

## Redefinition and Sexual Difference of Contact Calls in Belugas (*Delphinapterus leucas*)

Yuka Mishima,<sup>1,2,3</sup> Tadamichi Morisaka,<sup>3,4,5</sup> Yuki Mishima,<sup>6</sup>  
Tadashi Sunada,<sup>6</sup> and Yoshinori Miyamoto<sup>2</sup>

<sup>1</sup>Atmosphere and Ocean Research Institute, The University of Tokyo  
5-1-5, Kashiwanoha, Kashiwa-shi, Chiba, 277-8564, Japan

<sup>2</sup>School of Marine Resources and Environment, Tokyo University of Marine Science and Technology  
4-5-7, Konan, Minato-ku, Tokyo, 108-8477, Japan

<sup>3</sup>Department of Environmental and Societal Affairs, School of Marine Science and Technology, Tokai University  
3-20-1, Orido, Shimizu-ku, Shizuoka-shi, Shizuoka, 424-8610, Japan

<sup>4</sup>Institute of Oceanic Research and Development, Tokai University  
3-20-1, Orido, Shimizu-ku, Shizuoka-shi, Shizuoka, 424-8610, Japan

<sup>5</sup>Cetacean Research Center, Graduate School of Bioresources, Mie University  
1577, Kurimamachiya-cho, Tsu-shi, Mie, 514-8507, Japan

E-mail: chaka@bio.mie-u.ac.jp

<sup>6</sup>Shimane Aquarium 1117-2, Kushiro-cho, Hamada-shi, Shimane, 697-0004, Japan

### Abstract

Previous studies have suggested that belugas (*Delphinapterus leucas*) use pulsed calls both with and without tone-like components for contact calls. However, call classification differs among researchers, and the definition of contact calls remains ambiguous. The objective of this study is to organize and integrate this information to redefine the contact calls of belugas. Our previous studies termed their contact call as *PS1*, which was exchanged among the belugas and was the predominant call type in isolation. *PS1* is a broadband pulsed call that sounds like a door creaking and has a duration of 0.15 to 1.5 s. Individual distinctiveness was found within a typical pulse repetition pattern of *PS1*. The *PS1* characteristics initially described were based on one captive population. In this study, calls from another population of seven belugas of both sexes and various ages at Shimane Aquarium, Japan, were recorded from October 2014 to March 2015. The *PS1* definition was expanded to broadband pulsed calls continuing for > 0.15 s, and the *PS1* calls were explored from their calls. The belugas exchanged *PS1* calls, but the pulse repetition pattern had various forms instead of the typical pattern suggested in previous *PS1* studies. Additionally, all the *PS1* calls contained a tone-like component. By reflecting on these results and referring to other previous studies, we renamed these contact calls *creaking calls*. Moreover, this study shows sex differences in the creaking calls. The females and juvenile

male had creaking calls with an individually distinct pulse repetition pattern, while adult males had variations in pattern during free swimming. Each adult male, however, used only an individually distinct stereotype of pulse repetition pattern in a visual reunion and first-sighting context. This suggests that adult males have individualized and non-individualized creaking calls, and they use the former to advertise identity in separation, reunion, or greeting contexts.

**Key Words:** acoustic communication, vocal exchange, individuality, repertoire, classification, Delphinidae, cetacean

### Introduction

Social animals exchange *contact calls* to maintain group cohesion (Kondo & Watanabe, 2009). The vocal exchange is characterized as a call sequence when the preceding call is followed by the call-back of another individual within a particular time window (Sugiura, 1993). Calls produced by the receiver within the time window are considered to be responses. Callers wait for replies from other individuals during this time window. In this manner, a temporal rule to regulate vocal exchange exists in the contact calls of several species, including marine mammals such as sperm whales (*Physeter macrocephalus*; Schulz et al., 2008), bottlenose dolphins (*Tursiops truncatus*; Nakahara & Miyazaki, 2011), and beluga whales (*Delphinapterus leucas*; Vergara et al., 2010; Morisaka et al., 2013).

The type of information embedded in contact calls to discriminate associates is linked to the social organization of the organism. For example, bottlenose dolphins with fluid fission–fusion societies use contact calls termed *signature whistles* (Caldwell & Caldwell, 1965; Caldwell et al., 1990). These signature whistles encode strongly recognizable individual identities independent of the voice features (Janik et al., 2006; Sayigh et al., 2017). On the other hand, sperm whales and killer whales (*Orcinus orca*) with stable societies encode hierarchical information in their pulse-type contact calls: the strongly recognizable group identity and less prominent individual identity (Ford, 1989, 1991; Rendell & Whitehead, 2003, 2004; Nousek et al., 2006; Antunes et al., 2011; Gero et al., 2016; Oliveira et al., 2016).

Previous studies on the relationship between calls and behavioural activity of beluga whales indicate they use a diverse repertoire of call types, including whistles, pulsed calls, and mixed calls of a pulsed component and tonal or secondary pulsed component, for communication (Sjare & Smith, 1986a; Karlsen et al., 2002; Panova et al., 2012; Alekseeva et al., 2013). It has been suggested that some types of pulsed and mixed calls might function as contact calls. Van Parijs et al. (2003) collected calls from temporarily captured belugas in Svalbard, Norway. During the capture event of a mother–calf pair, they were kept in close contact with each other, but the mother produced pulsed calls many times and frequently moved her head towards the calf while producing calls. The calf also produced many pulsed calls which occasionally contained a tone. It suggests that they were using these calls to maintain acoustical contact with each other in the stressful situation.

Vergara & Barrett-Lennard (2008) and Vergara et al. (2010) recorded calls from captive belugas at the Vancouver Aquarium in Canada. The belugas originated from the Churchill River Estuary in Canada. The authors reported the emission of *Type A* calls between mothers and calves. These calls were pulsative in nature and sounded like a door creaking, and some variants of the calls contained a tonal component or a secondary pulsed component that had a different pulse repetition rate from the main pulse train. The *Type A* calls were often produced in sequence by the mother and her calf within 2 s, suggesting that they exchanged the calls following the temporal rule that responders call back within 2 s (Vergara et al., 2010). In addition, the mother and calf often produced *Type A* calls in separation contexts. Moreover, the mother predominantly produced *Type A* calls the day after giving birth and after the death of a calf (Vergara & Barrett-Lennard, 2008; Vergara et al., 2010). The description of *Type A* calls by Vergara

et al. (2010) from temporarily restrained belugas in the Nelson River Estuary and social groups in the St Lawrence Estuary in Canada suggest these calls are widely used by belugas. Chmelnitsky & Ferguson (2012) also described call types similar to *Type A* in belugas from the Churchill River Estuary. When the calls were recorded, at least one mother–calf pair was often observed. The findings also suggest that these calls serve as contact calls between mothers and calves.

Morisaka et al. (2013) and Mishima et al. (2015) recorded calls from belugas housed at the Port of Nagoya Public Aquarium in Japan. The belugas originated from the White Sea and the Russian Far East. The studies focused on one type of pulsed call, the *PS1*, which was defined acoustically as follows: the fixed pulse train sounds like a ratchet or a door creaking, and it is easy for humans to discriminate it as a *PS1* call. Energy is distributed in a broad band from less than 1 kHz up to at least 170 kHz, and the duration is more than 0.15 s. *PS1* calls had a typical pulse repetition pattern structure: the inter-pulse intervals decrease at first, become constant, and then rapidly increase at the end of the call. Because an individual produced tones with pulse trains resembling *PS1* calls, some mixed calls were occasionally considered to be *PS1* calls (Mishima et al., 2015). Morisaka et al. (2013) revealed that the *PS1* calls were exchanged following the temporal rule that responders call back within approximately 1 s, callers waited for responses within the 1-s time window, and the pulse repetition pattern of *PS1* calls was different among the three adults.

Because the study analysed a small dataset, Mishima et al. (2015) subsequently investigated the *PS1* characteristics of five belugas of both sexes and different ages at the Port of Nagoya Public Aquarium. The study used a separation context to elicit contact calls from the belugas. Each beluga was transferred to a separate pool alone with the exception of one male calf who was transferred with his mother or a subadult female. Thirty-minute recordings were made of the separation context for each beluga. *PS1* was the most frequently produced call type in the separation context, and the *PS1* call rate increased significantly when a mother–calf pair was separated. In addition, the pulse repetition pattern of a *PS1* call was individually distinctive and stereotyped, and some temporal and spectral parameters were also different among individuals. These two previous *PS1* studies suggest that *PS1* functions as a contact call among individuals, as well as between mothers and calves, and the pulse repetition pattern of *PS1* appeared to contain individual information.

Panova et al. (2017) showed additional evidence that belugas encode individual information in the

pulse repetition pattern of their contact calls. The study investigated stereotyped pulsed and mixed calls of captive belugas that originated from the Okhotsk Sea, and the calls were defined as similar calls to Type A and PS1 calls. The stereotyped calls were divided into types according to the pulse repetition pattern of the pulsed components, as well as the frequency contour of the tonal components. A newly introduced female in the Koktebel Dolphinarium in Ukraine produced only one type of stereotyped mixed call. In addition, four types of stereotyped pulsed and mixed calls were recorded from four females hosted together in Nilmaguba, Russia, when they interacted with underwater trainers, without disturbances, and when two of the females were removed. These results suggest that each type of stereotyped call may have belonged to one female; in other words, the females could have individually distinctive stereotyped pulsed or mixed calls. The type of information encoded in contact calls of belugas may have been selected in beluga societies with long-term associations in a fluid social structure (Bel'kovitch & Sh'ekotov, 1993; Michaud, 2005; Colbeck et al., 2012).

The pulsed calls in the Van Parijs et al. (2003) study, Type A calls (Vergara & Barrett-Lennard, 2008; Vergara et al., 2010), and PS1 calls (Morisaka et al., 2013; Mishima et al., 2015) share some acoustical and functional characteristics, and those characteristics are found in some call types of other previous studies on free-ranging belugas (Sjare & Smith, 1986a, 1986b; Recchia, 1994; Karlsen et al., 2002; Belikov & Bel'kovich, 2008; Chmelnitsky & Ferguson, 2012; Panova et al., 2012; Alekseeva et al., 2013; reviewed in Mishima et al., 2015). However, call classification was different among the studies, and the definition of contact calls in belugas remained ambiguous.

