

## Vocal sequences in narwhals (*Monodon monoceros*)<sup>a)</sup>

Sam F. Walmsley,<sup>1,b)</sup> Luke Rendell,<sup>1,c)</sup> Nigel E. Hussey,<sup>2</sup> and Marianne Marcoux<sup>3</sup>

<sup>1</sup>Sea Mammal Research Unit, School of Biology, University of St Andrews, Sir Harold Mitchell Building, St Andrews, KY16 9TH, Scotland

<sup>2</sup>Integrative Biology, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, N9B 3P4, Canada

<sup>3</sup>Arctic Aquatic Research Division, Fisheries and Oceans Canada, Winnipeg, Manitoba, R3T 2N6, Canada

### ABSTRACT:

Sequences are indicative of signal complexity in vocal communication. While vocal sequences are well-described in birds and terrestrial mammals, the extent to which marine mammals use them is less well understood. This study documents the first known examples of sequence use in the narwhal (*Monodon monoceros*), a gregarious Arctic cetacean. Eight female narwhals were fitted with animal-borne recording devices, resulting in one of the largest datasets of narwhal acoustic behaviour to date. A combination of visual and quantitative classification procedures was used to test whether subjectively defined vocalization patterns were organized into sequences. Next, acoustic characteristics were analyzed to assess whether sequences could disclose group or individual identity. Finally, generalized linear models were used to investigate the behavioural context under which sequences were produced. Two types of sequences, consisting of “paired” patterns and “burst pulse series,” were identified. Sequences of burst pulse series were typically produced in periods of high vocal activity, whereas the opposite was true for sequences of paired patterns, suggesting different functions for each. These findings extend the set of odontocetes which are known to use vocal sequences. Inquiry into vocal sequences in other understudied marine mammals may provide further insights into the evolution of vocal communication. © 2020 Acoustical Society of America.

<https://doi.org/10.1121/10.0000671>

(Received 16 July 2019; revised 4 January 2020; accepted 13 January 2020; published online 13 February 2020)

[Editor: Rebecca A. Dunlop]

Pages: 1078–1091

### I. INTRODUCTION

Species in many taxa combine vocalizations according to stereotyped organizational principles, producing vocal sequences (Kershenbaum *et al.*, 2016). Sequences can be composed of repetitions of the same unit or combinations of different units (Collier *et al.*, 2014; Kershenbaum *et al.*, 2016) and may be produced by a single individual or multiple individuals, as in the overlapping songs of great tits (*Parus major*; Krebs *et al.*, 1981). Many species’ vocal production capacities are largely innate, meaning they can only use a small repertoire of sounds (Podos, 1996). However, a set of vocal signals combined into structured sequences can contain more information than the same sounds produced in isolation. As such, constraints on vocal production are hypothesized to be an evolutionary driver sequence use, as in Campbell’s monkeys (*Cercopithecus campbelli*; Ouattara *et al.*, 2009). In species with more flexible vocal production, sequences may still contribute to information-rich communication, as is the case for human language. While possible relationships between the use of

sequences and the evolution of language are not resolved (Kershenbaum *et al.*, 2014; Ouattara *et al.*, 2009; Scott-Phillips *et al.*, 2014), it is agreed that vocal sequences are indicative of signal complexity. Even when individual units do not have specific meanings, the “developmental stress hypothesis” states that the complexity of a sequence (e.g., birdsong) can serve as an otherwise arbitrary indicator of mate quality (Nowicki and Searcy, 2004). Thus, without attention to the sequential structure of signals, the complexity and functions of vocal repertoires may be underestimated (Collier *et al.*, 2014; Kershenbaum *et al.*, 2012; Luís *et al.*, 2018).

Toothed whales are important model systems for studying the evolution of vocal communication. For example, bottlenose dolphins [*Tursiops* species (sp)] are the only non-human species for which vocal production learning, functional reference, signal innovation, and the capacity to understand syntax have all been demonstrated (Herman, 2006; Janik, 2013). However, in contrast to the complex sequences described in songbirds, terrestrial mammals, and mysticetes (Payne and McVay, 1971; McDonald *et al.*, 2006; Berwick *et al.*, 2011), most examples of vocal sequences in odontocetes are repetitions of a single call type. Existing work has identified nonrandom transitions between calls of different types in short-finned pilot whales (*Globicephala macrorhynchus*; Sayigh *et al.*, 2013), killer whales (*Orca orcinus*; Saulitis *et al.*, 2005), and bottlenose dolphins (Ferrer-i-Cancho and McCowan, 2012). Bottlenose dolphins

<sup>a)</sup>Portions of this work were presented in “Using biologging technology to study vocal sequences in narwhals” at the Easter meeting of the Association for the Study of Animal Behaviour, University of York, York, England, 2019, and “Rhythmically repeated patterns of pulsed vocalizations in wild narwhals,” MSc dissertation, University of St. Andrews, 2018.

<sup>b)</sup>Electronic mail: sjf.walmsley@gmail.com, ORCID: 0000-0001-8577-6884.

<sup>c)</sup>ORCID: 0000-0002-1121-9142.

are known to produce feeding-associated bray sequences composed of repeated pulsed units, as well as “multi-looped” sequences of signature whistles (Janik, 2000; Luís *et al.*, 2018; Esch *et al.*, 2009). Finally, sperm whales (*Physeter macrocephalus*) are known to produce temporally stereotyped click codas (Watkins and Schevill, 1977). However, the extent to which other odontocete species organize their vocalizations into sequences is poorly understood (Janik, 2009).

Technology has traditionally been a limiting factor in the study of marine mammal communication (Lammers and Oswald, 2015, p. 125), given that most behaviour occurs underwater, and vocalizations are often produced at frequencies beyond the range of human hearing (Mann, 2000). This has resulted in biases for the study of accessible (e.g., coastal) species or species that are readily kept in captivity. It is possible, then, that the apparent lack of vocal sequences in odontocetes is an artifact of methodological limitations. Alternatively, an advanced ability to produce and learn new signals (e.g., Abramson *et al.*, 2018) might allow odontocetes to produce large enough repertoires of singular signals to satisfy their communication needs without recourse to sequences.

Here, we used high-bandwidth animal-borne recording devices to investigate the vocal behaviour of individual narwhals (*Monodon monoceros*), a species for which the nature and function of communication signals remain understudied (Morisaka, 2012; Morisaka *et al.*, 2013). In winter, the narwhal’s environment is characterized by dense ice-cover and darkness (Laidre and Heide-Jørgensen, 2011), making acoustic signals critical for both sensing the environment and social interaction. Narwhals have been recorded producing a range of plausibly social sounds, typically classified as “pulsed calls” composed of clicks with very short inter-click intervals, as well as “tonal calls” or whistles, and “mixed calls,” which include overlaid pulsed and tonal components (Marcoux *et al.*, 2012; Stafford *et al.*, 2012; Shapiro, 2006). Initial exploration of a subset of our recordings revealed two types of possible sequences, one consisting of “paired” patterns of pulsed calls, and the other consisting of sets of short burst pulses, hereafter, termed “burst pulse series.”

To investigate the vocalizations in more depth, we first asked whether they were produced according to stereotyped organizational principles, satisfying the definition of sequences. Without being able to locate an individual using multiple simultaneous recording devices, it is challenging to attribute vocalizations to specific individuals, even when instrumented—for example, an animal swimming alongside a tagged individual might have its vocal production apparatus closer to the tag hydrophone than the tagged individual itself (Sayigh *et al.*, 2013). This makes it difficult to distinguish between sequences produced by single individuals and those produced in call exchanges (Kershenbaum *et al.*, 2014). For this reason, we then investigated frequency and amplitude characteristics of these patterns to understand whether differences in calls recorded from different individuals could be indicative of individual or group-specific signals. Finally we used generalized linear models (GLMs) to

investigate the behavioural context of sequences in order to make inferences about their function.

## II. METHODS

### A. Study system and data collection

Narwhals were captured and equipped with instrument packages that recorded acoustic and depth data (Acousonde™ Model 3B tags, Santa Barbara, California, USA) in Tremblay Sound, Nunavut (72° 28′ 49.5588″ N, 80° 54′ 18.2592″ W) between August 13 and September 11, 2017. Tags were attached adjacent to the left side of the dorsal ridge by suction cup, allowing them to be released after several days and float to the surface for recovery. The Acousonde™ tags were programmed to alternate every 30 min between low-frequency (LF) and high-power (HP) channels with sampling rates of 25.8 kHz and 232.3 kHz, respectively (see Wiggins, 2013, for details). The HP channel included a 22 dB low-pass anti-alias filter at 100 kHz, allowing for high quality representations of the primary bandwidth of narwhal clicks (20–70 kHz; Marcoux *et al.*, 2012; Koblitz *et al.*, 2016). Two tags (NW08, NW09) were programmed with an additional 20 dB gain, which was subtracted to match the relative amplitudes of the other six tags prior to acoustic analyses.

Because one aim of the overall tagging effort was to investigate year-round movements, narwhals recorded in this study were also fitted with “backpack” telemetry tags during capture, which were secured with pins through the dorsal ridge. Remotely deployed tags on similar species tend to last only a few weeks to months (Andrews *et al.*, 2008), making net-based capture and tagging an important tool in their study (Gonzalez, 2001; Blackwell *et al.*, 2018). Acousonde™ tags were sometimes secondarily secured by cable to these backpack tags, although the resulting telemetry data were not used in this analysis.

### B. Exploratory analyses

#### 1. Sound auditing and initial call selection

Sound files of the HP channel were visually and aurally audited in 15-s viewing frames in Raven Pro 1.3 [2007; fast Fourier transform (FFT) length 4096, 50% overlap, Hann window maximum frequency display 80 kHz]. This audit informed the design of a simple pre-classification procedure, where calls with sufficient signal-to-noise ratio (SNR) to be visually and aurally discernable were annotated and extracted. A small amount of overlap by surfacing noise, clicks, or other vocalizations was permitted (Kaplan *et al.*, 2014). Initial call classes included vocalizations broadly classified as tonal, pulsed, or mixed (containing both tonal and pulsed components). Echolocation trains and buzzes were readily distinguishable from other vocalizations by differences in click rate (Lammers *et al.*, 2004; Arranz *et al.*, 2016) and were typified by clicks at relatively low inter-click intervals (approximately 200 ms) speeding up to very high repetition rates with inter-click intervals of approximately 3 ms (Rasmussen *et al.*, 2015; Blackwell *et al.*, 2018).

Given our interest in identifying repeated call types, some of which are produced when animals are isolated (Janik and Slater, 1998), we included all recordings from the point of release onward for these analyses. We had no knowledge regarding social affiliations between individuals. Furthermore, while all recordings used in this analysis came from tags attached to female narwhals, our inability to fully discriminate between focal and non-focal vocalizations meant that we made no assumptions regarding the sex of calling animals.

