (doi:10.1371/journal.pbio.1001821)
- ₃₀ 681 31. Keehn RJJ, Iversen JR, Schulz I, Patel AD. 2019 Spontaneity and diversity of movement to music 31 682 are not uniquely human. Curr. Biol. 29(13), R621-R622. (doi: 10.1016/j.cub.2019.05.035)
- 32 683 32. Schusterman RJ. 1977 Temporal patterning in sea lion barking (Zalophus californianus). 33 684 Behavioral biology. 20(3), 404-408. (doi:10.1016/S0091-6773(77)90964-6)
- ³⁴ 685 33. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, Phoca $35_{36}_{686}_{687}$ vitulina. Canadian Journal of Zoology. 63(5), 1050-1056. (doi:10.1139/z85-157)
- 37 687 34. Cook P, Rouse A, Wilson M, Reichmuth C. 2013 A California sea lion (Zalophus californianus)
- 38 688 can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. J. Comp. 39 689 *Psychol.* **127(4)**, 412-27. (doi:10.1037/a0032345)
- 40 690 35. Torres Borda L, Jadoul Y, Rasilo H, Salazar Casals A, Ravignani A. 2021 Vocal plasticity in 41 691 harbour seal pups. Phil. Trans. R. Soc. B 376, 20200456. (doi:10.1098/rstb.2020.0456)
- ⁴² 692 36. Sauvé CC, Beauplet G, Hammill MO, Charrier I. 2015 Acoustic analysis of airborne, underwater,
- ⁴³ 693 44 693 and amphibious mother attraction calls by wild harbor seal pups (Phoca vitulina). Journal of 45⁶⁹⁴ Mammalogy 96(3), 591–602. (doi:10.1093/jmammal/gyv064)
- ₄₆ 695 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. 47 696 48 697 (doi:10.1139/z88-138)
- 49 698 38. Khan CB, Markowitz H, McCowan B. 2006 Vocal development in captive harbor seal pups, Phoca ⁵⁰ 699 vitulina richardii: age, sex, and individual differences. J. Acoust. Soc. Am. 120(3), 1684-94.
- (doi:10.1121/1.2226530)
- 51 700 52 701 53 701 39. Chow CP, Mitchell JF, Miller CT. 2015 Vocal turn-taking in a non-human primate is learned 54 702 during ontogeny. Proc. R. Soc. B 282:20150069. (doi:10.1098/rspb.2015.0069)
- 55 703 40. Landler L, Ruxton GD, Malkemper EP. 2018 Circular data in biology: advice for effectively
- 56 704 implementing statistical procedures. Behav. Ecol. Sociobiol. 72(8), 1-10. (doi:10.1007/s00265-018-57 705 2538-y)
- ⁵⁸ 706 41. Ruxton GD. 2017 Testing for departure from uniformity and estimating mean direction for circular ⁵⁹ 707 60 707 data. Biol. Lett. 13, 20160756. (doi:10.1098/rsbl.2016.0756)
- 708 42. Pewsey A, Neuhäuser M, Ruxton GD. 2013 Circular statistics in R. Oxford University Press.