The objective of the present study is to solidly define the contact calls of belugas. The definition of PS1 was based on just one captive population; therefore, the present study collected calls from another captive beluga population and explored PS1 calls to increase the population size. Taking the results of the present and previous PS1 studies together and comparing them to the pulsed and mixed calls described in other principal studies on beluga contact calls (Van Parijs et al., 2003; Vergara & Barrett-Lennard, 2008; Vergara et al., 2010; Panova et al., 2017), we provide a new definition of the generalized contact call in belugas. In addition, Panova et al. (2017) showed that two captive male belugas of Nilmaguba, Russia, shared one type of stereotyped mixed call in interactions with underwater trainers; we therefore conducted further investigations of the contact call usage of males.

Two experiments were conducted in the present study. Experiment 1 examined the vocal exchange

pattern, acoustic structure, and individuality of the PS1 calls. In Experiment 2, a visual reunion and first-sighting context was provided for the adult males to look for the existence of individually specific PS1 calls in their repertoires.

## Methods

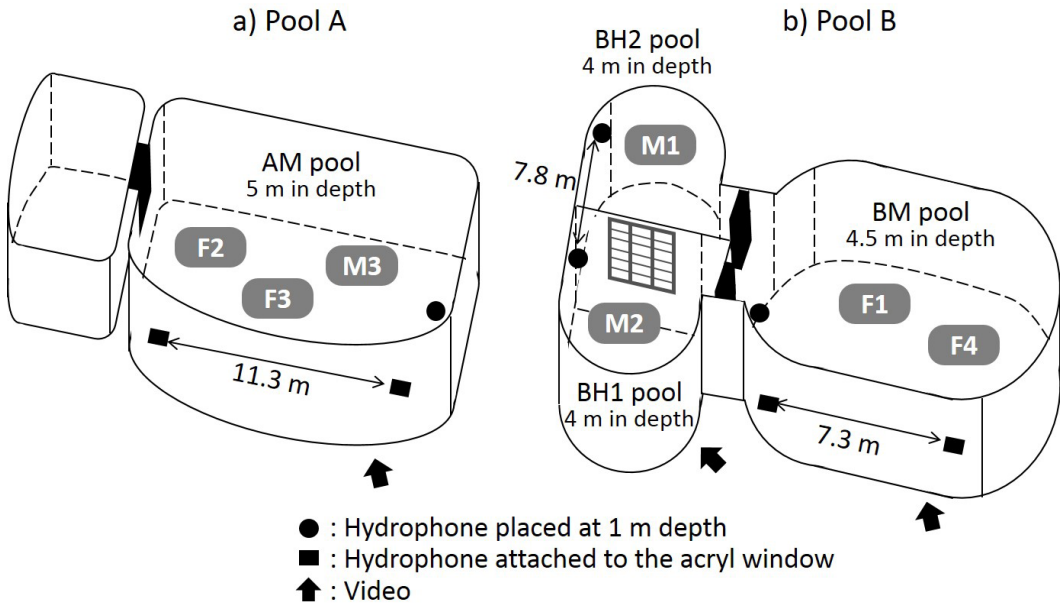
### *Experiment 1*

*Facility and Subjects*—Data were collected from the belugas housed at the Shimane Aquarium in Shimane, Japan, from October 2014 to March 2015. There were seven belugas: three female adults (F1, F2, and F3), one female calf (F4), two male adults (M1 and M2), and one male juvenile (M3). All adult belugas were captured in the Amur River in Russia and estimated to be 16 to 18 y old. F1, F2, and M2 came to the aquarium in 1999, and F3 and M1 in 2003. Both M3 and F4 were born in captivity from the same parents (F1 and M2). M3 was 5 y old and F4 was 2 mo old in October 2014.

There were two beluga pools, A and B. Pool A was composed of two subpools, and the main pool (AM pool) held F2, F3, and M3 (Figure 1a). Pool B was composed of three subpools: a main pool (BM pool) and two holding pools (BH1 and BH2 pools) (Figure 1b). The mother-calf pair (F1 and F4) was held in the BM pool. Each holding pool contained either M1 or M2. There was a metal lattice between the BH1 and BH2 pools, maintaining visual and acoustical contact between individuals M1 and M2. There were gates between the BH1 and BM pools, and between the BH2 and BM pools. These gates prevented visual contact, but sounds could pass through the gates, although they were attenuated.

There were three recording patterns: the AM, BM, and BH recordings were conducted in AM, BM, and BH pools, respectively. Each recording continued for 30 min, and a total of 12, 10, and six sessions were performed for AM, BM, and BH recordings, respectively. All sessions were performed in normal free-swimming contexts between 0900 and 1700 h, excluding feeding and training times.

*Recording System*—Three hydrophones were used for the AM and BM recordings (Figure 1). One was a TC 4013 underwater hydrophone (Reson Inc., Slangerup, Denmark) which exhibits a flat frequency response from 1 Hz to 170 kHz ( $211 \pm 3$  dB re 1V/ $\mu$ Pa) and was placed at a depth of 1 m. The hydrophone was housed in a polyvinyl chloride pipe, which did not prevent recordings because the pipe had a small window where the tip of the hydrophone was located. The other hydrophones used were Model AQH-100DTP (AquaSound Inc., Kobe, Japan). These hydrophones were attached to an acrylic observation window using grease.



**Figure 1.** Schematic view of the beluga pools in the Shimane Aquarium: (a) Pool A and (b) Pool B. Recordings were made at AM pool (the main pool of Pool A), BM pool (the main pool of Pool B), and BH pools (the holding pools of Pool B). The arrangements of the hydrophones and videos are shown in the figures.

The hydrophones collected the sounds generated under water behind the acrylic window from 3 to 130 kHz, although there were sound transmission losses between 3 to 40 kHz and gains between 50 to 130 kHz (Yoshida et al., 2016). The TC 4013 hydrophone was used for the analysis of acoustic parameters, and the two AQH-100DTP hydrophones were used to identify callers by measuring the time differences of sound arrivals. The three hydrophones installed in the BM pool also recorded the sounds from the males held in the BH pools, although the sound recordings were of a poor acoustic quality. The male calls were counted, but those PS1 callers could not be identified.

The BH recordings used two TC 4013 hydrophones submerged at a depth of 1 m (Figure 1b). They were used for the analysis of acoustic parameters and the identification of callers. The BH recordings also collected sounds from the mother–calf pair in the BM pool. The PS1 calls with poor acoustic quality were considered to be produced by the mother F1 since no PS1 calls were heard from the calf F4 during the experiment period.

The sound from TC 4013 hydrophones was filtered using an analog high-pass filter at 1 kHz and amplified by 32 dB using VP1000 preamplifiers (Reson Inc.). The sound from AQH-100DTP hydrophones was filtered using an

analog band-pass filter from 1 to 200 kHz and amplified by 50 dB using Aquafeeler III preamplifiers (AquaSound Inc.). All analog data were collected by EZ7510 data recorders (NF Corp., Yokohama, Japan), which digitized up to two channels of sound sampling at 500 kHz and 16 bits. Observations were made from underwater windows using a GZ-V675-R video camera (JVC Corp., Yokohama, Japan).

**PS1 Call Definition**—The definition of PS1 calls in the previous studies needed to be expanded to search for PS1 calls from another beluga population; therefore, the pulse repetition pattern structure properties defined in the previous studies in which “the inter-pulse intervals (IPIs) decrease at first, become constant, and then rapidly increase at the end of the call” (Morisaka et al., 2013; Mishima et al., 2015) were excluded from the definition of PS1 calls in the present study. PS1 calls were still defined here as fixed pulse trains that sounded like a ratchet or a door creaking. The energy had a broadband distribution from less than 1 kHz up to at least 170 kHz, and the duration was more than 0.15 s. Based on the updated definition, one of the authors (YM) found PS1 calls in the recordings using *Audacity*, Version 2.0.5 (The Audacity Team).

**Analysis for PS1 Call Exchange**—To test the prediction that PS1 calls produced by the belugas in the



present study were also used for vocal exchanges, as observed in the previous study (Morisaka et al., 2013), the between-individual intervals (BIIs) and within-individual intervals (WIIs) of the PS1 calls were investigated. An interval in this study was defined as the period from the peak of the first pulse in the preceding PS1 call to the peak of the first pulse in the following PS1 call. The interval criterion was different from the Morisaka et al. (2013) study, which defines an interval as the period from the end of a PS1 to the beginning of the next PS1 because there were many overlapping PS1 calls produced by two or three belugas in the present study. It was easier to read the beginning of the PS1 calls than the end of the PS1 calls in those call sequences. In cases where intervals could not be measured correctly and could not be classified as either BII or WII, the samples were excluded from the interval data. BII and WII frequency distributions were constructed for each of the AM and BH recordings, and a temporal production pattern was examined. BM recordings were excluded in this analysis because in BM recordings, the callers located in the BH pools could not be identified, and it was difficult to classify most intervals as either BII or WII.

A simulation for BII was performed using a bootstrapping technique (Yosida et al., 2007; Kondo et al., 2010) to investigate whether the observed BII distribution showed a temporal rule of vocal exchange or was just an incidental result. In the simulation, imaginary individuals produced PS1 calls independently at their own pace. First, a number of samples were randomly selected from the observed WII data pool using the average number of observed WIIs per individual and per session as the specific number (8 and 10 samples for AM and BH recordings, respectively), and the selected WII samples were rearranged. The value of the first WII sample was regarded as the onset time of the first PS1 call produced by an imaginary individual. The cumulative sum of the first and second WII samples was considered as the onset time of the second PS1 call produced by the imaginary individual. Similarly, all onset times of PS1 calls were calculated. The procedure was repeated three times for three imaginary individuals, which was the same as the number of individuals engaged in PS1 call production in Pools A and B because F4 did not produce PS1 calls. The imaginary individuals were presumed to independently produce PS1 calls in parallel. The simulated BII was defined as the time difference in the onset times of PS1 call sequences by different imaginary individuals. This procedure was repeated until we obtained a comparable number of simulated and observed BII samples. The frequency distributions between the observed and simulated BIIs were compared using two-sample Kolmogorov-Smirnov tests.