## 2. Two types of patterns: Classification of paired patterns and burst pulse series

Given that patterns of pulsed vocalizations have not yet been described in narwhals, we had to create bespoke definitions of sequence types. We defined *paired patterns* as two stereotyped, pulsed (i.e., click-based) units that co-occur in the same order within a short time interval (<2 s), the pair of which is repeated at least three times within 30 s (e.g., “A-B-A-B-A-B”; Fig. 1).

We also identified a series of short pulsed vocalizations, possibly similar to the “chitter” described by Stafford *et al.* (2012). These were termed as *burst pulse series* (Fig. 2). We limited our analysis to sets of at least three burst pulse series, each produced within 10 s of the next. Subunits greater than 0.5 s apart were assumed to belong to separate series.

Time and repetition thresholds were informed by exploratory analyses as well as the timing of rhythmically repeated vocalizations in other odontocetes (Janik and Sayigh, 2013; Riesch *et al.*, 2006; Zwamborn and Whitehead, 2017). Quantitative support for the distinction between paired patterns and burst pulse series was verified retrospectively with a two-sample *t*-test comparing the

duration of subunits of each type (i.e. the AB units in Fig. 1 compared to the sub-units in Fig. 2).

## C. Testing stereotyped organization of units

### 1. Visual classification of paired patterns

We used a visual classification task with independent human observers to test our prediction that narwhals combined recognizable types of “A-B” pairs into repetitive sequences (see Appendix A for details of the classification task). Fleiss’ kappa statistic was used to quantify agreement between observers (irr Package in R; Gamer *et al.*, 2015; Landis and Koch, 1977). To assess the distribution of call types across time and recording devices, we labelled each pair according to the type matched by the majority of raters if at least 4/6 agreed on the same type. Any remaining patterns were excluded from further analyses.

### 2. Discriminant function analysis of paired patterns

In principle, animals could produce vocalizations containing gaps of silence that are still perceived as single calls. Support for the interpretation that paired patterns are combinations of units would come from the finding that their constituent units are also produced alone. To this end, we performed an initial classification of any “lone” calls that appeared to match a single unit of a paired pattern. These matches were verified with a discriminant function analysis (DFA), classifying the “A” units of all stereotyped patterns (as classified in the visual classification task), including the additional “lone units.”

For the DFA, we first filtered each A unit using a Butterworth four-pole bandpass filter bounded by user-defined LF and high-frequency limits from the selection process, similar to Marcoux *et al.* (2012). Individual clicks

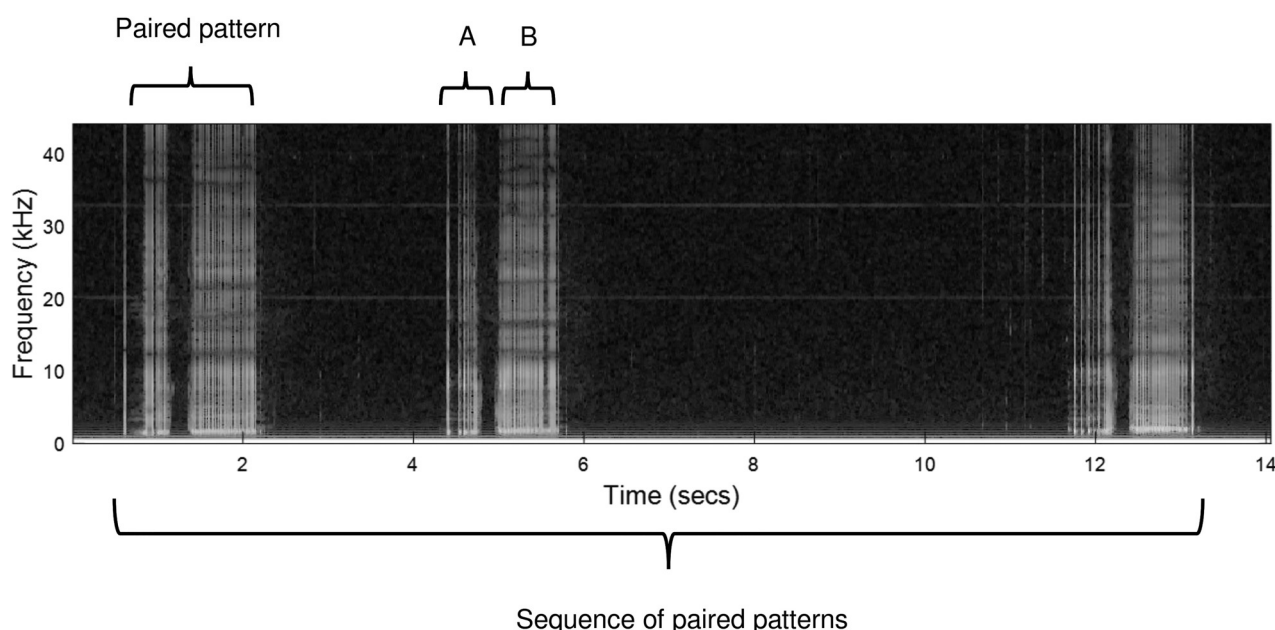


FIG. 1. Sample spectrogram showing definition of repeated “paired” vocalizations (FFT length, 4096; window, Hann; overlap, 50%).



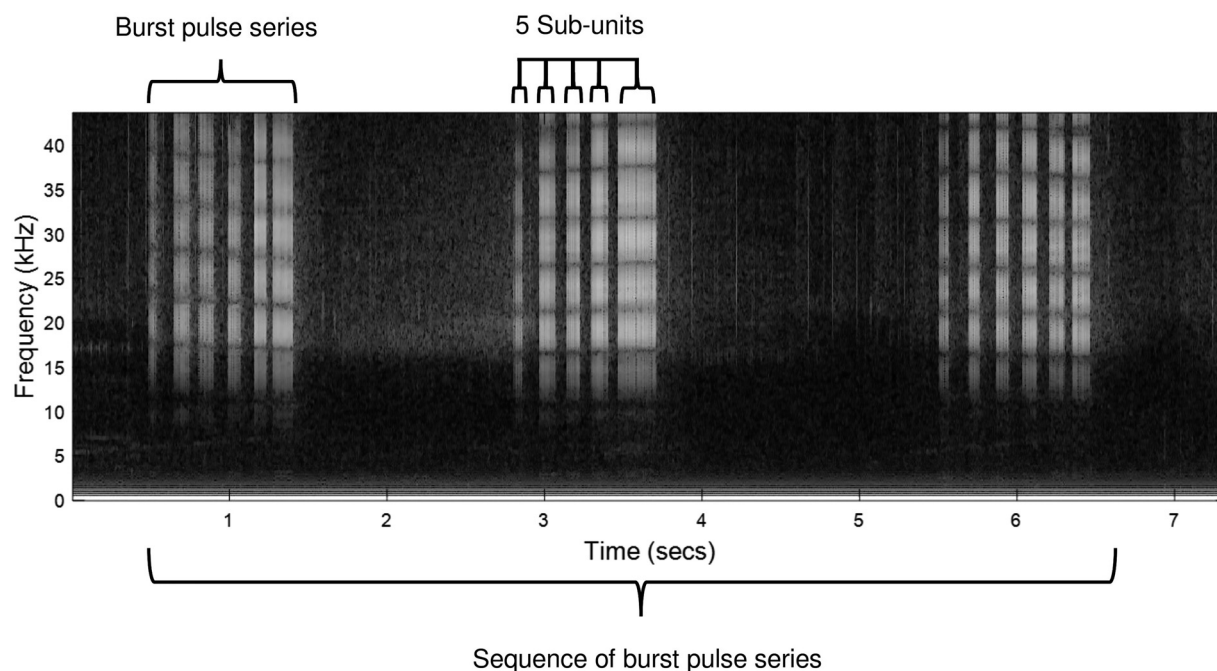


FIG. 2. Sample spectrogram showing the definition of the burst pulse series (FFT length, 4096; window, Hann; overlap, 50%).

were located using the “findpeaks” function from the Signal Processing Toolbox in MATLAB (R2017b; The MathWorks, Natick, MA). Starting and ending pulse repetition rate (PRR, measured in Hz) were then calculated from the initial and final quartiles of the call, respectively. These PRR measurements and duration (s) were checked for normality and homogeneity of variance and then entered into a DFA. (HH Package in R; Heiberger, 2018). Only paired pattern types representing four or more A-B pairs were used so that the number of variables in the DFA did not exceed the size of the smallest group.

### 3. Levenshtein distance analysis of burst pulse series

Using a custom MATLAB (R2017b) script, we manually marked the start- and end-times of each subunit in each series. These data were transformed into a binary string that represented, with one value per recording sample, the precise timing of the burst pulse series as “on” (1) or “off” (0). These strings were then down-sampled to every 500th value, allowing us to reduce the computational load of subsequent analyses while retaining high temporal resolution (465 values/s).

Levenshtein distance (LD) is the minimum number of insertions, deletions, and substitutions, required to transform one sequence into another, and has been used to compare sequences as diverse as deoxyribonucleic acid (DNA), language (Petroni and Serva, 2010), and humpback song (Garland et al., 2012; Kershenbaum and Garland, 2015). To test for repetition in sequences of burst pulse series, we calculated LD between series using the binary call representations, first for *observed transitions*, defined as adjacently produced burst pulse series within originally identified sequences and

then also for an equal number of *random comparisons*, pairwise comparisons of the same dataset of burst pulse series reshuffled in random order. Random comparisons were calculated within samples from each tag to accommodate for any individual-specific differences in vocal behaviour, which could bias subsequent estimates of stereotypy. We interpreted lower LD (i.e., higher similarity) values for observed compared to random transitions as evidence of repeated use of stereotyped sequences.

We fit a Poisson family GLM using the canonical link function to compare LDs between observed and random transitions according to the following equations:

$$\overline{LD} = \exp(\beta_0 + \phi_{\text{Random}}\beta_1),$$

$$LD \sim \text{Poisson}(\overline{LD}),$$

where LD represents actual measures of LD and  $\overline{LD}$  represents expected measures.  $\beta_0$  is the model intercept, representing mean LD for observed transitions between burst pulse series.  $\beta_1$  is a contrast representing the difference between LD measures of random comparisons relative to observed transitions. We also fit an identical model with the quasi-Poisson family to check for overdispersion in LD values. Only recordings with more than one sequence of burst pulse series were included.

