Individuality embedded in contact calls of beluga whales Delphinapterus leucas

Doctoral Dissertation

INDIVIDUALITY EMBEDDED IN CONTACT CALLS OF BELUGA WHALES *Delphinapterus leucas*

March 2016

Graduate School of Marine Science and Technology

Tokyo University of Marine Science and Technology

Doctoral Course of Applied Marine Environmental Studies

YUKA MISHIMA

Acknowledgements

This dissertation could not have been completed without the support of many people.

Foremost, I am extremely grateful to my advisor Prof. Yoshinori Miyamoto for his advice and support throughout my five years in graduate school. He has provided the extensive technical knowledge necessary for recordings and analyses, and let me move toward to completion of this thesis.

Dr. Tadamichi Morisaka of Tokai University introduced me to the research field and has provided me with a variety of biological expertise. He has supported me over the years, and invaluable discussions with him have helped me determine the route of my study. I would like to express my sincere gratitude to him.

The deepest appreciation goes to the members of the Port of Nagoya Public Aquarium for their cooperation, advice, and encouragement, especially Tomoko Mori, Miho Itoh, Sayo Nishimoto, Mahiro Ryono, Toyoshi Saitou, Yuichiro Akune, Masanori Kurita, Hiroshi Nitto, and Makoto Soichi.

I am also indebted to the members of Shimane Aquarium for their cooperation, advice, and encouragement, especially Yoko Mori, Yuki Mishima, Koji Adachi, Isao Ohtsuji, and Tadashi Sunada.

The members of both aquariums have been welcoming and fantastic to work alongside. Their understanding of my study, scheduling of experiments, support for recordings, and abundant knowledge have been indispensable to my study.

Development of the broadband underwater speaker could not have been possible without the cooperation of Dr. Toyoki Sasakura and Yuzo Abe of Fusion Inc. They have also supported me technically over the years. I wish to show my deep gratitude to them for their continuous help and immense knowledge.

Special thanks to Prof. Ikuo Matsuo of Tohoku Gakuin University who taught me signal

processing and wrote the software used for sound source localization. Further, the evaluation of the broadband transmitting system could not have been conducted without his suggestions and support.

Dr. Kazuo Amakasu of Research Center for Advanced Science and Technology shared with me the fun of acoustics, and I could not have created the broadband underwater speaker without his help. I deeply acknowledge him for his valuable comments and advice.

Many thanks to Dr. Keiichi Uchida of our laboratory. He helped me to perform the field experiments of the broadband transmitting system. I also appreciate his genuine concern and interest in my study.

I would like thank Prof. Toshiharu Kakihara of our laboratory for his kindness and thoughtful opinions. He has always supported me throughout my student life in graduate school.

Prof. Hisayuki Arakawa of another laboratory provided me with much advice and encouragement. I owe him a debt of thanks.

I am not through yet. Students in my lab have helped me in many ways, especially Aiko Sakaguchi, who assisted me with collecting data at the Port of Nagoya Public Aquarium. Further, I had a lot of fun being able to spend time with the lab members. I do hereby appreciate all of them.

Last but not the least, I would like to extend my indebtedness to all my family and friends for their understanding, encouragement, and support throughout my thesis research.

This study was financially supported by JSPS KAKENHI Grant Numbers 2510729.

Table of contents

1. Introduction

1.1 Individuality in contact calls

Individual identity advertisement and individual recognition are important for inter-individual social networks. The degrees of individuality linked to social structure and complex social interactions tend to favor strongly recognizable individual identity. The individual information can be carried by physical characteristics (Parr et al., 2000), odor (Johnston, 2003), and vocalizations such as contact calls. Contact calls are used for vocal exchange to maintain group cohesion and affiliation. Individuality in contact calls is reported in various species groups such as birds (Cortopassi & Bradbury, 2006) and mammals, including elephants (Soltis et al., 2005), bats (Arnold & Wilkinson, 2011), primates (Marler & Hobbet, 1975), and odontocetes (Caldwell & Caldwell, 1965). The present study focuses on contact calls in odontocetes. Systematic studies of contact calls in odontocetes have been conducted on only three species; the sperm whale (*Physeter macrocephalus*), killer whale (*Orcinus orca*), and bottlenose dolphin (*Tursiops truncatus*). Before a description of their contact calls, odontocete vocalizations are introduced.