*Acoustic Parameter Extraction from PS1 Calls*—The acoustic characteristics of PS1 calls with good signal-to-noise ratios were analysed using *Avisoft SASLab Pro*, Version 5.2 (Avisoft Bioacoustics, Glienicke/Nordbahn, Germany). PS1 was composed of many pulses, and the pulses were automatically detected; however, in cases where reflecting pulses had been counted or where direct pulses with lower amplitudes had not been counted, the miscounted pulses were corrected manually. Five temporal parameters were extracted: (1) the number of pulses ( $N_p$ ), (2) duration of call (DUR), (3) pulse repetition rate (PRR), (4) mean IPI of pulse numbers 11 to 20 (IPI 1), and (5) mean IPI of pulse numbers 11 to 20 from the final pulse of the call (IPI 2). The two parameters of IPIs, IPI 1 and IPI 2, were selected according to the previous study (Mishima et al., 2015), which will allow us to compare statistically the pulse repetition patterns of PS1 calls. The changes in IPIs as a function of time, termed *IPI contours*, were also depicted with time in seconds as the x axis and IPIs in milliseconds as the y axis.

The spectral characteristics did not seem to change among pulses within a call as suggested in Mishima et al. (2015); therefore, the middle pulses were selected as representative of pulses consisting of PS1 calls for spectral analyses. The averaged power spectra of the middle pulses were calculated by a 256-point FFT with a Hamming window and then smoothed using a five-point window. The maximum source level (SL) above 1 kHz was set to zero to compare relative spectra. There were no noise effects on the spectra of PS1 calls since noise spectra calculated using non-call windows before the onset of the PS1 calls had no distinct noise above 1 kHz. Four spectral parameters were calculated: (1) peak frequency of the middle pulse ( $F_p$ ); (2) 10 dB bandwidth (10 BW), which is the frequency band at a level of -10 dB from the peak; and (3 & 4) the lower and upper frequencies of the 10 BW ( $F_l$  and  $F_u$ ).

*IPI Contour Classification by Human Observers*—Using the visual classification methods by human observers in Janik (1999) and Sayigh et al. (2007), the individual distinctiveness of IPI contours was investigated. Six IPI contours were randomly selected for each of the six belugas. The graphic scales of all of the 36 samples were set to 0 to 1.2 s for x axes (1 s = 15 cm) and 0 to 14 ms for y axes (1 ms = 0.6 cm) to fit the maximum size of samples. The samples were printed without the axis legends, scales, and gridlines. Five observers who had no prior experience in classifying animal sounds categorized the randomly ordered samples into classes based on the overall contour similarity with no information on the appropriate number of classes. If at least three out of five observers

agreed on a pair to be grouped together, the pair was regarded as the same call type produced by the same individual or a similar call type produced by different individuals.

*Statistical Analyses of Acoustic Parameters*—The five temporal parameters ( $N_p$ , DUR, PRR, IPI 1, and IPI 2) and four spectral parameters ( $F_p$ , 10 BW,  $F_i$ , and  $F_u$ ) were compared inter-individually by using the Kruskal-Wallis test or one-way ANOVA. To identify which parameters possibly carry individual information, the potential for individual coding (PIC) was calculated. The PIC is the ratio of the coefficient of variation between individuals to that within individuals; thus, acoustic parameters showing a PIC score greater than 1 may be a useful feature for the recognition of individuals (Charrier et al., 2002). Furthermore, a discriminant function analysis (DFA) was performed to classify PS1 calls into individuals based on the acoustic variables with a high PIC score. To decrease the disparity in sample sizes and to increase the effectiveness of the DFA before running it (McGarigal et al., 2000), 53 samples were randomly selected from 100 PS1 calls of F2, which was the same number used for F1, and the sample size of each individual fell within the range of 21 to 53. Because the dataset did not satisfy multivariate normality (Shapiro-Wilk test,  $p < 0.01$ ) and homogeneity of the variance-covariance matrices (Box's M test,  $p < 0.01$ ), a quadratic DFA was selected. The predominant discriminator was found by a stepwise DFA. All statistical analyses were performed using *R*, Version 3.1.0 (The R Foundation for Statistical Computing, Vienna, Austria), and statistical significance was accepted when  $p < 0.01$ .

### Experiment 2

*Facility and Subjects*—Experiment 1 revealed that only the adult males had several types of PS1 calls; therefore, Experiment 2 was conducted to determine the existence of individually specific PS1 calls in their repertoires. Experiment 2 was carried out in Pool B at the Shimane Aquarium in March 2015. As described in Experiment 1, Pool B contained three adults, M1, M2, and F1, and one calf, F4 (Figure 1b). The mother-calf pair (F1 and F4) was held in the BM pool, and each holding pool contained either M1 or M2. A metal lattice divided the BH1 and BH2 pools, while gates divided the BH1 and BM pools and the BH2 and BM pools. M1 and M2 could communicate visually and acoustically with each other. The males could not communicate with the mother-calf pair visually, but they could communicate acoustically, although the sounds were attenuated.

In Experiment 2, the gates between the BM pool and BH pools were opened while the metal lattice remained in place, meaning the individuals could not move between the subpools. This was

a visual reunion after a long separation for the males and F1, and a first opportunity for the males to see the calf F4.

*Recording System*—Sounds were recorded using all the hydrophones used in the BM and BH recordings of Experiment 1 to identify callers. The equipment and settings were the same as in Experiment 1. Behaviour was recorded from the underwater windows using a GZ-V675-R video camera.

*Context-Specific Use of PS1 Calls in Adult Males*—The adult belugas in Pool B were already known to each other, and acoustic communication was made possible through the closed gates. However, if the adult males had individualized PS1 calls used for identity advertisement in their PS1 call repertoire, the rate of the individualized PS1 calls might increase with the gate open, where both acoustical and visual contact were fully available to verify individual identification. Recordings were classified into three sections: (1) the 12 min before the gate opens, (2) during the 12-min gate opening, and (3) for 12 min after the gate opens. First, the number of PS1 calls by each individual was investigated in each recording section. Second, focusing on the PS1 calls produced by the adult males during the 12-min gate opening, the IPI contours were depicted for calls with good acoustic quality.

To investigate whether the adult males use an individually distinctive stereotyped IPI contour within their repertoire during the gate opening, a classification was performed by human observers. We arranged IPI-contour samples from the four datasets: (1) all M1 samples during free swimming of Experiment 1, (2) all M1 samples during the gate opening, (3) all M2 samples during free swimming of Experiment 1, and (4) all M2 samples during the gate opening. All of the samples were printed with x axes of 0 to 2.6 s (1 s = 15 cm) and y axes of 0 to 20 ms (1 ms = 0.6 cm) to fit the maximum size of samples, and they were randomly ordered. Five naïve observers without prior experience in classifying animal sounds, who were different from the observers in Experiment 1, were instructed to categorize the samples into a number of classes as she or he thought best based on the overall contours. On each pair of all samples, the number of observers who grouped the pair together were counted. The percentage of the number of pairs grouped together by at least three observers to the number of total pairs were calculated within each dataset and between datasets to examine the relationship of IPI contours within and between contexts and individuals. Finally, the duration of the individually distinctive stereotyped PS1 calls was compared between the two different contexts using the

two-sample *t* test or the Welch two-sample *t* test to verify the previously suggested prediction that the duration varies depending on contexts (Mishima et al., 2015).

## Results

### Experiment 1

In total, 965, 862, and 421 PS1 calls were collected from 6-h AM (the main pool of Pool A), 5-h BM (the main pool of Pool B), and 3-h BH (two holding pools of Pool B) recordings, respectively. All belugas, excluding the calf F4, produced PS1 calls. Of the 965 PS1 calls in the AM recordings, 867 calls were from identified callers, including 489, 211, and 167 from F2, F3, and M3, respectively. Of the 862 PS1 calls in the BM recordings, 220 calls were identified as calls from F1. All of the 421 PS1 calls in the BH recordings were identified with regard to callers, including 242, 108, and 71 from M1, M2, and F1, respectively, but the F1 samples in the BH recordings are excluded from this analysis because of poor acoustic quality.

**Temporal Rule of PS1 Call Exchange**—Measured intervals of the PS1 calls were divided into BII (between-individual interval) and WII (within-individual interval). The frequency distribution of the BIIs and WIIs in the AM and BH data can be observed in Figure 2. In the AM recordings, the observed WII distribution had a sharp peak at an interval of 1.6 s, while the observed BII distribution had two peaks, with a sharp peak at 0.6 s and a small peak at 1.2 s. These peaks indicate clear differences in duration between the BIIs and WIIs, with 95.5% of the WIIs occurring after 1.0 s and 50.4% of the BIIs occurring within 1.0 s. It was found that 20.2% of the BIIs were intervals of overlapping call sequences. In the BH recordings, the observed WII distribution showed two

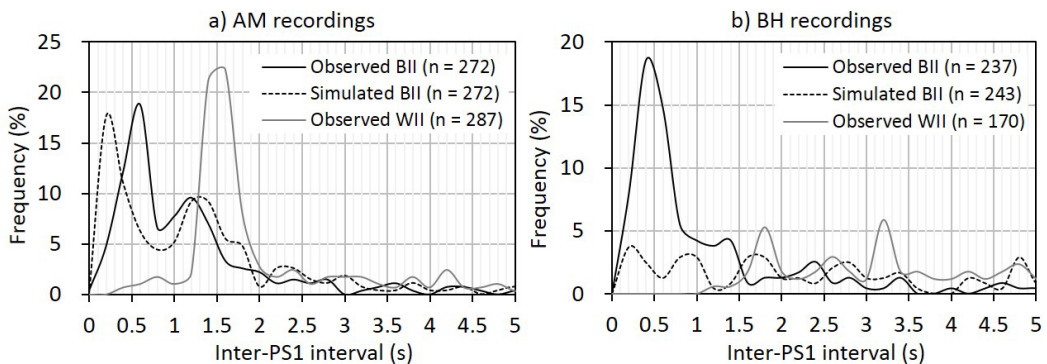
small peaks at intervals of 1.8 and 3.2 s, while the observed BII distribution had a sharp peak at 0.4 s. All of the WIIs occurred after 1.0 s, whereas 51.1% of the BIIs occurred within 1.0 s, and 47.3% of the BIIs were intervals of overlapping call sequences.