## D. Inferring patterns of production across individuals

### 1. Discriminating between sequences and call exchanges

A lack of overlap between repeated signals is often used to support inferences that a sequence is produced by a

single individual (Luís *et al.*, 2018; Sayigh *et al.*, 2013; Zwamborn and Whitehead, 2017). However, some species produce call exchanges with precise and stereotyped timing (Mann *et al.*, 2006; Pika *et al.*, 2018), meaning that nonoverlapping sequences can also result from exchanges between multiple individuals (Miller *et al.*, 2004). To assess whether sequences of vocalizations are produced by a single animal, we examined not only overlap, but also variation in amplitude and frequency of pulsed vocalizations produced in sequences, expecting that the exceptionally directional clicks of narwhals (Koblitz *et al.*, 2016) should not be consistently recorded with the same acoustic characteristics from multiple individuals.

## 2. Exclusion of patterns produced by non-focal individuals

Clicks recorded from suction-cup-tagged cetaceans are hypothesized to include additional LF components that are conducted through the body (Johnson *et al.*, 2006; Johnson *et al.*, 2009; Zimmer *et al.*, 2005; Blackwell *et al.*, 2018). We examined the frequency spectra of clicks from each call sequence to determine which were unlikely to be produced by the focal individual (see Appendix B for details of the discrimination method). We compared the remaining “possibly focal” sequences across tag recordings to consider evidence of individual- or group-specific call characteristics.

## E. Modeling the behavioural context of sequence production

We fit GLMs to explore the context, and thereby potential function, of narwhal call sequences. Models were fit according to the following equations:

$$(\bar{S}) = \frac{\exp(\beta_0 + \beta_1 v + \beta_2 d + \beta_3 \phi_Z)}{1 + \exp(\beta_0 + \beta_1 v + \beta_2 d + \beta_3 \phi_Z)},$$

$$S \sim \text{Binomial}(\bar{S}),$$

where  $\bar{S}$  represents the expected probability of a sequence occurring in a given 60 second period.  $\beta_0$  is the intercept of the linear predictor.  $\beta_1$  represents the effect of the number of other calls ( $v$ ),  $\beta_2$  represents the effect of depth ( $d$ ) in meters, and  $\beta_3$  represents the difference in model intercept when buzzes were detected. Call counts and depth values were  $\log_{10}$ -transformed as the raw data were positively skewed. We expected that a logarithmic representation would be biologically appropriate because a change in depth of 1 m vs 5 m represents a proportionally much larger difference in pressure than 100 m vs 105 m, for example. Depth profiles were calibrated by zero-offset correction using a custom script in MATLAB from the Animal Tag Tools Calibration Toolkit.<sup>1</sup>

This model structure was applied separately to patterns of pulsed calls and burst pulse series, and for each, according to the 60 s preceding a call (“previous activity models”), and the 60 s following a call (“subsequent activity models”). For pseudo-absences, we randomly selected 60 additional 60-s periods from the same tags where the sequences in

question were identified but in which no sequences were heard. These pseudo-absences were only selected from periods following the first vocalization identified on a tag recording and prior to tag detachment as determined from the dive profile. All four models were also fit using GLMs with quasi-binomial families to check for overdispersion, and diagnostic plots were used to validate all model structures prior to consideration of estimated parameters. Statistical analyses were done in R 3.4.2 (R Core Team 2017). Analysis scripts and basic datasets are available via the Open Science Framework.<sup>2</sup>

## III. RESULTS

### A. General features of recordings

Eight female narwhals were equipped with hydrophone tag packages. We identified 3261 vocalizations from the resulting high-frequency recordings, including tonal calls, pulsed calls, mixed calls, and patterns of pulsed calls (Table I). Most tags (7/8) detached from the narwhal prior to maximum recording duration, resulting in varying recording lengths per tag (Table I). Periods of multiple overlapping vocalizations at rates of up to 3 calls/s were identified on 4/8 tags, and presumed to be the result of narwhals aggregating into larger herds. Consistent with other studies, pulsed vocalizations outnumbered whistles (Stafford *et al.*, 2012; Miller *et al.*, 1995), although we did note several periods when many overlapping whistles were produced, as described by Ford and Fisher (1978).

Contrary to suggestions that narwhals tend not to feed in the summertime (Stafford *et al.*, 2012), buzzes indicative of prey capture attempts were identified on all but one tag (7/8; Table I). Buzzes tended to occur in series with a short gap after the terminal buzz, and occurred at very high rates on some recordings (up to 11 buzzes/min). Narwhals did not appear to withhold from vocalizing immediately after tagging, as has been found in other studies. Even when sampling 50% of possible recordings (only the HP channel), we found that narwhals used buzzes comparatively soon after tagging [ $8.5 \pm 2.8$  [standard error (SE)] hours;  $N=7$ , excluding one animal that did not buzz at all] compared to  $22.7 \pm 5.1$  h of post-release silence as found by Blackwell *et al.* (2018;  $N=6$ ).

### B. Summary of identified patterns

We identified eight plausible call types consistent with our definition of paired patterns. We also included one pattern of two stereotyped pulsed calls with an additional small set of clicks preceding a clear pair of calls (type I; Fig. 3) and one pattern for which both units were linked by a tonal “bridge” (type VI; Fig. 3). Each of these ten patterns was included as a template in the subsequent matching task. This can be seen in Mm. 1.

**Mm. 1.** Combined sound file of ten templates. This is a file of type “wav” (13.3 Mb).

TABLE I. Details of female *M. monoceros* fit with Acousonde™ acoustic tags in Tremblay Sound, Nunavut. Here, “pattern” refers to any plausible stereotyped patterns of pulsed calls as well as series of burst pulses. Note that “time to vocalize” was calculated using the LF channel as well, so may exceed the duration of the HP channel recording. No buzzes were detected on the recording from NW14.

Narwhal ID	Deployment date	Recording length of HP channel	Time to vocalize (min)		Calls detected			
			Call	Buzz	Tonal	Pulsed	Mixed	Pattern
NW08	Aug 13, 2017	41 h	53	30	15	110	0	199
NW09	Aug 16, 2017	16 h	38	246	2	17	0	0
NW11 <sup>a</sup>	Aug 30, 2017	4 h	<5	158	4	36	1	23
NW12 <sup>a</sup>	Sept 2, 2017	4 h	<5	754	408	696	240	583
NW14	Sept 3, 2017	12 h	10	N/A	45	91	8	46
NW18 <sup>a</sup>	Sept 11, 2017	11 h	34	870	12	99	5	10
NW19 <sup>b</sup>	Sept 11, 2017	8 h	38	271	23	41	0	5
NW20 <sup>b</sup>	Sept 11, 2017	16 h	31	1231	13	119	27	32

<sup>a</sup>Associated with a calf.

<sup>b</sup>Captured together.

Aural assessment at 0.1 × speed confirmed that the burst pulse series consisted of sets of clicks interspersed with clearly defined periods of silence. This can be seen in [Mm. 2](#). Burst pulse series were detected on nearly all tag recordings (7/8) but were distributed unevenly. Sequences of at least three burst pulse series with intervals of less than 10 s between each (e.g., [Fig. 2](#)) were identified on fewer tag recordings (3/8) but still comprised approximately 40% of all burst pulse series detected. Within these sequences, we also noticed that single subunits matching the amplitude, frequency, and duration of subunits in the more typical burst pulse series were sometimes produced. These singletons, while rare (7% of all burst pulse series), were also included in the analysis.

[Mm. 2](#). Example of burst pulse series. This is a file of type “wav” (3.6 Mb).

In total, we isolated 36 sequences comprising varying numbers of burst pulse series ( $N_{series} = 212$ ). Aural assessment suggested that PRR was relatively consistent within and across the burst pulse series. Some sequences were composed of series with different numbers and temporal organization of subunits. However, most sequences appeared to be repetitive with small modifications such as the addition or subtraction of a single set of clicks ([Fig. 2](#)). The length of individual subunits in the burst pulse series often appeared to increase over the course of the call, and some (approximately 14%) of the burst pulse series were initiated by an especially short subunit (<50 ms).

Subunits of paired patterns [mean = 515.8 ± 40.8 (SE) ms;  $N = 21$ ] were significantly longer than those of the burst pulse series [mean 104.9 ± 1.5 (SE) ms;  $N = 900$ ; two-sample *t*-test;  $t(20.06) = 10.05$ ,  $p < 0.001$ ], supporting

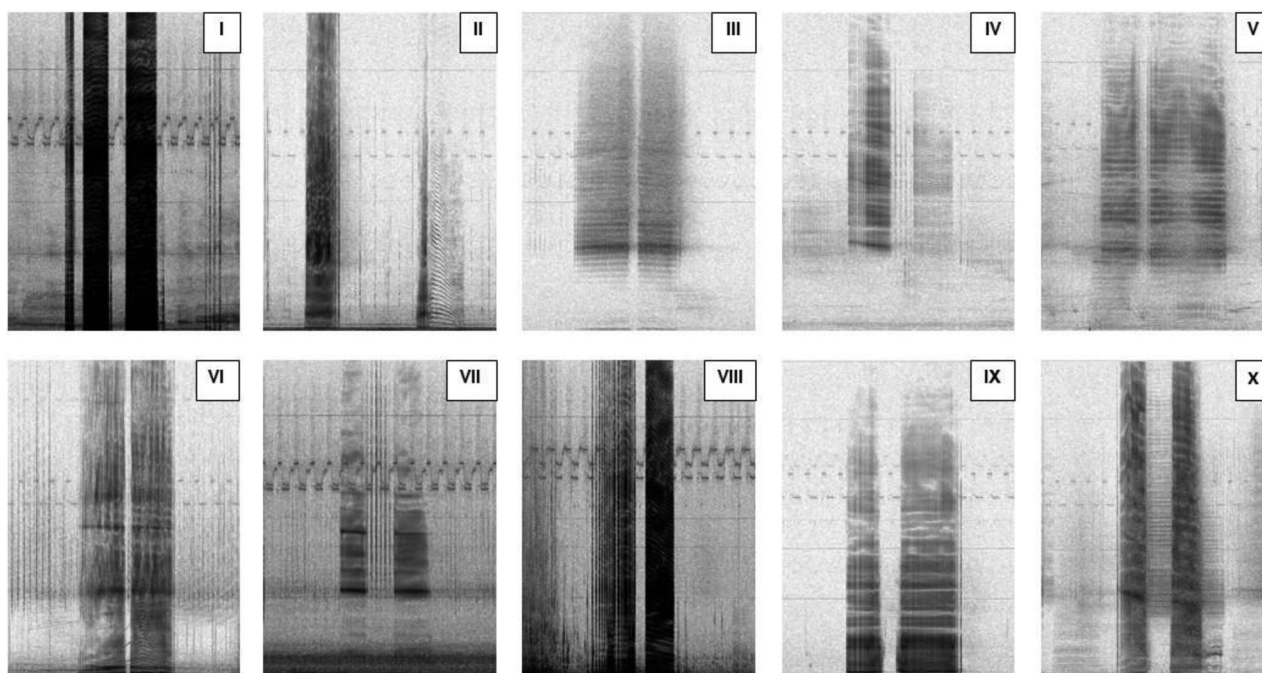


FIG. 3. Templates of paired pattern types used in the visual classification task. All templates are shown with identically calculated spectrograms, 80 kHz maximum frequency, and 3 s duration (FFT length, 4096; window, Hann; overlap, 50%).