1.1.1 Primary odontocete vocalizations

Odontocete vocalizations are mainly classified into two types; whistles and pulse trains (Fig. 1.1). Whistles are narrowband tonal calls of relatively long duration and are used for communication. Pulse trains are composed of a series of several very short broadband pulses. Pulse trains are further acoustically divided—primarily by inter-pulse intervals—into two types; burst pulses and clicks (Lammers et al., 2004; Morisaka et al., 2011) (Fig. 1.1). Burst pulses are composed of shorter spaced pulses, and they are used for communication. Burst pulses with extremely short inter-pulse intervals, often called pulsed tones, have a tonal quality to human ears and are spectrographically sideband-structured (Watkins, 1967). Clicks

composed of more widely spaced pulses are generally called echolocation clicks. They are primarily used to detect prey or obstacles, and to obtain information on the surrounding environment. However, some species use clicks for communication as well as echolocation as described in subsection 1.1.2.

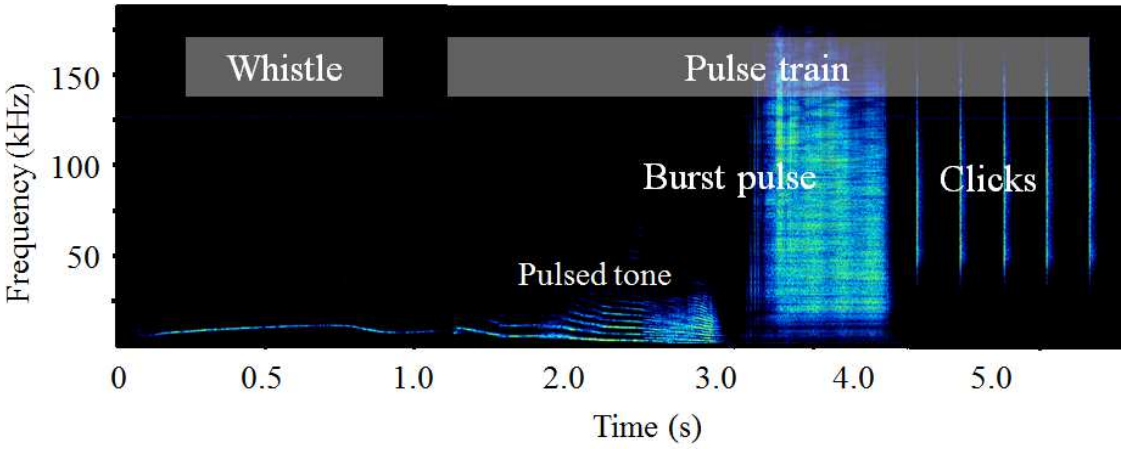


Fig. 1.1 Primary odontocete call types.

1.1.2 Contact calls of sperm whales and killer whales

Sperm whales belong to the family Physeteridae; the earliest diverging group of odontocetes. Matrilineal “units” of sperm whales live in tropical and subtropical waters, and unit composition is stable for a long time. The matrilineal units form temporary “groups” with one or more other units for periods of only hours to days (Whitehead & Weilgart, 2000). Sperm whales exchange stereotyped clicks called “codas” to reinforce social bonds (Whitehead & Weilgart, 1991; Schulz et al., 2008). Codas are composed of 3–40 broadband clicks (0–>20 kHz), last less than 3 s (Watkins & Schevill, 1977), and can be classified into distinct types according to the click repetition patterns and the number of clicks they contain (Weilgart & Whitehead, 1997). The coda types indicate unit identity rather than individual identity, since several coda types are shared among members within matrilineal units (Rendell & Whitehead, 2004; Schulz et al., 2011). In addition, units forming a group have a similar

coda repertoire and some of the coda types are also shared by other groups. The units with a similar coda repertoire are called an acoustic “clan,” and the coda dialects are believed to be the result of cultural transmission based on social learning (Rendell & Whitehead, 2003). Thus, coda may function as a means not only of unit identity, but also of clan identity, allowing units to identify other units of the same clan when they form groups.

It is suggested that individual differences exist in timing around the stereotyped rhythm of a certain coda type (Antunes et al., 2011). Further, codas contain individual differences attributed to individuality in their species-specific sound-producing organ, or spermaceti organ (Gordon, 1991). Each click composing the coda comprises several equally spaced pulses, including a direct pulse generated at the phonic lips, and reflected pulses from air sacs at the anterior and posterior ends of the spermaceti organ. Therefore, the inter-pulse intervals of clicks are related to the size of spermaceti organ, or as observed in several species, individuality in inter-pulse intervals is a by-product of individual differences in the vocal tract (Boughman & Moss, 2003) and is thought to be used for individual recognition within units. Together, these findings reveal that codas are hierarchically coded contact calls with strongly recognizable unit and/or clan identity information and less prominent individual identity information.