The simulated BII distribution of AM recordings had two peaks: (1) a sharp peak at 0.2 s and (2) a small peak at 1.3 s (Figure 2a). Although the second peak was similar, the first peak was different from that of the observed BII distribution. On the other hand, the simulated BII distribution of BH recordings had no particular peaks (Figure 2b). The simulated and observed BIIs had significantly different frequency distributions in both the AM and BH recordings (two-sample Kolmogorov-Smirnov tests:  $D = 0.143$ ,  $p < 0.01$  and  $D = 0.447$ ,  $p < 0.01$ , respectively).

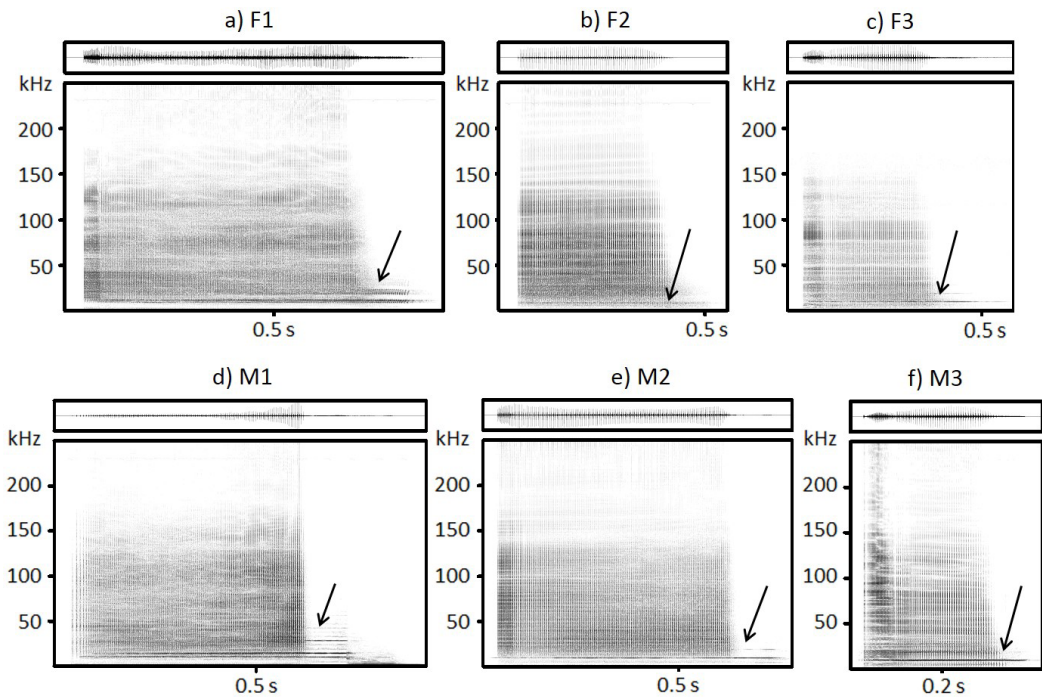
**Acoustic Characteristics and Individuality of PS1 Calls**—Acoustic parameters were extracted from PS1 calls with good signal-to-noise ratios. The sample numbers were 53, 100, 21, 38, 24, and 32 for F1, F2, F3, M1, M2, and M3, respectively. Since F2 had several samples with good acoustic quality, 100 samples were randomly selected.

Spectrogram examples of PS1 calls are shown in Figure 3. All PS1 calls contained a tone-like component: a tonal component or a secondary pulsed component composed of low-frequency, narrow-band, and high-repetition-rate pulses, which overlap in both the temporal and frequency domains. The tone-like components were not modulated over the call duration in all belugas. The dominant frequencies of the components were  $11.3 \pm 0.3$  (mean  $\pm$  SD) kHz for F1,  $7.8 \pm 0.1$  kHz for F2,  $9.2 \pm 0.1$  kHz for F3,  $9.7 \pm 0.5$  kHz for M1,  $8.5 \pm 0.8$  kHz for M2, and  $9.1 \pm 0.1$  kHz for M3.

Fifteen IPI contours of the main pulse trains in PS1 calls were randomly selected for each individual (Figure 4). Each beluga had individually



**Figure 2.** Frequency distributions of the intervals between successive PS1 calls. BII = between-individual interval and WII = within-individual interval. The percentages were calculated for each of the observed BIIs (black solid line), simulated BIIs (black broken line), and observed WIIs (grey solid line) in (a) AM recordings and (b) BH recordings (bin width: 0.2 s).



**Figure 3.** Examples of PS1 calls from six belugas: (a) F1, (b) F2, (c) F3, (d) M1, (e) M2, and (f) M3. The top graphs represent waveforms, and the bottom graphs represent spectrograms (FFT size: 1,024 points; window: Hamming; overlap: 50%). The arrows indicate tone-like components co-occurring with the main pulse trains.

distinctive, stereotyped IPI contours, excluding the adult males, M1 and M2, which had variations in the IPI contours. All individuals tended to have abrupt changes of IPIs within the first 0.1 s of the IPI contour.

The result of the classification of IPI contours of PS1 calls by human observers are visually illustrated as similarity networks in Figure 5. The IPI contours from each of F1, F2, F3, and M3 were found to form an independent cluster, with the exception of one M3 sample (M3-1), which was grouped together with F3 samples (F3 was a pool mate of M3 but not his mother). The samples of M1 and M2 were split, and there was a mixed cluster of the M1 and M2 samples as well as an M1 cluster and an M2 cluster; however, none of the M1 and M2 samples were grouped together with other beluga samples.

The averaged power spectra of the middle pulses within PS1 calls were calculated, and 15 randomly selected samples for each individual are shown in Figure 6. Contrary to IPI contours, there was no obvious individual distinctiveness and consistency.

The characteristics and results of statistical analyses for each acoustic parameter of PS1 calls are summarized in Table 1. Univariate statistical

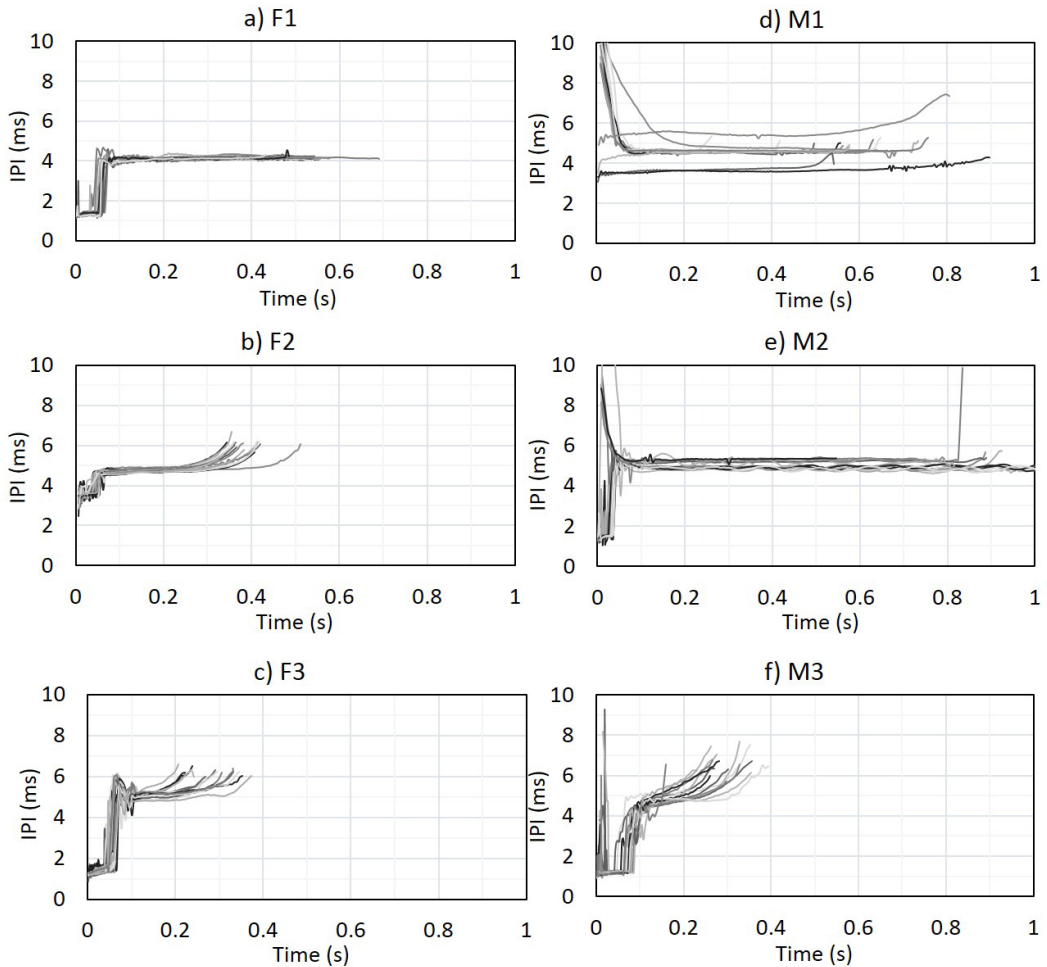
analyses demonstrated that all temporal parameters and two spectral parameters, 10 BW and  $F_1$ , differed significantly among individuals ( $p < 0.01$ ), but  $F_p$  and  $F_a$  were not significantly different ( $p = 0.05$  and  $p = 0.96$ , respectively). Furthermore, all temporal parameters had PICs  $> 2.0$ , especially the PRR and IPI 1 which had PICs  $> 2.5$ , but spectral parameters had PICs of around 1.

The DFA was performed using the five temporal parameters with high PIC values. To avoid multicollinearity,  $N_p$  and PRR with high variance inflation factors were excluded from the dataset. The quadratic DFA based on DUR, IPI 1, and IPI 2 resulted in correct classification rates of 100, 96.2, 81.0, 97.4, 87.5, and 43.8% for F1, F2, F3, M1, M2, and M3, respectively, with an overall correct classification rate of 87.3%. About half of the PS1 calls of M3 were misattributed to F3. The stepwise DFA revealed that the most powerful discriminator was IPI 2, followed by IPI 1.