TABLE II. Summary statistics of Fleiss' kappa test applied to multiple observer classification of patterns of *M. monoceros* paired patterns. Tag specificity indicates the recording on which each pattern was identified with the number of A-B pairs in parentheses. Bolded tag identifications (IDs) designate calls that were likely produced by the tagged animal. Reliability descriptors are taken from Landis and Koch (1977).

	<i>k</i>	<i>z</i>	<i>p</i>	Reliability descriptor	Tag specificity ( <i>N</i> )
Pattern					
I	0.813	31.472	<0.001	Near-perfect	<b>NW08 (8)</b>
II	0.921	35.683	<0.001	Near-perfect	<b>NW12 (7)</b>
III	0.399	15.469	<0.001	Fair	NW12 (5)
IV	0.757	29.316	<0.001	Substantial	NW12 (5)
V	0.660	25.551	<0.001	Substantial	NW12 (8)
VI	0.492	19.060	<0.001	Moderate	NW12 (1), <b>NW14 (5)</b>
VII	0.533	20.633	<0.001	Moderate	NW08 (4)
VIII	0.772	29.893	<0.001	Substantial	<b>NW08 (16)</b> , NW08 (1), <b>NW11 (1)</b>
IX	0.607	23.525	<0.001	Substantial	<b>NW11 (6)</b> , NW12 (2)
X	0.373	14.439	<0.001	Fair	<b>NW12 (3)</b>
No match	0.321	12.447	<0.001	Fair	(All)
Overall					
All classes	0.584	63.4	<0.001	Substantial	

our decision to split them into distinct classes of vocal sequences.<sup>3</sup>

### C. Test of stereotyped organization of units

Six participants naive to the original order and context of vocalizations consistently matched paired patterns to ten templates, although interobserver agreement varied between types (Table II). Quantitative classification was also successful: using jackknifed cross-validation, the DFA classified paired patterns with 80.6% accuracy [Wilk's  $\Lambda = 0.04$ ,  $X^2 = 205.5$ , DF (degrees of freedom) = 21.0,  $p < 0.001$ ].<sup>3</sup> Here, classification accuracy was dependent on pattern type: four classes had very high accuracies (I, 100%; II, 100%; IV, 100%; VIII, 94%; IX, 88.9%), and three classes were not accurately discriminated (III, 0%; V, 50%; VI, 0%).

We identified 29 sequences of paired patterns, consisting of 72 pairs in total. Classifiers confirmed that each sequence contained just a single type of A-B pair (Fig. 3). We identified lone units matching the A units of two types of paired patterns (II and IX). These were classified with their expected types with 100% accuracy in the DFA,

confirming that individual units of the paired patterns are sometimes produced alone. We detected no lone "B" units for any patterns. Lone A units were sometimes rhythmically repeated (Fig. 4) and often produced in close proximity to full "pairs" (Fig. 5). When both units of a pattern were produced together, they were always produced in the same order. We also noticed that the mean PRR of the B unit was higher than that of the A unit for 9/10 types.

Measures of LD between burst pulse series were overdispersed (dispersion parameter 104.9), so we interpreted effect coefficients from the quasi-Poisson GLM. Pairs of burst pulse series from observed transitions were more similar than pairs from randomly generated transitions (mean  $LD_{\text{Observed}} = 150.7$ , mean  $LD_{\text{Random}} = 200.7$ ,  $p < 0.001$ ). Burst pulse series did not appear to change gradually across the study period,<sup>3</sup> indicating that narwhals combine burst pulse series into repetitive "sequences of series."

### D. Patterns of production across individuals

Frequency and amplitude were highly consistent between units of paired patterns and across subsequent A-B

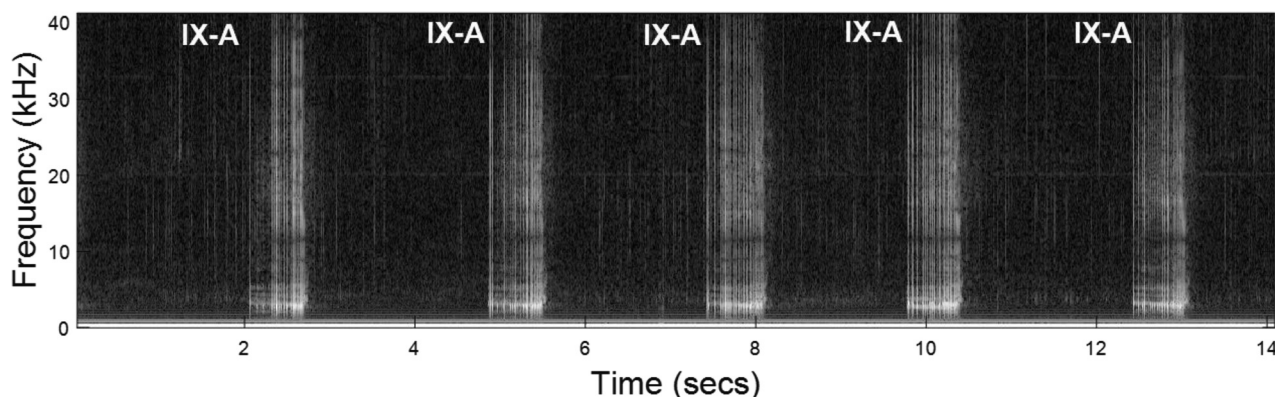


FIG. 4. Spectrogram showing lone A units of a stereotyped pattern of calls repeated in rapid succession (type IX; FFT length, 4096; window, Hann; overlap, 50%). Classified visually ( $kappa$  statistic 0.607,  $p < 0.001$ ) and by DFA (88.9% classification for type IX).

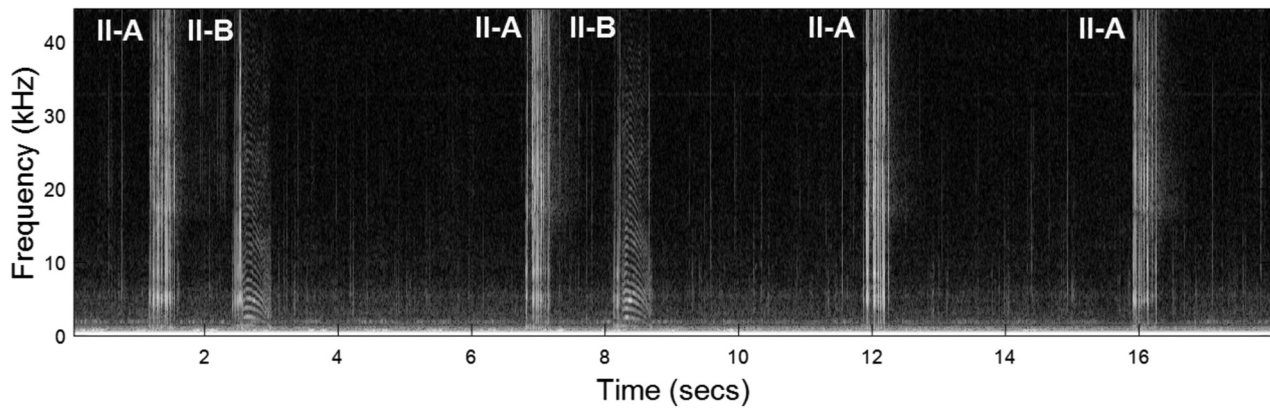


FIG. 5. Spectrogram showing a stereotyped pattern of calls followed by two additional lone A units (type II; FFT length, 4096; window, Hann; overlap, 50%). Classified visually ( $\kappa$  statistic 0.921,  $p < 0.001$ ) and by DFA (100% classification).

pairs in the same sequence. There was one exception: a sequence of sub-type VIII included one pair produced with clicks of higher frequency than others in the sequence.<sup>3</sup> We did not identify any instances of overlapping A and B units of paired patterns, nor was there any overlap of repeated pairs. Sequences of burst pulse series were generally produced with consistent frequencies and amplitudes, although we detected two sequences with overlapping series.

When excluding calls attributable to non-focal individuals, 98% (45/46) of remaining paired patterns were identified on a single recording only (bold sequences in Table II). These tag-specific patterns appeared to be stable over our recording durations: types I, II, VIII, and IX were identified 7.2, 10.2, 38.6, and 40.7 h apart within their respective recordings. Although sequences of burst pulse series (as defined above) were only detected on three tags, and only two tags had more than one sequence, we detected tag-specific differences in burst pulse series characteristics: series recorded from NW08 had fewer subunits than series recorded from NW12 (two-sample  $t$ -test; mean subunits<sub>NW08</sub> = 3.29, mean subunits<sub>NW12</sub> = 4.79;  $t(207) = -6.1$ ,  $p < 0.001$ ).

### E. Behavioural context of sequence production

Relative to randomly selected samples, paired patterns were 50% more likely, and burst pulse series 77% less likely, respectively, to be followed by buzzes, although these estimates were associated with very large standard error and were not statistically significant (Table III). Depth did not appear to influence the likelihood of detecting vocal sequences, matching our observation that sequences were produced at various positions in the water column, except the deepest dives (i.e., below 300 m; Table III, Fig. 6).

A one-unit increase in  $\log_{10}$ -transformed counts of vocalizations in a 60-s recording sample corresponded to a 330% increase in the probability of the sample being preceded (and 460% increase of being followed) by a burst pulse series, suggesting they are associated with vocal exchanges generally (Table III). In contrast, paired patterns were significantly less likely to be produced in the context of other calls (Fig. 7). Three sequences of burst pulse series

from one tag (NW12) were excluded from behavioural context models as pressure sensors on the tag failed midway through the deployment.

### IV. DISCUSSION

We have shown that narwhals produce at least two kinds of vocal sequences. Paired patterns are composed of distinct units combined into strictly ordered sets, providing parallels to vocal sequences in avian and other mammalian species. Burst pulse series, similar to vocalizations in other odontocetes, are shown to be combined into repetitive sequences of series.

TABLE III. Parameter estimates for GLMs relating the occurrence of sequences of paired patterns ( $n = 29$ ) or burst pulse series ( $n = 21$ ) produced by *M. monoceros* to the behavioural context preceding and following the call. Call occurrences were supplemented with 60 randomly selected 60-s samples to provide pseudo-absences.