- 2 709 43. Zar JH. 1999 Circular distributions: hypothesis testing. *Biostatistical analysis*, 624–668.
- 3 710 44. Landler L, Ruxton GD, Malkemper EP. 2021 Advice on comparing two independent samples of
- 4 711 circular data in biology. *Scientific reports*. **11(1)**, 1-10. (doi:10.1038/s41598-021-99299-5) 5
- 712 45. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral 6 methods. Econometrica. 37, 424-438. (doi: 10.2307/1912791) 713 7
- 8 714 46. Ravignani A, Norton P. 2017 Measuring rhythmic complexity: A primer to quantify and compare
- 9 715 temporal structure in speech, movement, and animal vocalizations. Journal of Language Evolution 10716 4–19. (doi:10.1093/jole/lzx002)
- 11717 47. Zeileis A, Hothorn T. 2002 Diagnostic Checking in Regression Relationships. R News, 2(3), 7-
- ¹²718 10. (https://CRAN.R-project.org/doc/Rnews/)
- 13 719 14 719 48. Keller PE, Novembre G, Hove MJ. 2014 Rhythm in joint action: Psychological and 15 720 neurophysiological mechanisms for real-time interpersonal coordination. Philos. Trans. R. Soc. Lond. *B Biol. Sci.* **369(1658)**, 20130394. (doi:10.1098/rstb.2013.0394) 16721
- 49. Keller PE, Novembre G, Loehr J. 2016 Musical ensemble performance: Representing self, other 17722 18723 and joint action outcomes. In S. S. Obhi & E. S. Cross (Eds.), Shared representations: Sensorimotor ¹⁹724 foundations of social life. pp. 280–310.
- ²⁰ 725 Cambridge: Cambridge University Press. (doi:10.1017/CBO9781107279353.015) 21
- 22 726 50. Mills PF, van der Steen MC, Schultz BG, Keller PE. 2015 Individual differences in temporal
- 23 727 anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception* 3(1-2),
- 24 7 28 13-31. (doi:10.1163/22134468-03002040)
- 25 7 29 51. van der Steen MC, Schwartze M, Kotz SA, Keller PE. 2015 Modeling effects of cerebellar and 26730 basal ganglia lesions on adaptation and anticipation during sensorimotor synchronization. Annals of ²⁷ 731 the New York Academy of Sciences, **1337**, 101–110. (doi:10.1111/nvas.12628)
- ²⁸732 29732 52. Lemasson A, Pereira H, Levréro F. 2018 Social basis of vocal interactions in western lowland ²₃₀733 gorillas (Gorilla g. gorilla). J. Comp. Psychol. 132(2), 141–151. (doi: 10.1037/com0000105)
- 31 7 34 53. Yang XJ, Ma XR, Slabbekoorn H. 2014 Timing vocal behaviour: experimental evidence for song 32 7 35 avoidance in Eurasian wrens. Behav Processes 84-90. overlap 103. 33 7 3 6 (doi:10.1016/j.beproc.2013.11.011)
- ³⁴737 54. Song J, Sun R, Fang K, Zhang B, Fang G. 2020 Flexibility as a strategy for avoiding call overlap ³⁵738 male Anhui Treefrogs. Asian Herpetological Research 11(3), 230-239. in 36 37 37 38 740 (doi:10.16373/j.cnki.ahr.1900331)
- 55. Serrano A, Terhune JM. 2002 Antimasking aspects of harp seal (Pagophilus groenlandicus) 39 741 underwater vocalizations. J. Acoust. Soc. Am. 112(6), 3083-3090. (doi:10.1121/1.1518987)
- 40 7 4 2 56. Verga L, Sroka MG, Varola M, Villanueva S, Ravignani, A. 2022 Spontaneous rhythm 41 7 4 3 discrimination in a mammalian vocal learner. Biology Letters, 18(10), 20220316.
- ⁴² 744 57. Deecke VB, Slater PJ, Ford JK. 2002 Selective habituation shapes acoustic predator recognition in ⁴³ 745 ⁴⁴ 746 ⁴⁵ 746 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
- 46 747 2-, 4-, and 6-Month-Old Infants. *Infancy*, **4(1)**, 111–140. (doi:10.1207/S15327078IN0401 6)
- 47 748 59. Henning A, Striano T. 2011 Infant and Maternal Sensitivity to Interpersonal Timing. Child 48 749 Development, 82(3), 916–931. (doi:10.1111/j.1467-8624.2010.01574.x)
- 49750 60. Malloch S, Trevarthen C. 2018 The Human Nature of Music. Frontiers in Psychology, 9(1680). ⁵⁰751 (doi:10.3389/fpsyg.2018.01680)
- 51 52 52 53 753 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" 54 754
- 55 755 study of face-to-face communication between 2- and 4-month-olds and their mothers. Infant Behavior 56756 and Development, 27(2), 195-203. (doi:10.1016/j.infbeh.2003.09.004)
- 57 7 57 63. Fairhurst MT, Janata P, Keller PE. 2013 Being and feeling in sync with an adaptive virtual partner:
- ⁵⁸758 Brain mechanisms underlying dynamic cooperativity. Cerebral Cortex, 23, 2592-2600.
- ⁵⁹759 60 (doi:10.1093/cercor/bhs243)

- 2 760 64. Fairhurst MT, Janata P, Keller PE. 2014 Leading the follower: an fMRI investigation of dynamic
- $\frac{3}{4}$ 761 cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner.
- ⁴ 762 *Neuroimage*, **84**, 688–697. (doi:10.1016/j.neuroimage.2013.09.027)
- ⁵ 763
 ⁶ 763
 ⁶ 764
 ⁶ 764
- 8 765 66. Harry B, Keller PE. 2019 Tutorial and simulations with ADAM: an adaptation and anticipation 9 766 model of sensorimotor synchronization. *Biological Cybernetics*, **113(4)**, 397–421.
- 10767 (doi:10.1007/s00422-019-00798-6)
- ¹¹768 67. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization

http://mc.manuscriptcentral.com/issue-ptrsb

- $\frac{12}{13}$ 769 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)

0/360

Vocal Interaction

2-way

1-way



413x232mm (118 x 118 DPI)

http://mc.manuscriptcentral.com/issue-ptrsb



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)

http://mc.manuscriptcentral.com/issue-ptrsb



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)



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	Circular statistics	A-I	Pups will not vocalise at random points in time.	Y	Rayleigh test: Unimodal distribution of response phases.
calling partner affect the call timing of individual muns?			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
<i>p</i>			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 A, B, 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).
			Vocally interacting pups will have a significant peak at the 4:1 SIR.	Ν	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.	Υ	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.