#### Experiment 2

The number of PS1 calls was compared among the three recording sections: (1) 12 min before the gate opens, (2) 12 min during the gate opening, and (3) 12 min after the gate opens. Only M1





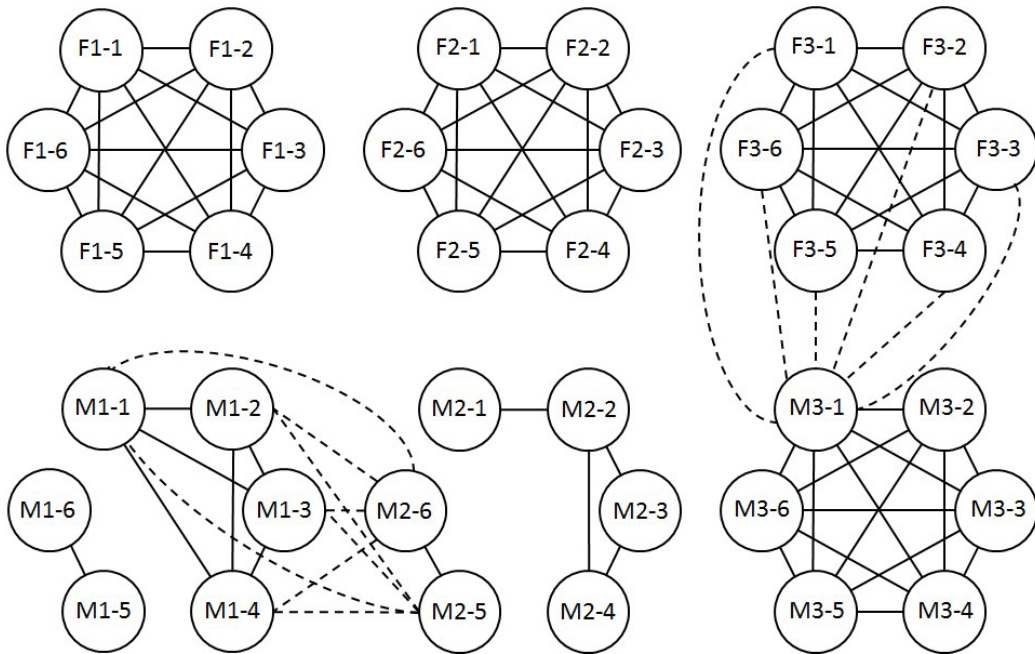
**Figure 4.** IPI contours of PS1 calls from six belugas: (a) F1, (b) F2, (c) F3, (d) M1, (e) M2, and (f) M3. Fifteen examples were randomly selected from each of the belugas.

produced PS1 calls ( $n = 19$ ) for 12 min before the gate opened, and there were no PS1 calls for 12 min after the gate opened; whereas PS1 calls were produced by F1 ( $n = 6$ ), M1 ( $n = 43$ ), and M2 ( $n = 6$ ) during the gate opening. Fifteen out of 43 M1 samples and four out of six M2 samples during the gate opening had a good acoustic quality; their IPI contours are depicted in Figure 7.

Human observers classified a total of 81 samples (38 M1 samples and 24 M2 samples in the free-swimming contexts of Experiment 1; 15 M1 samples and four M2 samples in the gate-opening contexts of Experiment 2), and the results showed that the samples during the gate opening were more individually distinctive and consistent than during free swimming (Table 2). However, 24 out of 38 M1 samples during free swimming

were grouped together by the observers with all M1 samples during the gate opening. Similarly, 12 out of 24 M2 samples during free swimming were grouped together with all M2 samples during the gate opening.

The contextual comparison of duration was conducted on the PS1 calls with the individually distinctive, stereotyped IPI contours as the result of classification performed by human observers. The duration of the M1 samples during the gate opening (mean  $\pm$  SD) was  $1.33 \pm 0.17$  s (0.98 to 1.67 s) ( $n = 15$ ), which was significantly longer than the duration of  $0.57 \pm 0.13$  s (0.26 to 0.85 s) ( $n = 24$ ) during free swimming (two-sample  $t$  test:  $t_{37} = -15.02$ ,  $p < 0.01$ ). Also, the duration of M2 samples during the gate opening had a significantly longer duration of  $2.11 \pm 0.41$  s (0.98 to



**Figure 5.** Similarity network depicted based on the human observer classification of the randomly ordered 36 IPI contours (randomly selected six samples from each of six belugas). The numbers following hyphens indicate the serial numbers of samples from each beluga. If at least three out of five observers agreed on a pair of samples to be grouped together, the pair was connected by a line: when the pair was samples from a single beluga, they were connected by a solid line, and when the pair was samples from two different belugas, they were connected by a broken line.

1.67 s) ( $n = 4$ ) than  $1.02 \pm 0.11$  s (0.83 to 1.19 s) ( $n = 12$ ) during free swimming (Welch two-sample  $t$  test:  $t_{3,1} = -4.61, p < 0.01$ ).

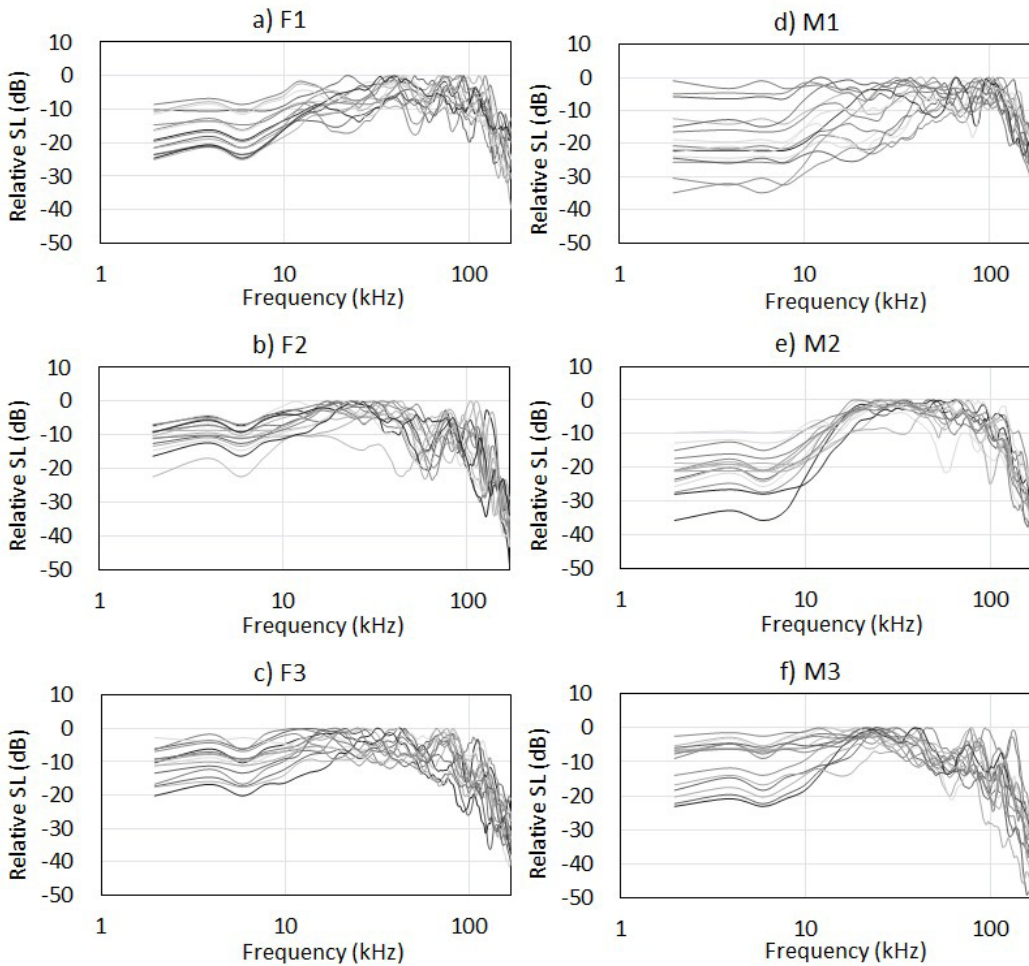
### Discussion

#### “Creaking Calls” – The New Name for Contact Calls in Belugas

All the belugas, excluding a calf, in the present study produced PS1 calls. Given that the PS1 calls were used for vocal exchange and contained individual information, they were likely to have the same function as the earlier described PS1 calls. When taking the results of the present and previous PS1 studies together (Morisaka et al., 2013; Mishima et al., 2015) and referring to the call types in other principal studies on beluga contact calls (Van Parijs et al., 2003; Vergara & Barrett-Lennard, 2008; Vergara et al., 2010; Panova et al., 2017), the contact calls of belugas are generalized and newly named *creaking calls*. The creaking call is a fixed pulse train that sounds like a ratchet or a door creaking. Broadband energy is distributed from less than 1 up to at least 170 kHz, and the duration is approximately 0.15 to 2.50 s.

Several creaking calls contain tone-like components co-occurring with the main pulse train. The creaking call is likely obtained through developmental processes and/or vocal learning. Belugas use creaking calls for vocal exchanges regulated by an approximately 1-s response rule such that another animal responds to the production of a creaking call within 1 s. The IPI contour of the main pulse train in a creaking call, especially in its initial part (the first 0.1 s), is individually specific, although adult males seem to possess not only the individualized creaking call but also non-individualized creaking calls. The duration of creaking calls changes depending on context. The definition was based on the following findings.

For the vocal exchange in the present study, most of the observed WIIs occurred after 1.0 s, with a first peak at 1.6 to 1.8 s, but more than half of the observed BIIs occurred within 1.0 s, with a sharp peak at 0.4 to 0.6 s (Figure 2). The difference between the observed WIIs and BIIs indicated that intervals of adjacent PS1 calls produced by different individuals were shorter than those produced by the same whale. The frequency distribution of the observed BIIs was significantly different from that of the simulated BIIs, and it



**Figure 6.** Power spectra calculated at the middle pulse location for PS1 calls from six belugas: (a) F1, (b) F2, (c) F3, (d) M1, (e) M2, and (f) M3 (FFT size: 256 points; window: Hamming; smoothing: five points). Fifteen examples were randomly selected from each beluga.

indicates that the observed BII distribution was not an incidental result.