	Estimate	SE	$z$	$p$
<b>Paired patterns</b>				
<i>Previous activity model</i>				
Intercept	-0.304	0.488	-0.624	0.532
Calls	-0.865	0.393	-2.202	0.028
Depth	-0.005	0.351	-0.014	0.989
Feeding buzzes	-0.489	0.688	-0.711	0.477
<i>Subsequent activity model</i>				
Intercept	-0.550	0.493	-1.116	0.265
Calls	-0.476	0.388	-1.228	0.219
Depth	-0.205	0.327	-0.626	0.531
Feeding buzzes	0.407	0.574	0.709	0.478
<b>Burst pulse series</b>				
<i>Previous activity model</i>				
Intercept	-1.790	0.679	-2.637	0.008
Calls	1.231	0.499	2.469	0.014
Depth	0.484	0.496	0.975	0.329
Feeding buzzes	-0.488	0.912	-0.535	0.593
<i>Subsequent activity model</i>				
Intercept	-2.281	0.819	-2.786	0.005
Calls	1.526	0.533	2.863	0.004
Depth	0.835	0.495	1.688	0.091
Feeding buzzes	-1.472	0.868	-1.697	0.090



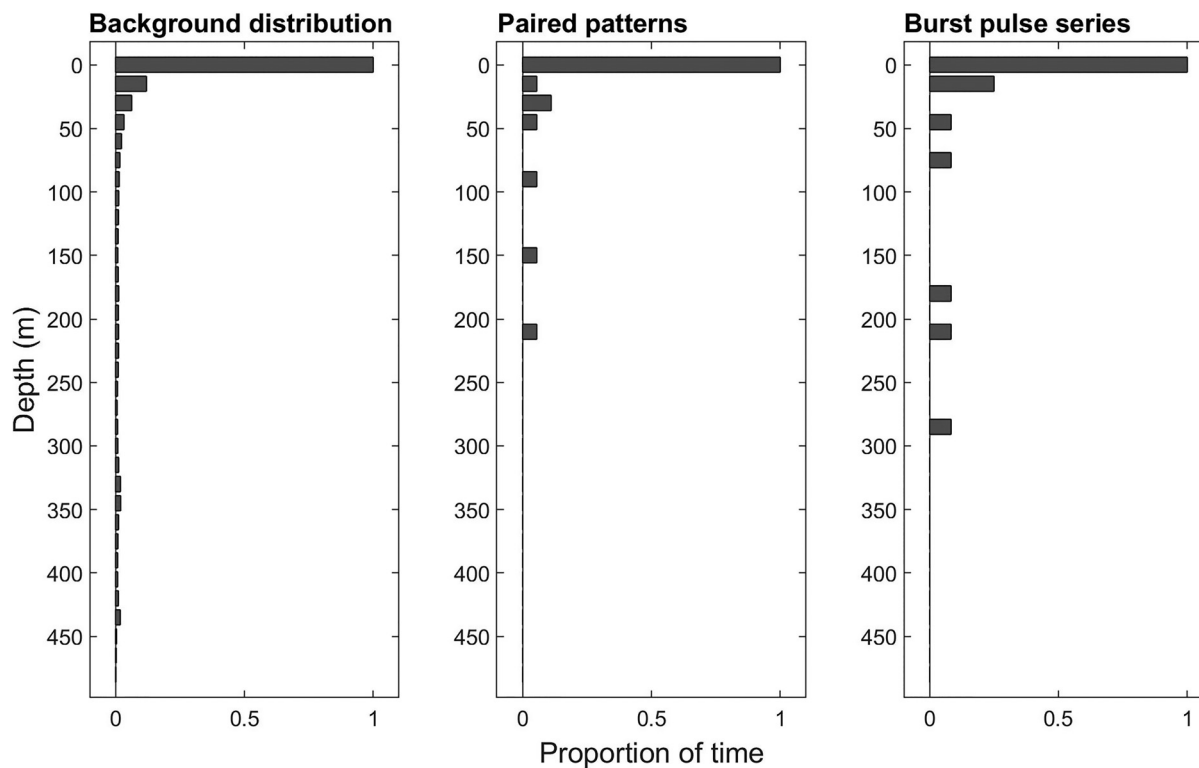


FIG. 6. Distributions of depths registered on the tags (left) compared to depths at which sequences of paired patterns (middle) and burst pulse series (right) were produced.

Consensus between the visual classification task and the DFA for most types of paired patterns demonstrates that they are stereotyped and rhythmically repeated. The finding that the first units of two types (II, IX) were also produced alone provides further support for the interpretation of these

as combinations of calls, rather than single calls with gaps of silence (Kershenbaum *et al.*, 2016). We are unable to conclude with certainty that other types are also composed of divisible units as we did not identify any candidate cases. However, this may have largely been an artefact of the rarity

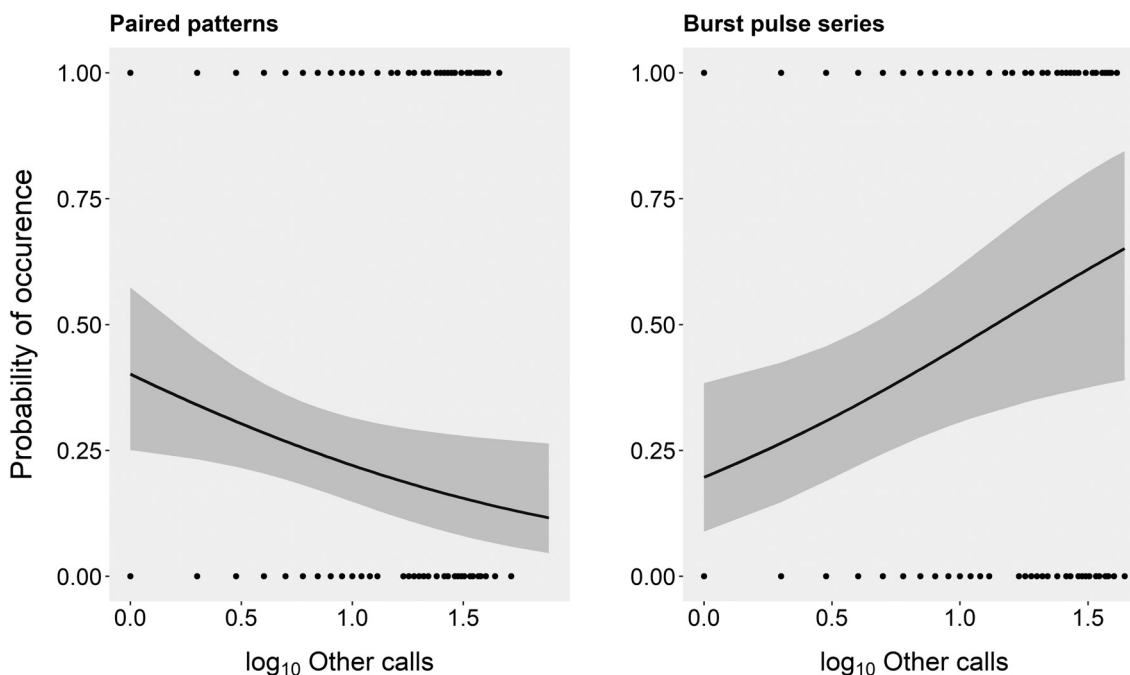


FIG. 7. Marginal effects of other vocalizations preceding the occurrence of sequences of pulsed vocalizations. Effects were estimated using binomial family GLMs.

of certain types as the two patterns for which we identified lone units were relatively common (Table I). To our knowledge, there are only two examples of ordered sequences in odontocetes: the “N7-N8” calls of Northern resident killer whales and the bray sequences of bottlenose dolphins (Ford, 1988, Luís *et al.*, 2018). Structurally similar combinations of vocalizations are found more commonly in other taxa, where they have been linked to communicative complexity. For example, meerkats (*Suricata suricatta*) produce similar “di-drrr” calls composed of two units, the pair of which are rhythmically repeated, and interpreted as evidence for hierarchical communication (Collier *et al.*, 2014). Putty-nosed monkeys (*Cercopithecus nictitans*) produce “pyow” and “hack” calls, which are regularly combined into ordered sequences but are also sometimes produced alone. Playback experiments have demonstrated that the ordering of these calls conveys distinct meaning and is considered a rare example of syntax in a wild organism (Arnold and Zuberbühler, 2008; Kershenbaum *et al.*, 2016). As such, the finding that B units of paired patterns were only ever found following A units is noteworthy as it is indicative of a “finite-state” grammar similar to those documented in primate and avian taxa (Berwick *et al.*, 2011; Fitch and Hauser, 2009; Shettleworth, 2010). While inflexible A-B grammars are fundamentally different from the recursive grammatical structures that support human language (Berwick *et al.*, 2011; Fitch and Hauser, 2009), they can be important vehicles for the transfer of information (Bradbury and Vehrencamp, 1998, p. 395). Of course, it remains to be tested whether the rule-like structure of narwhal paired patterns supports compositional communication, i.e., whether A and B units convey distinct information when combined (Zuberbühler, 2018).

Whereas burst pulse series are varied (e.g., in number of subunits), narwhals combine similar series into repetitive sequences. Vocalizations similar to burst pulse series have been detected in recordings of northern right whale dolphins (*Lissodelphis borealis*; Rankin *et al.*, 2007), dusky dolphins (*Lagenorhynchus obscurus*; Vaughn-Hirshorn *et al.*, 2012), pacific white-sided dolphins (*Lagenorhynchus obliquidens*; Henderson *et al.*, 2011), and Heaviside’s dolphins (*Cephalorhynchus heavisidii*; Martin *et al.*, 2019), although there are nontrivial differences between species. For example, the “patterned burst pulses” in northern right whale dolphins are the only other series that appear to be rhythmically repeated, and the “burst pulse sequences” produced by dusky dolphins rarely had more than two subunits. Calves of Araguaian river dolphins (*Inia araguaiaensis*) produce calls similar to the narwhals’ burst pulse series, although it remains to be tested whether they are combined into sequences of series (Melo-Santos *et al.*, 2019).

We are unable to rule out the possibility that sequences were sometimes produced by multiple individuals in coordinated call exchanges, which has important implications for the interpretation of communicative complexity (as discussed above). However, several findings were inconsistent with this explanation. First, the strong directionality of

narwhal clicks makes it unlikely that multiple individuals would be able to consistently produce clicks resulting in near-identical frequency ranges on an animal-attached tag (Koblitz *et al.*, 2016). Second, overlap between burst pulse series was very rare and was not identified in paired patterns, supporting the interpretation that they were produced by a single individual (Sayigh *et al.*, 2013). If these sequences were the result of call exchanges, they would imply a high degree of temporal synchrony, akin to duetting or even “turn taking,” given the highly stereotyped time intervals between units (Pika *et al.*, 2018). Therefore the most parsimonious interpretation is that sequences of paired patterns were produced by single individuals. The same can be concluded for sequences of burst pulse series, whereas the detection of some cases of overlap suggests that they may also be produced in call exchanges.