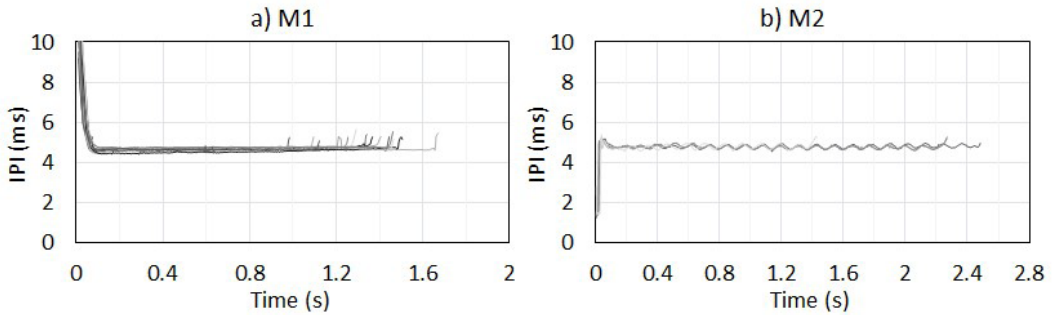
The previous PS1 study found that the frequency distribution of the WIIs had a sharp peak at 1.0 s, and the frequency distribution of the BIIs had a gradual peak at -0.5 s (Morisaka et al., 2013). The interval criterion in the previous study was the latency period from the end of a PS1 to the beginning of the next PS1 and was a shorter duration than the criterion in the present study. When the distributions of the previous study were shifted to the right along the temporal axis by 0.85 s—the mean PS1 duration of Morisaka et al. (2013)—their distribution resembled that of the present study. We thus concluded that the production of PS1 calls was temporally regulated:

responders called back within approximately 1 s, they called frequently at 0.4 to 0.6 s, and callers kept silent in the 1-s time window.

For Type A calls, a 2-s cutoff was selected for the BII distribution, but the WII distribution was not investigated (Vergara et al., 2010). The Type A call study divided the BIIs shorter than 2 s into the time difference between “overlapping” calls and “adjacent” calls. If the second call was produced by different individuals before the termination of the initial call, they were termed *overlapping calls*. If the second call was produced by different individuals within 2 s but after the termination of the initial call, they were termed *adjacent calls*. The interval distributions of overlapping and adjacent calls have a peak at 1 and 2 s, respectively.

**Table 1.** Characteristics and results of the statistical analyses for each acoustic parameter of the PS1 calls.  $N_p$  = number of pulses, DUR = duration, PRR = pulse repetition rate, IPI 1 = average inter-pulse interval of pulse numbers 11 to 20, IPI 2 = average inter-pulse interval of pulse numbers 11 to 20 from the final pulse of the call,  $F_p$  = peak frequency, 10 BW = 10 dB bandwidth,  $F_l$  = the lower frequency of the 10 BW, and  $F_u$  = the upper frequency of the 10 BW.

Parameter	$N_p$	DUR (s)	PRR (pulses/s)	IPI 1 (ms)	IPI 2 (ms)	$F_p$ (kHz)	10 BW (kHz)	$F_l$ (kHz)	$F_u$ (kHz)
Mean $\pm$ SD	115 $\pm$ 43	0.47 $\pm$ 0.19	250.6 $\pm$ 55.4	3.17 $\pm$ 1.62	4.84 $\pm$ 0.42	48.3 $\pm$ 27.8	60.8 $\pm$ 27.7	19.5 $\pm$ 21.8	80.2 $\pm$ 31.1
Max.	260	1.19	525.5	6.12	6.19	123.0	136.7	93.8	146.5
Min.	53	0.16	178.8	1.04	3.48	2.0	15.6	2.0	25.4
Kruskal-Wallis test	$H_5 = 196.23$ $p < 0.01$	$H_5 = 207.58$ $p < 0.01$	$H_5 = 189.81$ $p < 0.01$	$H_5 = 198.36$ $p < 0.01$	$H_5 = 196.14$ $p < 0.01$	--	$H_5 = 44.60$ $p < 0.01$	$H_5 = 57.53$ $p < 0.01$	--
One-way ANOVA	--	--	--	--	--	$F_{5,266} = 3.98$ $p = 0.05$	--	--	$F_{5,266} = 0.002$ $p = 0.96$
PIC	2.07	2.33	2.71	2.57	2.27	1.09	1.07	1.10	1.17



**Figure 7.** IPI contours of PS1 calls with good acoustic quality during the gate opening in Experiment 2: (a) M1 and (b) M2. Fifteen and four samples are represented for M1 and M2, respectively.

**Table 2.** Percentage of the pairs of IPI contours grouped together by at least three observers to the total pairs of samples in each combination of two from the four datasets: (1) 38 M1 samples during the free swimming, (2) 15 M1 samples during the gate opening, (3) 24 M2 samples during the free swimming, and (4) four M2 samples during the gate opening. The numbers in parentheses indicate the number of pairs grouped together/the total number of pairs.

ID & context	M1 Free swimming	M1 Gate opening	M2 Free swimming	M2 Gate opening
M1 Free swimming	56.8% (399/703)			
M1 Gate opening	66.7% (380/570)	100.0% (105/105)		
M2 Free swimming	24.7% (225/912)	23.1% (83/360)	34.8% (96/276)	
M2 Gate opening	0.0% (0/152)	0.0% (0/60)	50.0% (48/96)	100.0% (6/6)



When the overlapping and adjacent distributions are added together to compare the interval distributions of PS1 studies, the constructed distribution has a sharp peak at 1 s. The reason why BII distributions between PS1 and Type A calls was slightly different is unclear, but there is a possibility that this difference may be because the subjects of the PS1 and Type A call studies originated from Russian and Canadian populations, respectively. Further investigations on population differences in between call intervals should be conducted. From the exchange characteristics of PS1 and Type A calls, we defined that creaking calls were used for vocal exchanges regulated by an approximately 1-s response rule.

Each beluga, excluding adult males, in the present study produced only PS1 calls with a stereotyped IPI contour that was individually specific (Figure 4). This was supported by the results of the classification by human observers (Figure 5), univariate statistical analyses, PIC (Table 1), and quadratic DFA. The classification by observers showed that the IPI contours from three females were similar within individuals and different among individuals. The juvenile male, M3, had a similar pattern to the females, and one sample (M3-1) was also classified with F3 samples. M1 and M2 had not only individually distinctive types of IPI contours but also shared types. In addition, all temporal parameters had significant inter-individual differences and high PIC scores; in particular, PRR and IPI 1 had PIC scores greater than 2.5, indicating that PRR and IPI 1 had higher inter-individual variations than intra-individual variations. Furthermore, a quadratic DFA based on DUR, IPI 1, and IPI 2 resulted in an overall correct classification rate of 87.3%, and the most informative parameters in the DFA were IPI 2 followed by IPI 1. These results suggest that the pulse repetition pattern has a high potential as an individual identification medium. This finding coincides with the results of previous PS1 studies (Morisaka et al., 2013; Mishima et al., 2015) and is supported by the findings in Panova et al. (2017). Each of the Type A variants, which were classified based on the pulse repetition rate and energy distribution, did not belong to an individual (Vergara et al., 2010). However, the possibility that Type A calls could exhibit identity coding if they are further classified based on the IPI contours remains unexplored.

The initial part of the IPI contours seems to contain more information for individual identification than the final part because an abrupt change was found in the initial part of the IPI contours in both the present and previous studies (Morisaka et al., 2013; Mishima et al., 2015; Figure 4). Furthermore, overlapping exchanges often occurred in the present

study as was reported by the previous PS1 study (Morisaka et al., 2013) and the Type A call study (Vergara et al., 2010). Thus, belugas seem to answer immediately after hearing the initial part of the call; therefore, the initial part may be more important for individual recognition. From these findings, creaking calls were characterized such that the IPI contour of the main pulse train, especially in its initial part (the first 0.1 s), is individually specific.

Statistical analyses indicated that the durations of the PS1 calls had an individual distinctiveness in the normal free-swimming contexts in the present study, but the duration within individuals seemed to vary depending on the context. The durations of the individualized PS1 calls of adult males were longer in the gate-opening context of Experiment 2 than in the normal free-swimming contexts of Experiment 1. The durations of the PS1 calls in the previous study were not individually specific in the separation contexts (Mishima et al., 2015). In addition, the duration was less important for the classification of individually specific stereotyped pulsed or mixed call types in Panova et al. (2017). These results imply that the duration may not be used for individual recognition; rather, it could contain information about the callers' motivational state. Similarly, the duration of the signature whistles in bottlenose dolphins is affected by their motivational state (Esch et al., 2009). When the values of the durations in the PS1 calls of Experiments 1 and 2 and the earlier described PS1 calls were integrated, the durations ranged from approximately 0.15 to 2.50 s. The pulse trains in the Van Parijs et al. (2003) study had average durations of 0.3 to 1.9 s; the Type A calls in the Vergara et al. (2010) study had average durations of 1.2 to 1.9 s; and the durations of the stereotyped pulsed or mixed calls in the work by Panova et al. (2017) were between 0.17 to 1.34 s. The previous studies recorded calls in various contexts, but their durations were within the range of the PS1 studies. Thus, we defined that the call duration of creaking calls varies depending on the context, but it is approximately in the 0.15 to 2.50 s range.

Individual specificity could not be found in the spectra of PS1 calls (Figure 6). In addition, PIC scores of all spectral parameters were around 1, indicating that the intra-individual and inter-individual variations were comparable, although the univariate statistical analyses showed that some spectral parameters were individually distinctive (Table 1). These findings were similar to the results of the previous PS1 study in which a visual comparison of the spectra showed that there were no intra-individual consistency and inter-individual differences in the PS1 calls even though statistical analyses indicated that some

spectral parameters were individually distinctive (Mishima et al., 2015). Therefore, we can conclude that spectral cues have less individuality.

As described above, there is some consistency and similarity between the PS1 calls described in the present and previous studies; however, PS1 calls in the present study have some acoustic characteristics to which the previous PS1 definition do not apply. Various types of IPI contours were observed in the present study, and most of them were different from the typical decrease-constant-increase contours observed in the previous studies (Morisaka et al., 2013; Mishima et al., 2015; Figure 4). Thus, it can be concluded that there is no typical form of the IPI contours in creaking calls, and there is the potential for a high degree of individuality in IPI contours.

In addition, in the present study, all the PS1 calls of all the belugas had a tone-like (tonal or secondary pulsed) component that co-occurred with the main broadband pulse train and was consistent intra-individually (Figure 3); whereas in the previous study, a tonal component was found in the PS1 calls of only one adult male (Mishima et al., 2015). Similar tone-like components were found in some of the pulsed calls in the Van Parijs et al. (2003) study and the Type A calls in the Vergara & Barrett-Lennard (2008) and Vergara et al. (2010) studies, and stereotyped pulsed or mixed calls in the Panova et al. (2017) study. In the Panova et al. study, not only the pulse repetition patterns of the main pulsed components but also the frequency contour of the tone-like components were used for classification of the stereotyped pulsed or mixed calls, and each type of pulse repetition pattern and tone-like component may belong to one individual. These results suggest that several belugas have a tone-like component in their creaking calls, which may function as a supportive individual identity. Including or not including the component may also be one of the forms of individuality.

The 2-mo-old female calf, F4, in the present study did not produce PS1 calls. The 1-y-old male calf in Mishima et al. (2015) produced PS1 calls, but his IPI contour fluctuated over the duration of the call. Further, the male calf in the Vergara & Barrett-Lennard (2008) study gradually developed the pulsed and tonal components of Type A calls in his first year. These findings imply that beluga calves gradually obtain their individualized creaking calls in their first few years through developmental processes and/or vocal learning.

Finally, it should be noted that when acoustic parameters of the PS1 calls were analysed in the present study, a large number of samples were unable to be analysed. There were various factors causing poor acoustic quality such as saturation and low amplitude, but the primary factors were

multiple reflections and the overlapping of two consecutive PS1 calls by different individuals. However, the impact of the unused samples on the results seemed to be low because there were aural and spectrographic similarities between the analysed and unused PS1 calls.

### *Sexual Difference of Creaking Calls*

From here, we use the term *creaking calls* instead of PS1 calls. Females and the single juvenile male produced only individually specific creaking calls, but the two adult males produced several types of creaking calls. The creaking call rate increased during the gate opening and decreased after the gate opened. Further, the classification by human observers suggested that both adult males used not only the individually distinctive, stereotyped IPI contours but also a shared type of IPI contour and other non-individualized contours during free swimming in Experiment 1. In contrast, they produced only the individually distinctive, stereotyped creaking calls during the gate opening, which also seemed to be the most frequently used type during free swimming (Table 2). Mishima et al. (2015) showed that an adult male mostly produced one type of creaking call with an individually distinctive IPI contour in the separation context, although two creaking calls with clearly different IPI contours were found in his call repertoire (Mishima et al., 2015). Thus, in the vocal repertoire of males, the predominantly produced creaking call type in separation, reunion, or greeting contexts where they may have needed to advertise their identity appears to be their individually specific creaking call.

If the individually specific creaking call is used predominantly, it is unclear why adult males also produced non-individualized creaking calls. For comparison, bottlenose dolphins produce individually specific signature whistles, but they also produce several non-signature whistles (Tyack, 1986; Sayigh et al., 1990; Janik et al., 1994; McCowan & Reiss, 1995; Janik & Slater, 1998; Watwood et al., 2004, 2005; Nakahara & Miyazaki, 2011). Signature whistles were most frequently produced in isolation, reunion, or greeting contexts to identify each other (Caldwell et al., 1990; Sayigh et al., 1990; Janik et al., 1994; Janik & Slater, 1998; Watwood et al., 2005; Quick & Janik, 2012), whereas non-signature whistles were often produced in other contexts (Janik et al., 1994; Janik & Slater, 1998; Watwood et al., 2005; King & Janik, 2015). Bottlenose dolphin males disperse from their matrilineal group after weaning and make alliances with other males (Connor & Krützen, 2015). They have a broader whistle repertoire than females (Tyack, 1986; Sayigh et al., 1990; Watwood et al., 2004), and

the whistle repertoires of males expand at about the time they separate from their mother (Sayigh et al., 1990). Allied males in a consortship with a female mostly produced non-signature whistles (Watwood et al., 2005). Whistle convergence has also been found among allied males (Smolker & Pepper, 1999; Watwood et al., 2004), and these whistles are likely to play a role in the alliance signature to form alliance bonds (Smolker & Pepper, 1999). Adult male belugas also appeared to facilitate long-term social bonds with other mature males (Michaud, 2005; Colbeck et al., 2012). The creaking call repertoires of male belugas might expand as they mature as seen in bottlenose dolphins. The juvenile male in the present study was housed with the females rather than the males, and he had just one type of creaking call. The wild juvenile males leave their natal groups and may spend time with other adult males, which could be where they learn several types of creaking calls. Although we could not clarify the function of the non-individualized creaking calls of males, the call sharing found in the males of the present study (Figure 5; Table 2) and the Panova et al. (2017) study might occur to strengthen the male bond.

#### Future Plans

The present study discussed the definition and sex differences in contact calls of belugas. However, our conclusions were primarily based on studies in captivity, which could affect the data. Given that pulsed calls similar to creaking calls were produced from wild belugas in various behavioural activities (Sjare & Smith, 1986a; Karlsen et al., 2002; Panova et al., 2012; Alekseeva et al., 2013), the limited behavioural context in captivity might lead to more stereotyped IPI contours. Further studies of wild belugas in various behavioural contexts should be conducted to gain insight into the acoustical and functional characteristics of creaking calls. In addition, as a next step, an optimum classification method of the IPI contours should be established to divide creaking calls into types for application to research on individual identification of wild belugas and for the examination of creaking call repertoires in adult males. From the results, the classification performed by human observers appeared to be useful, similar to the analysis done for signature whistles of bottlenose dolphins (Janik, 1999); however, it is time-consuming and laborious for a large dataset. The automatic classification methods of frequency or pulse-rate contours used for the contact calls of bottlenose dolphins and killer whales, respectively, might be applicable to classify the IPI contours of creaking calls in belugas (Deecke & Janik, 2006).

#### Acknowledgments

We are grateful to the members of the Shimane Aquarium for their support in this research, especially Y. Mori, K. Adachi, and I. Ohtsuji. We are also deeply indebted to T. Sasakura and Y. Abe of Fusion Inc. for their technical support. Thanks to S. Sendo, A. Enokizu, A. Furuyama, I. Kanda, and K. Gondo of Mie University and C. Namaizawa, S. Asai, T. Kasai, Y. Shimosaki, and A. Nakamura of Tokyo University of Marine Science and Technology who helped us with the IPI contour classification. We also thank Dr. E. Henderson and two anonymous reviewers for their valuable and helpful comments on our manuscript. This research was financially supported by JSPS KAKENHI Grant Numbers 2510729 and 18J40100 for Y. Mishima; 23220006, 15H05709, and 18H05071 for T. Morisaka; and 24657015 for Y. Miyamoto.

#### Literature Cited

- Alekseeva, Y. I., Panova, E. M., & Bel'kovich, V. M. (2013). Behavioral and acoustical characteristics of the reproductive gathering of beluga whales (*Delphinapterus leucas*) in the vicinity of Myagostrov, Golyi Sosnovets, and Roganka Islands (Onega Bay, the White Sea). *Biology Bulletin*, 40(3), 307-317. <https://doi.org/10.1134/S1062359013030023>
- Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J., & Rendell, L. (2011). Individually distinctive acoustic features in sperm whale codas. *Animal Behaviour*, 81(4), 723-730. <https://doi.org/10.1016/j.anbehav.2010.12.019>
- Bel'kovich, V. M., & Sh'ekotov, M. N. (1993). *The belukha whale: Natural behaviour and bioacoustics*. Woods Hole, MA: Woods Hole Oceanographic Institution Press. <https://doi.org/10.1575/1912/75>
- Belikov, R. A., & Bel'kovich, V. M. (2008). Communicative pulsed signals of beluga whales in the reproductive gathering off Solovetskii Island in the White Sea. *Acoustical Physics*, 54(1), 115-123. <https://doi.org/10.1134/S1063771008010168>
- Caldwell, M. C., & Caldwell, D. K. (1965). Individualized whistle contours in bottlenose dolphins (*Tursiops truncatus*). *Nature*, 207, 434-435. <https://doi.org/10.1038/207434a0>
- Caldwell, M. C., Caldwell, D. K., & Tyack, P. L. (1990). Review of the signature-whistle hypothesis for the Atlantic bottlenose dolphin. In S. Leatherwood & R. R. Reeves (Eds.), *The bottlenose dolphin* (pp. 199-234). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-0-12-440280-5.50014-7>
- Charrier, I., Mathevon, N., & Jouventin, P. (2002). How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. *Journal of Experimental Biology*, 205(5), 603-612.
- Chmelnitsky, E. G., & Ferguson, S. H. (2012). Beluga whale, *Delphinapterus leucas*, vocalizations from the