If sequences are typically produced by a single individual, the variation in sequence characteristics between tags opens the possibility that they support individual or group recognition, as has been suggested for other narwhal vocalizations (Shapiro, 2006). For example, call types associated with specific tag recordings have been taken as preliminary evidence for individual-specific signals in short-finned pilot whales (*G. macrorhynchus*; Quick *et al.*, 2018). We identified possible individual specificity in the paired patterns, which were re-produced on the same tags up to 40 h apart. However, two considerations suggest that these patterns should not be interpreted as “signature sequences.” First, we would expect a signature signal to be the predominant vocalization in an individual’s repertoire (Cook *et al.*, 2004), whereas paired patterns were relatively rare. Second, for species with large vocal repertoires, a subsampling should result in recording-specific differences, even if repertoire composition is identical across individuals, making us cautious to label these as individual specific without further inquiry. We were surprised to find that burst pulse series recorded from two different individuals had different and continuous (i.e., not obviously incomplete or under-sampled) distributions of numbers of subunits. Nevertheless, these patterns seem unlikely to disclose individual identity, given the wide range of numbers of subunits used.

In addition to the consideration of signal properties, linking acoustic cues to behaviour can help to provide insights into their function (Papale *et al.*, 2017). Burst pulse series were more likely to be produced in contexts of high vocal activity. This matches well with theory that graded signals are primarily used in social contexts where animals are in close proximity (Ford, 1988). We detected the opposite effect for paired patterns. These calls were often produced in very quiet periods. Accordingly, it may be that the highly stereotyped paired patterns may serve a role in longer-distance communication. Neither paired patterns nor burst pulse series were significantly related to prey-capture attempts, and were not recorded when individual narwhals were below 300 m (Fig. 6), suggesting that they are not related to feeding, as are the “bray sequences” of bottlenose dolphins (Janik, 2000; King and Janik, 2015). Combining

our findings with visual observations (e.g., from shore or uncrewed aerial vehicle) might reveal further associations between context and sequence use. For example, call sequences may facilitate contact between mothers and calves (Smolker *et al.*, 1993), like the “type A” pulsed vocalizations of belugas (Vergara *et al.*, 2010). Female narwhals give birth in midsummer (Furgal and Laing, 2012; Heide-Jørgensen, 2009), and paired patterns were recorded on all tags attached to narwhals observed with calves during capture (Table I).

Both paired patterns and burst pulse series are characterized by rhythmic repetition. While the energetic cost of vocalizing underwater may not be a significant limitation (Jensen *et al.*, 2012), the risk of detection by eavesdropping acoustic predators (e.g., killer whales) is likely important for narwhals (Deecke *et al.*, 2002; Breed *et al.*, 2017; Furgal and Laing, 2012; Laidre *et al.*, 2006), making the use of repeated signals surprising. One benefit of repetitive signaling is to increase the likelihood of successful transmission, perhaps especially important in the icy, reverberation-prone environments inhabited by narwhals (Brumm and Slater, 2006; Ey and Fisher, 2009; Vergara *et al.*, 2010). For example, links between noise and signal redundancy have been identified in blue whales (*Balaenoptera musculus*; Miller *et al.*, 2000) and killer whales (Foote *et al.*, 2004). Alternatively, the repetitive nature of these sequences may suggest that call production is sustained until a response is achieved in a target animal or the number of repetitions itself encodes specific information (Payne and Pagel, 1997; Janik and Sayigh, 2013; Sloan and Hare, 2004). For example, the number of “dee” units in call sequences of black-capped chickadees (*Poecile atricapilla*) is inversely proportional to the size of an encroaching predator (Templeton *et al.*, 2005). Encoding information in the temporal structure of a signal is also consistent with the optimization of transmission success in noisy habitats; these features should degrade *less* severely than other acoustic properties over long distances (Bradbury and Vehrencamp, 1998, p. 129). This reasoning is proposed for the use of temporally stereotyped vocalizations in white-eyed mangabeys (*Cercocebus* sp.; Wasser, 1982, p. 134), sperm whales (Gero *et al.*, 2016), and forest-dwelling chingolo sparrows (*Zonotrichia capensis*; Bradbury and Vehrencamp, 1998, p. 136).

## V. CONCLUSIONS

We conclude that narwhals produce at least two kinds of vocal sequences. The paired patterns appear to be one of few ordered, multiunit sequences in odontocetes, and to our knowledge, the first in the *Monodontidae*. Whereas we propose that they may play a role in long-distance communication, their specific function remains unknown. The use of repetition in successive burst pulse series suggests that their signaling benefit outweighs the costs associated with redundant signaling. Further inquiry into the function and population-distribution of sequence use (e.g., whether they are produced by narwhals in East Greenland and Western

Hudson Bay) could lead to insights regarding the ecological, genetic, and cultural development of narwhal communication (Garland *et al.*, 2011; Janik, 2000; May-Collado and Wartzok, 2008; Rendell and Whitehead, 2001). Together, our findings provide a small contribution to the understanding of the phylogenetic distribution of vocal sequence production and suggest that methodological limitations may explain the apparent paucity of vocal sequences in other odontocetes. We hope that future inquiry will continue to elucidate the function(s) of these sequences in narwhals and widen the set of species in which sequence use is tested, ultimately supporting comparative studies to better understand evolution of vocal communication.

## ACKNOWLEDGMENTS

We thank Patrick Miller and Katie Harris for providing helpful comments on early iterations of this work. We also thank Mikhail Barabanov, Vanessa Simons, Rowan Gardner, Nick Jones, and Diana Mirella Pabón Figueroa for participating in the visual classification task. Funding for S.F.W. was provided by an R and A Ransome Scholarship. Data used in this analysis were collected by the Ecosystem Approach to Tremblay Sound project led by the Department of Fisheries and Oceans, Canada. We thank Robert Hodgson and all participants involved in this project and the associated funding bodies, including the Nunavut Wildlife Management Board, the World Wildlife Fund, Golder Associates, and Ocean Wise. We are grateful to the Mittimatalik Hunters & Trappers Organization and the Mittimatalik Inuit Community for their support and involvement in this project. We also thank the Polar Continental Shelf Program, the Government of Nunavut, Parks Canada, the Ocean Tracking Network, l’Université du Québec à Rimouski, the University of Windsor, the University of Manitoba, the University of Calgary, l’Université de Montréal, the University of British Columbia, and the University of New Hampshire for their partnership. Finally, we thank two anonymous reviewers whose comments improved the manuscript.

## APPENDIX A: MULTIPLE-OBERVER CLASSIFICATION

Visual classification has been shown to be an effective method for the classification for vocal sequences (Janik and Sayigh, 2013; Kershenbaum *et al.*, 2016). While spectrographic visualizations of pulsed signals are highly contingent on FFT and time dimensions used, differences in click rate (the expected form of stereotypy for paired patterns) should be consistently distinguishable as long as spectrogram parameters are kept constant.

One randomly selected pair from each of the paired pattern types we defined served as a template (Fig. 3). We then printed 100 patterns of pulsed calls to be matched to these templates. This number of patterns was chosen as a balance between including adequate variation to assess patterns across tags, while not being overly cumbersome for human classifiers, which could result in reduced performance (Rendell and Whitehead, 2003). These included the



remaining paired patterns from the original sequences, as well as possible pair-like patterns identified elsewhere in the classification procedure, but which were not rhythmically repeated. Given the propensity of narwhals to aggregate in groups where many (and often overlapping) vocalizations are heard, this test sample was likely to include pulsed calls that were produced close together in time by chance, i.e., “null” patterns.

Six participants matched call patterns to one of ten possible templates or indicated that no match was found. Participants had varying degrees of expertise in bioacoustics (range 1–20 yr), although none had previous experience classifying narwhal vocalizations. Participants were instructed to ignore any background noise or differences in frequency range for the classification task. Index numbers for calls were randomized prior to the task.

## APPENDIX B: DISCRIMINATING FOCAL AND NON-FOCAL VOCALIZATIONS

To improve our ability to make inferences regarding the individual specificity of call types, we devised a threshold for “unusual” low-sound frequency based on previous reports of narwhal click frequency ranges. Estimates of the peak frequency of narwhal clicks using standard hydrophones include 19 kHz (Miller *et al.*, 1995) and 12–20 kHz (Marcoux *et al.*, 2012). Others have reported that narwhals produce clicks at variable frequency ranges, the lowest having a peak frequency of 3.5–5 kHz and no energy below 3 kHz [Fig. 2(a) in Stafford *et al.*, 2012]. In contrast, recordings of tagged narwhals contain clicks with substantial energy below these frequencies (Shapiro, 2006; Blackwell *et al.*, 2018). As such, pulsed patterns with clearly visible energy below 3 kHz were identified as possibly focal.

On average, clicks produced by non-focal individuals should have lower amplitudes than those plausibly produced by the tagged animal, acknowledging that non-focal vocalizations may sometimes produce higher received levels (Johnson *et al.*, 2009). We used a two-sample *t*-test to test the prediction that pulsed vocalizations with visible LF energy (<3 kHz) should have greater root mean square (RMS) amplitudes, using the first unit of each paired pattern. A significant difference would provide further support for the hypothesis that click-based sounds lacking energy below 3 kHz are produced by non-focal whales. Units were filtered with a 3 kHz high-pass filter to remove the confounding effects of the LF energy, itself, influencing amplitude. As predicted, we found that pulsed vocalizations with energy below 3 kHz had higher amplitudes than calls lacking energy below 3 kHz (two-sample *t*-test; mean amplitude<sub>lowFreq</sub> = 156.8 dB RMS, mean amplitude<sub>noLowFreq</sub> = 142.4 dB RMS, *t*(108) = 9.57, *p* < 0.001), even when energies below 3 kHz were filtered out, supporting the hypothesis that calls lacking LF energy were produced by non-focal whales.

That being said, it is thought that high-amplitude pulsed vocalizations produced by nearby, non-focal individuals can also result in additional LF energy on a tag recording

(Blackwell *et al.*, 2018). As in-depth consideration of the conduction and propagation of LF energy was beyond the scope of this analysis, we assumed that our discrimination process should allow for an increased (but not total) ability to relate sequences of calls to specific individuals.

<sup>1</sup>See <http://www.animaltags.org/doku.php?id=tagwiki:tools:calibration> (Last viewed 1 June 2018).

<sup>2</sup>See [osf.io/wk9a5](https://osf.io/wk9a5) (Last viewed 31 January 2020).

<sup>3</sup>See supplementary material at <https://doi.org/10.1121/10.0000671> for distributions of subunit durations; boxplots of all paired pattern units showing distributions of duration, number of clicks, and mean PRR; the plot of dissimilarity between burst pulse series as a function of temporal spacing; and spectral density curves.

Abramson, J. Z., Hernández-Lloreda, M. V., García, L., Colmenares, F., Aboitiz, F., and Call, J. (2018). “Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*),” *Proc. R. Soc. B Biol. Sci.* **285**, 20172171.

Andrews, R. D., Pitman, R. L., and Ballance, L. T. (2008). “Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross Sea, Antarctica,” *Polar Biol.* **31**, 1461–1468.

Arnold, K., and Zuberbühler, K. (2008). “Meaningful call combinations in a non-human primate,” *Curr. Biol.* **18**, 202–203.

Arranz, P., DeRuiter, S. L., Stimpert, A. K., Neves, S., Friedlaender, A. S., Goldbogen, J. A., Visser, F., Calambokidis, J., Southall, B. L., and Tyack, P. L. (2016). “Discrimination of fast click-series produced by tagged Risso’s dolphins (*Grampus griseus*) for echolocation or communication,” *J. Exp. Biol.* **219**, 2898–2907.

Berwick, R. C., Okanoya, K., Beckers, G. J. L., and Bolhuis, J. J. (2011). “Songs to syntax: The linguistics of birdsong,” *Trends Cogn. Sci.* **15**, 113–121.

Blackwell, S. B., Tervo, O. M., Conrad, A. S., Sinding, M. H. S., Hansen, R. G., Ditlevsen, S., and Peter, M. (2018). “Spatial and temporal patterns of sound production in East Greenland narwhals,” *PLoS One* **13**, 1–23.

Bradbury, J. W., and Vehrencamp, S. L. (1998). “Sound propagation,” in *Principles of Animal Communication* (Sinauer Associates, Sunderland, MA), pp. 113–139.

Breed, G. A., Matthews, C. J. D., Marcoux, M., Higdon, J. W., LeBlanc, B., Petersen, S. D., Orr, J., Reinhart, N. R., and Ferguson, S. H. (2017). “Sustained disruption of narwhal habitat use and behavior in the presence of Arctic killer whales,” *Proceedings of the National Academy of Sciences* **114**(10), 2628–2633.

Brumm, H., and Slater, P. J. B. (2006). “Ambient noise, motor fatigue, and serial redundancy in chaffinch song,” *Behav. Ecol. Sociobiol.* **60**, 475–481.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., and Townsend, S. W. (2014). “Language evolution: Syntax before phonology?,” *Proc. R. Soc. B* **281**, 20140263.

Cook, M. L. H., Sayigh, L. S., Blum, J. E., and Wells, R. S. (2004). “Signature-whistle production in undisturbed free-ranging bottlenose dolphins (*Tursiops truncatus*),” *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 1043–1049.

Deecke, V. B., Slater, P. J. B., and Ford, J. K. B. (2002). “Selective habituation shapes acoustic predator recognition in harbour seals,” *Nature* **729**, 2001–2003.

Esch, H. C., Sayigh, L. S., and Wells, R. S. (2009). “Quantifying parameters of bottlenose dolphin signature whistles,” *Mar. Mammal Sci.* **25**, 976–986.

Ey, E., and Fischer, J. (2009). “The ‘acoustic adaptation hypothesis’—a review of the evidence from birds, anurans and mammals,” *Bioacoustics* **19**, 21–48.

Ferrer-i-Cancho, R., and McCowan, B. (2012). “The span of correlations in dolphin whistle sequences,” *J. Stat. Mech. Theory Exp.* **2012**, P06002.

Fitch, W. T., and Hauser, M. D. (2009). “Computational constraints on syntactic processing in a nonhuman primate,” *Science* **377**, 377–380.

Foote, A. D., Osborne, R. W., and Hoelzel, A. R. (2004). “Whale-call response to masking boat noise,” *Nature* **428**, 910.

- Ford, J. K. (1988). "Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia," *Can. J. Zool.* **67**, 727–745.
- Ford, J. K. B., and Fisher, H. D. (1978). "Underwater acoustic signals of the narwhal (*Monodon monoceros*)," *Can. J. Zool.* **56**(4), 552–560.
- Furgal, C., and Laing, R. (2012). "A synthesis and critical review of the traditional ecological knowledge literature on narwhal (*Monodon monoceros*) in the eastern Canadian Arctic fish," *Ocean. Canada, Can. Sci. Adv. Secr.*, Fisheries and Oceans Canada, Canadian science advisory secretariat, Vol. 3848, 47 pp.
- Gamer, M., Lemon, J., Fellows, I., and Singh, P. (2015). "Package irr: Various coefficients of interrater reliability and agreement," a package for R software (R Foundation for Statistical Computing, Vienna, Austria) <https://cran.r-project.org/web/packages/irr/index.html>.
- Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D., Poole, M. M., Robbins, J., and Noad, M. J. (2011). "Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale," *Curr. Biol.* **21**, 687–691.
- Garland, E. C., Lilley, M. S., Goldizen, A. W., Rekdahl, M. L., Garrigue, C., and Noad, M. J. (2012). "Improved versions of the Levenshtein distance method for comparing sequence information in animals' vocalisations: Tests using humpback whale song," *Behaviour* **149**, 1413–1441.
- Gero, S., Whitehead, H., and Rendell, L. (2016). "Individual, unit and vocal clan level identity cues in sperm whale codas," *R. Soc. Open Sci.* **3**, 150372.
- Gonzalez, N. (2001). *Inuit Traditional Ecological Knowledge of the Hudson Bay Narwhal (Tuugaalik) Population*, Fisheries and Oceans Canada, Iqaluit, Nunavut, Canada, pp. 1–26.
- Heiberger, R. M. (2018). *HH: Statistical Analysis and Data Display: Heiberger and Holland*, a package for R software, <https://rdr.io/cran/HH/>.
- Heide-Jørgensen, M. P. (2009). "Monodon monoceros," in *Encyclopedia of marine mammals*, 2nd ed., edited by W. F. Perrin, B. Würsig, and H. Theissen (Academic Press, San Diego, California), pp. 754–758.
- Henderson, E. E., Hildebrand, J. A., and Smith, M. H. (2011). "Classification of behavior using vocalizations of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*)," *J. Acoust. Soc. Am.* **130**, 557–567.
- Herman, L. M. (2006). "Intelligence and rational behaviour in the bottlenosed dolphin," in *Rational animals?*, edited by S. Hurley and M. Nudds (Oxford University Press, Oxford, UK), pp. 439–467.
- Janik, V. M. (2000). "Food-related bray calls in wild bottlenose dolphins (*Tursiops truncatus*)," *Proc. Biol. Sci.* **267**, 923–927.
- Janik, V. M. (2009). "Acoustic communication in delphinids," *Adv. Study Behav.* **40**, 123–157.
- Janik, V. M. (2013). "Cognitive skills in bottlenose dolphin communication," *Trends Cogn. Sci.* **17**, 157–159.
- Janik, V. M., and Sayigh, L. S. (2013). "Communication in bottlenose dolphins: 50 years of signature whistle research," *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **199**, 479–489.
- Janik, V. M., and Slater, P. J. B. (1998). "Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls," *Anim. Behav.* **56**, 829–838.
- Jensen, F. H., Beedholm, K., Wahlberg, M., Bejder, L., and Madsen, P. T. (2012). "Estimated communication range and energetic cost of bottlenose dolphin whistles in a tropical habitat," *J. Acoust. Soc. Am.* **131**, 582–592.
- Johnson, M., De Soto, N. A., and Madsen, P. T. (2009). "Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: A review," *Mar. Ecol. Prog. Ser.* **395**, 55–73.
- Johnson, M., Madsen, P. T., Zimmer, W. M. X., de Soto, N. A., and Tyack, P. L. (2006). "Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation," *J. Exp. Biol.* **209**, 5038–5050.
- Kaplan, M. B., Aran Mooney, T., Sayigh, L. S., and Baird, R. W. (2014). "Repeated call types in Hawaiian melon-headed whales (*Peponocephala electra*)," *J. Acoust. Soc. Am.* **136**, 1394–1401.
- Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., Deruiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T. M., Garland, E. C., Gustison, M., Harley, H. E., Huetz, C., Hughes, M., Hyland Bruno, J., Ilany, A., Jin, D. Z., Johnson, M., Ju, C., Karnowski, J., Lohr, B., Manser, M. B., McCowan, B., Mercado, E., Narins, P. M., Piel, A., Rice, M., Salmi, R., Sasahara, K., Sayigh, L., Shiu, Y., Taylor, C., Vallejo, E. E., Waller, S., and Zamora-Gutierrez, V. (2016). "Acoustic sequences in non-human animals: A tutorial review and prospectus," *Biol. Rev.* **91**, 13–52.
- Kershenbaum, A., Bowles, A. E., Freeberg, T. M., Jin, D. Z., Lameira, A. R., and Bohn, K. (2014). "Animal vocal sequences: Not the Markov chains we thought they were," *Proc. R. Soc. B Biol. Sci.* **281**(1792), 20141370.
- Kershenbaum, A., and Garland, E. C. (2015). "Quantifying similarity in animal vocal sequences: Which metric performs best?," *Methods Ecol. Evol.* **6**, 1452–1461.
- Kershenbaum, A., Ilany, A., Blaustein, L., and Geffen, E. (2012). "Syntactic structure and geographical dialects in the songs of male rock hyraxes," *Proc. R. Soc. B Biol. Sci.* **279**, 2974–2981.
- King, S. L., and Janik, V. M. (2015). "Come dine with me: Food-associated social signalling in wild bottlenose dolphins (*Tursiops truncatus*)," *Anim. Cogn.* **18**, 969–974.
- Koblitz, J. C., Stolz, P., Rasmussen, M. H., and Laidre, K. L. (2016). "Highly directional sonar beam of narwhals (*Monodon monoceros*) measured with a vertical 16 hydrophone array," *PLoS One* **11**(11), e0162069.
- Krebs, J. R., Aschcroft, R., and Orsdol, K. V. (1981). "Song matching in the great tit *Parus major*," *Anim. Behav.* **29**, 918–923.
- Laidre, K. L., and Heide-Jørgensen, M. P. (2011). "Life in the lead: Extreme densities of narwhals *Monodon monoceros* in the offshore pack ice," *Mar. Ecol. Prog. Ser.* **423**, 269–278.
- Laidre, K. L., Heide-Jørgensen, M. P., and Orr, J. R. (2006). "Reactions of narwhals, *Monodon monoceros*, to killer whale, *Orcinus orca*, attacks in the eastern Canadian Arctic," *Can. Field-Naturalist* **120**, 457–465.
- Lammers, M. I., and Oswald, J. I. (2015). "Analyzing the acoustic communication of dolphins," in *Dolphin communication and cognition: Past, present, and future*, edited by D. L. Herzog and C. M. Johnson (MIT Press, Cambridge, MA).
- Lammers, M. O., Au, W. W. L., Aubauer, R., and Nachtigall, P. E. (2004). "A comparative analysis of the pulsed emissions of free-ranging Hawaiian spinner dolphins," in *Echolocation Bats Dolphins*, edited by J. Thomas, C. Moss, and M. Vater (University of Chicago Press, Chicago, IL), pp. 414–419.
- Landis, J. R., and Koch, G. G. (1977). "The measurement of observer agreement for categorical data," *Biometrics* **33**, 159–174.
- Luis, A. R., Alves, I. S., Sobreira, F. V., Couchinho, M. N., and dos Santos, M. E. (2018). "Brays and bits: Information theory applied to acoustic communication sequences of bottlenose dolphins," *Bioacoustics* **46**(2), 1–11.
- Mann, J. (2000). "Unraveling the dynamics of social life," in *Cetacean Societies: Field Studies of Dolphins and Whales*, edited by J. Mann, R. C. Connor, P. L. Tyack, and H. Whitehead (University of Chicago Press, London), pp. 45–64.
- Mann, N. I., Dingess, K. A., and Slater, P. J. B. (2006). "Antiphonal four-part synchronized chorusing in a neotropical wren," *Biol. Lett.* **2**, 1–4.
- Marcoux, M., Auger-Méthé, M., and Humphries, M. M. (2012). "Variability and context specificity of narwhal (*Monodon monoceros*) whistles and pulsed calls," *Mar. Mammal Sci.* **28**, 649–665.
- Martin, M. J., Elwen, S. H., Kassanjee, R., and Gridley, T. (2019). "To buzz or burst-pulse? The functional role of Heaviside's dolphin, *Cephalorhynchus heavisidii*, rapidly pulsed signals," *Anim. Behav.* **150**, 273–284.
- May-Collado, L. J., and Wartzok, D. (2008). "A comparison of bottlenose dolphin whistles in the Atlantic Ocean: Factors promoting whistle variation," *J. Mammal.* **89**, 1229–1240.
- McDonald, M. A., Mesnick, S. L., and Hildebrand, J. A. (2006). "Biogeographic characterisation of blue whale song worldwide: Using song to identify populations," *J. Cetacean Res. Manag.* **8**, 55–65.
- Melo-Santos, G., Figueiredo Rodrigues, A. L., Tardin, R. H., de Sá Maciel, I., Marmontel, M., Da Silva, M. L., and May-Collado, L. J. (2019). "The newly described Araguaian river dolphins, *Inia araguaiaensis* (Cetartiodactyla, Iniidae), produce a diverse repertoire of acoustic signals," *PeerJ* **7**, e6670.
- Miller, L. A., Pristed, J., Møshl, B., and Surlykke, A. (1995). "The click-sounds of narwhals (*Monodon monoceros*) in Inglesfield Bay, Northwest Greenland," *Marine Mammal Science* **11**(4), 491–502.
- Miller, P. J. O., Biassoni, N., Samuels, A., and Tyack, P. L. (2000). "Whale songs lengthen in response to sonar," *Nature* **405**, 903.
- Miller, P. J. O., Shapiro, A. D., Tyack, P. L., and Solow, A. R. (2004). "Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*," *Anim. Behav.* **67**, 1099–1107.