- Churchill River, Manitoba, Canada. *The Journal of the Acoustical Society of America*, 131(6), 4821-4835. <https://doi.org/10.1121/1.4707501>
- Colbeck, G. J., Duchesne, P., Postma, L. D., Lesage, V., Hammill, M. O., & Turgeon, J. (2012). Groups of related belugas (*Delphinapterus leucas*) travel together during their seasonal migrations in and around Hudson Bay. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1752), 20122552. <https://doi.org/10.1098/rspb.2012.2552>
- Connor, R. C., & Krützen, M. (2015). Male dolphin alliances in Shark Bay: Changing perspectives in a 30-year study. *Animal Behaviour*, 103, 223-235. <https://doi.org/10.1016/j.anbehav.2015.02.019>
- Deecke, V. B., & Janik, V. M. (2006). Automated categorization of bioacoustic signals: Avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America*, 119(1), 645-653. <https://doi.org/10.1121/1.2139067>
- Esch, H. C., Sayigh, L. S., Blum, J. E., & Wells, R. S. (2009). Whistles as potential indicators of stress in bottlenose dolphins (*Tursiops truncatus*). *Journal of Mammalogy*, 90(3), 638-650. Retrieved from [www.jstor.org/stable/30224517](http://www.jstor.org/stable/30224517); <https://doi.org/10.1644/08-MAMM-A-069R.1>
- Ford, J. K. B. (1989). Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. *Canadian Journal of Zoology*, 67(3), 727-745. <https://doi.org/10.1139/z89-105>
- Ford, J. K. B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Canadian Journal of Zoology*, 69(6), 1454-1483. <https://doi.org/10.1139/z91-206>
- Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, 3(1), 150372. <https://doi.org/10.1098/rsos.150372>
- Janik, V. M. (1999). Pitfalls in the categorization of behaviour: A comparison of dolphin whistle classification methods. *Animal Behaviour*, 57(1), 133-143. <https://doi.org/10.1006/anbe.1998.0923>
- Janik, V. M., & Slater, P. J. B. (1998). Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Animal Behaviour*, 56(4), 829-838. <https://doi.org/10.1006/anbe.1998.0881>
- Janik, V. M., Dehnhardt, G., & Todt, D. (1994). Signature whistle variations in a bottlenosed dolphin, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, 35(4), 243-248. Retrieved from [www.jstor.org/stable/4601006](http://www.jstor.org/stable/4601006); <https://doi.org/10.1007/BF00170704>
- Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences, U.S.A.*, 103(21), 8293-8297. <https://doi.org/10.1073/pnas.0509918103>
- Karlsen, J. D., Bisther, A., Lydersen, C., Haug, T., & Kovacs, K. M. (2002). Summer vocalizations of adult male white whales (*Delphinapterus leucas*) in Svalbard, Norway. *Polar Biology*, 25(11), 808-817. <https://doi.org/10.1007/s00300-002-0415-6>
- King, S. L., & Janik, V. M. (2015). Come dine with me: Food-associated social signalling in wild bottlenose dolphins (*Tursiops truncatus*). *Animal Cognition*, 18(4), 969-974. <https://doi.org/10.1007/s10071-015-0851-7>
- Kondo, N., & Watanabe, S. (2009). Contact calls: Information and social function. *Japanese Psychological Research*, 51(3), 197-208. <https://doi.org/10.1111/j.1468-5884.2009.00399.x>
- Kondo, N., Watanabe, S., & Izawa, E. (2010). A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithological Science*, 9(1), 83-91. <https://doi.org/10.2326/osj.9.83>
- McCowan, B., & Reiss, D. (1995). Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): A re-evaluation of the signature whistle hypothesis. *Ethology*, 100(3), 194-209. <https://doi.org/10.1111/j.1439-0310.1995.tb00325.x>
- McGarigal, K., Cushman, S., & Stafford, S. (2000). *Multivariate statistics for wildlife and ecology research*. New York: Springer Press. <https://doi.org/10.1007/978-1-4612-1288-1>
- Michaud, R. (2005). Sociality and ecology of the odontocetes. In K. E. Ruckstuhl & P. Neuhaus (Eds.), *Sexual segregation in vertebrates: Ecology of the two sexes* (pp. 303-326). Cambridge, UK: Cambridge University Press.
- Mishima, Y., Morisaka, T., Itoh, M., Matsuo, I., Sakaguchi, A., & Miyamoto, Y. (2015). Individuality embedded in the isolation calls of captive beluga whales (*Delphinapterus leucas*). *Zoological Letters*, 1, 27. <https://doi.org/10.1186/s40851-015-0028-x>
- Morisaka, T., Yoshida, Y., Akune, Y., Mishima, H., & Nishimoto, S. (2013). Exchange of “signature” calls in captive belugas (*Delphinapterus leucas*). *Journal of Ethology*, 31(2), 141-149. <https://doi.org/10.1007/s10164-013-0358-0>
- Nakahara, F., & Miyazaki, N. (2011). Vocal exchanges of signature whistles in bottlenose dolphins (*Tursiops truncatus*). *Journal of Ethology*, 29(2), 309-320. <https://doi.org/10.1007/s10164-010-0259-4>
- Nousek, A. E., Slater, P. J. B., Wang, C., & Miller, P. J. O. (2006). The influence of social affiliation on individual vocal signatures of northern resident killer whales (*Orcinus orca*). *Biology Letters*, 2(4), 481-484. <https://doi.org/10.1098/rsbl.2006.0517>
- Oliveira, C., Wahlberg, M., Silva, M. A., Johnson, M., Antunes, R., Wisniewska, D. M., . . . Madsen, P. T. (2016). Sperm whale codas may encode individuality as well as clan identity. *The Journal of the Acoustical Society of America*, 139(5), 2860-2869. <https://doi.org/10.1121/1.4949478>
- Panova, E., Agafonov, A., Belikov, R., & Melnikova, F. (2017). Vocalizations of captive beluga whales, *Delphinapterus leucas*: Additional evidence for contact signature “mixed” calls in belugas. *Marine Mammal Science*, 33(3), 889-903. <https://doi.org/10.1111/mms.12393>



- Panova, E. M., Belikov, R. A., Agafonov, A. V., & Bel'kovich, V. M. (2012). The relationship between the behavioral activity and the underwater vocalization of the beluga whale (*Delphinapterus leucas*). *Oceanology*, *52*(1), 79-87. <https://doi.org/10.1134/S000143701201016X>
- Quick, N. J., & Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society of London B: Biological Sciences*, *279*(1738), 2539-2545. <https://doi.org/10.1098/rspb.2011.2537>
- Recchia, C. A. (1994). *Social behaviour of captive belugas*, *Delphinapterus leucas* (Doctoral dissertation). Woods Hole Oceanographic Institution/Massachusetts Institute of Technology. Retrieved from [www.dtic.mil/dtic/tr/fulltext/u2/a289549.pdf](http://www.dtic.mil/dtic/tr/fulltext/u2/a289549.pdf)
- Rendell, L., & Whitehead, H. (2003). Vocal clans in sperm whales (*Physeter macrocephalus*). *Proceedings of the Royal Society of London B: Biological Sciences*, *270*(1512), 225-231. <https://doi.org/10.1098/rspb.2002.2239>
- Rendell, L., & Whitehead, H. (2004). Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurement. *Animal Behaviour*, *67*(5), 865-874. <https://doi.org/10.1016/j.anbehav.2003.04.018>
- Sayigh, L. S., Wells, R. S., & Janik, V. M. (2017). What's in a voice? Dolphins do not use voice cues for individual recognition. *Animal Cognition*, *20*, 1067-1079. <https://doi.org/10.1007/s10071-017-1123-5>
- Sayigh, L. S., Esch, H. C., Wells, R. S., & Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour*, *74*(6), 1631-1642. <https://doi.org/10.1016/j.anbehav.2007.02.018>
- Sayigh, L. S., Tyack, P. L., Wells, R. S., & Scott, M. D. (1990). Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: Stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, *26*(4), 247-260. <https://doi.org/10.1007/BF00178318>
- Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: Insights into communication function. *Animal Behaviour*, *76*(6), 1977-1988. <https://doi.org/10.1016/j.anbehav.2008.07.032>
- Sjare, B. L., & Smith, T. G. (1986a). The relationship between behavioral activity and underwater vocalizations of the white whale, *Delphinapterus leucas*. *Canadian Journal of Zoology*, *64*(12), 2824-2831. <https://doi.org/10.1139/z86-406>
- Sjare, B. L., & Smith, T. G. (1986b). The vocal repertoire of white whales, *Delphinapterus leucas*, summering in Cunningham Inlet, Northwest Territories. *Canadian Journal of Zoology*, *64*(2), 407-415. <https://doi.org/10.1139/z86-063>
- Smolker, R., & Pepper, J. W. (1999). Whistle convergence among allied male bottlenose dolphins (*Delphinidae*, *Tursiops* sp.). *Ethology*, *105*(7), 595-617. <https://doi.org/10.1046/j.1439-0310.1999.00441.x>
- Sugiura, H. (1993). Temporal and acoustic correlates in vocal exchange of coo calls in Japanese macaques. *Behaviour*, *124*(3), 207-225. <https://doi.org/10.1163/156853993X00588>
- Tyack, P. (1986). Whistle repertoires of two bottlenosed dolphins, *Tursiops truncatus*: Mimicry of signature whistles? *Behavioral Ecology and Sociobiology*, *18*(4), 251-257. Retrieved from [www.jstor.org/stable/4599887](http://www.jstor.org/stable/4599887); <https://doi.org/10.1007/BF00300001>
- Van Parijs, S. M., Lydersen, C., & Kovacs, K. M. (2003). Sounds produced by individual white whales, *Delphinapterus leucas*, from Svalbard during capture (L). *The Journal of the Acoustical Society of America*, *113*(1), 57-60. <https://doi.org/10.1121/1.1528931>
- Vergara, V., & Barrett-Lennard, L. G. (2008). Vocal development in a beluga calf (*Delphinapterus leucas*). *Aquatic Mammals*, *34*(1), 123-143. <https://doi.org/10.1578/AM.34.1.2008.123>
- Vergara, V., Michaud, R., & Barrett-Lennard, L. (2010). What can captive whales tell us about their wild counterparts? Identification, usage, and ontogeny of contact calls in belugas (*Delphinapterus leucas*). *International Journal of Comparative Psychology*, *23*(3), 278-309. Retrieved from <https://escholarship.org/uc/item/4gt03961>
- Watwood, S. L., Tyack, P. L., & Wells, R. S. (2004). Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, *55*(6), 531-543. <https://doi.org/10.1007/s00265-003-0724-y>
- Watwood, S. L., Owen, E. C. G., Tyack, P. L., & Wells, R. S. (2005). Signature whistle use by temporarily restrained and free-swimming bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour*, *69*(6), 1373-1386. <https://doi.org/10.1016/j.anbehav.2004.08.019>
- Yoshida, Y. M., Morisaka, T., Nakahara, F., & Saito, S. (2016). Receiving sensitivity of the hydrophones attached to acrylic plate: Towards simplification of dolphin sound study at aquariums. *Bulletin of Institute of Oceanic Research and Development, Tokai University*, *37*, 1-6.
- Yosida, S., Kobayashi, K., Ikebuchi, M., Ozaki, R., & Okanoya, K. (2007). Antiphonal vocalization of a subterranean rodent, the naked mole-rat (*Heterocephalus glaber*). *Ethology*, *113*(7), 703-710. <https://doi.org/10.1111/j.1439-0310.2007.01371.x>