- Morisaka, T. (2012). "Evolution of communication sounds in odontocetes: A review," *Int. J. Comp. Psychol.* **25**, 1–20.
- Morisaka, T., Yoshida, Y., Akune, Y., Mishima, H., and Nishimoto, S. (2013). "Exchange of 'signature' calls in captive belugas (*Delphinapterus leucas*)," *J. Ethol.* **31**, 141–149.
- Nowicki, S., and Searcy, W. A. (2004). "Song function and the evolution of female preferences: Why birds sing, why brains matter," *Ann. N. Y. Acad. Sci.* **1016**, 704–723.
- Quattara, K., Lemasson, A., and Zuberbühler, K. (2009). "Campbell's monkeys concatenate vocalizations into context-specific call sequences," *Proceedings of the National Academy of Sciences* **106**(51), 22026–22031.
- Papale, E., Perez-Gil, M., Castrillon, J., Perez-Gil, E., Ruiz, L., Servidio, A., Tejedor, M., Giacoma, C., and Martín, V. (2017). "Context specificity of Atlantic spotted dolphin acoustic signals in the Canary Islands," *Ethol. Ecol. Evol.* **29**, 311–329.
- Payne, R. J. H., and Pagel, M. (1997). "Why do animals repeat displays?," *Anim. Behav.* **54**, 109–119.
- Payne, R. S., and McVay, S. (1971). "Songs of humpback whales," *Science* **173**, 587–597.
- Petroni, F., and Serva, M. (2010). "Measures of lexical distance between languages," *Physica A: Statistical Mechanics and its Applications* **389**, 2280–2283.
- Pika, S., Wilkinson, R., Kendrick, K. H., and Vernes, S. C. (2018). "Taking turns: Bridging the gap between human and animal communication," *Proc. R. Soc. B Biol. Sci.* **285**, 20180598.
- Podos, J. (1996). "Motor constraints on vocal development in a songbird," *Anim. Behav.* **51**, 1061–1070.
- Quick, N., Callahan, H., and Read, A. J. (2018). "Two-component calls in short-finned pilot whales (*Globicephala macrorhynchus*)," *Mar. Mammal Sci.* **34**, 155–168.
- R Core Team (2014). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria.
- Rankin, S., Oswald, J., Barlow, J., and Lammers, M. (2007). "Patterned burst-pulse vocalizations of the northern right whale dolphin, *Lissodelphis borealis*," *J. Acoust. Soc. Am.* **121**, 1213–1218.
- Rasmussen, M. H., Koblitz, J. C., and Laidre, K. L. (2015). "Buzzes and high-frequency clicks recorded from narwhals (*Monodon monoceros*) at their wintering ground," *Aquat. Mamm.* **41**, 256–264.
- Raven Pro. (2007). "Raven Pro: Interactive Sound Analysis Software," (Cornell Lab of Ornithology, Ithaca, NY)
- Rendell, L., and Whitehead, H. (2001). "Culture in whales and dolphins," *Behav. Brain Sci.* **24**, 309–382.
- Rendell, L. E., and Whitehead, H. (2003). "Comparing repertoires of sperm whale codas: A multiple methods approach," *Bioacoustics* **14**, 61–81.
- Riesch, R., Ford, J. K. B., and Thomsen, F. (2006). "Stability and group specificity of stereotyped whistles in resident killer whales, *Orcinus orca*, off British Columbia," *Anim. Behav.* **71**, 79–91.
- Saulitis, E. L., Matkin, C. O., and Fay, F. H. (2005). "Vocal repertoire and acoustic behavior of the isolated AT1 killer whale subpopulation in southern Alaska," *Can. J. Zool.* **83**, 1015–1029.
- Sayigh, L., Quick, N., Hastie, G., and Tyack, P. (2013). "Repeated call types in short-finned pilot whales, *Globicephala macrorhynchus*," *Mar. Mammal Sci.* **29**, 312–324.
- Scott-Phillips, T. C., Gurney, J., Ivens, A., Diggle, S. P., and Popat, R. (2014). "Combinatorial communication in bacteria: Implications for the origins of linguistic generativity," *PLoS One* **9**, e95929.
- Shapiro, A. D. (2006). "Preliminary evidence for signature vocalizations among free-ranging narwhals (*Monodon monoceros*)," *J. Acoust. Soc. Am.* **1695**, 1695–1705.
- Shettleworth, S. J. (2010). *Cognition, Evolution, and Behaviour*, 2nd ed. (Oxford University Press, New York).
- Sloan, J. L., and Hare, J. F. (2004). "Monotony and the information content of Richardson's ground squirrel (*Spermophilus richardsonii*) repeated calls: Tonic communication or signal certainty?," *Ethology* **110**, 147–156.
- Smolker, R. A., Mann, J., and Smuts, B. B. (1993). "Use of signature whistles during separations and reunions by wild bottlenose dolphin mothers and infants," *Behav. Ecol. Sociobiol.* **33**, 393–402.
- Stafford, K. M., Laidre, K. L., and Heide-Jørgensen, M. P. (2012). "First acoustic recordings of narwhals (*Monodon monoceros*) in winter," *Mar. Mammal Sci.* **28**, E197–E207.
- Templeton, C. N., Greene, E., and Davis, K. (2005). "Allometry of alarm calls: Black-capped chickadees encode information about predator size," *Science* **308**, 1934–1938.
- Vaughn-Hirshorn, R. L., Hodge, K. B., Würsig, B., Sappenfield, R. H., Lammers, M. O., and Dudzinski, K. M. (2012). "Characterizing dusky dolphin sounds from Argentina and New Zealand," *J. Acoust. Soc. Am.* **132**, 498–506.
- Vergara, V., Michaud, R., and Barrett-Lennard, L. G. (2010). "What can captive whales tell us about their wild counterparts? Identification, usage, and ontogeny of contact calls in belugas (*Delphinapterus leucas*)," *Int. J. Comp. Psychol.* **23**, 278–309.
- Waser, P. M. (1982). "The evolution of male loud calls among mangabeys and baboons," in *Primate Communication*, edited by C. T. Snowdon, C. H. Brown, and M. R. Petersen (Cambridge University Press, New York), pp. 117–143.
- Watkins, W. A., and Schevill, W. E. (1977). "Sperm whale codas," *J. Acoust. Soc. Am.* **62**, 1485–1490.
- Wiggins, S. M. (2013). *Acousonde Acoustic Tag Calibration at the Transducer Evaluation Center (TRANSDEC)*, a technical report (La Jolla, California), pp. 1–28.
- Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., and Tyack, P. L. (2005). "Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*)," *J. Acoust. Soc. Am.* **117**, 3919–3927.
- Zuberbühler, K. (2018). "Combinatorial capacities in primates," *Curr. Opin. Behav. Sci.* **21**, 161–169.
- Zwamborn, E. M. J., and Whitehead, H. (2017). "Repeated call sequences and behavioural context in long-finned pilot whales off Cape Breton, Nova Scotia, Canada," *Bioacoustics* **26**, 169–183.