Killer whales are the most basal species within Delphinidae. Killer whales are cosmopolitan and are distributed in all oceans worldwide. Long-term, stable matrilineal “units” are the fundamental element of killer whale society, and closely related matrilineal units, which associate regularly with each other, are termed “pods” (Baird, 2000). Resident killer whales exchange stereotyped pulsed tones, or “discrete calls” (Miller et al., 2004) to maintain pod cohesion (Ford, 1989). The pulse repetition rate of discrete calls is modulated over their duration with several abrupt shifts; in addition, some discrete calls contain a simultaneous tonal component (Ford, 1991; Miller & Bain, 2000; Yurk et al., 2002). Discrete calls continue

for 0.25–2.5 s in duration, and harmonics of the tonal component distribute to over 100 kHz (Schevill & Watkins, 1966; Miller & Bain, 2000). The discrete call types with an overlapping tonal component are used as long-distance contact calls, whereas those without an overlapping tonal component are used as close-range contact calls (Miller, 2006; Filatova et al., 2009). The discrete call types indicate pod identity rather than individual identity, because pod members share 7–17 types (Ford, 1989; Ford, 1991). Moreover, some of the call types are also shared by other pods. The acoustically related pods are called an acoustic "clan" (Ford, 1991), and the vocal dialect is explained by cultural transmission based on social learning (Yurk et al., 2002). Thus, discrete calls may serve a function in clan identity as well as pod identity.

A neural network technique showed some degree of individuality within the stereotyped pulse repetition pattern of a certain discrete call type (Nousek et al., 2006). Killer whales thus also have hierarchical information in their contact calls, with strongly recognizable pod and/or clan identity information and less prominent individual identity information.

Sperm whales and killer whales therefore have a similar degree of individuality in pulse-type contact calls, or a slight individuality akin to voice cues. In contrast, bottlenose dolphins—with a complicated social structure—possess strongly recognizable individuality in a whistle-type contact call.

1.1.3 Contact calls of bottlenose dolphins

Bottlenose dolphins are a Delphinidae species that emerged after the divergence of killer whales, and are the most extensively studied odontocete species. They are found in pelagic and coastal areas of temperate and tropical waters worldwide. The society of matrilineal groups is of a highly fission-fusion type, in which individuals associate in small groups that change in composition, often on a daily or hourly basis, whereas mature males form long-

lasting, strong bond groups called alliances (Connor et al., 2000). Bottlenose dolphins need to facilitate strong individual identification to maintain inter-individual relationships in this fluid society. Caldwell and Caldwell found that isolated bottlenose dolphins frequently produced particular whistle types and that their frequency modulation pattern was stereotyped intra-individually and differed inter-individually (Caldwell & Caldwell, 1965; Caldwell & Caldwell, 1968; Caldwell et al., 1990). Those whistles are called “signature whistles,” and are defined as the dominant whistle type in an isolation context (Caldwell et al., 1990; Sayigh et al., 1990).

Signature whistles are characterized by low-frequency tonal sounds from 1 kHz to 30 kHz, and 0.1–4.0 s in duration (Janik & Sayigh, 2013). Playback studies revealed that dolphins could recognize signature whistles of familiar associates (Caldwell et al., 1972; Sayigh et al., 1999) and that the frequency modulation pattern of signature whistles conveys individual information independent of voice features (Janik et al., 2006). Thus, frequency contours themselves function as referential signal of individuals, akin to a “name” in humans.

During the first year of life, bottlenose dolphins develop signature whistles via production learning (Sayigh et al., 1990; Fripp et al., 2005). Production learning is defined as when “signals themselves are modified in form as a result of experience with those of other individuals” (Janik & Slater, 2000), and this learning has been observed in only selected animals; songbirds, hummingbirds, parrots, elephants, bats, pinnipeds, and cetaceans (Janik & Slater, 1997; Poole et al., 2005). Bottlenose dolphin calves model their signature whistles on those of community members, but by modifying some features of the models, they create their own signature whistles that are novel types for the community (Fripp et al., 2005). Signature whistles are consistent for a long time (Sayigh et al., 1990; Watwood et al. 2005), and dolphins memorize conspecifics’ signature whistles for at least 20 years (Bruck, 2013).

Further, bottlenose dolphins evolve elaborate individual recognition systems. Dolphins often exchange their own signature whistles, but matching occurs at a lower rate, where a

responder called back with a copy of the preceding signature whistles (Janik, 2000b; Nakahara & Miyazaki, 2011; King et al., 2013). This copying was often found between close associates such as mother–calf pairs and male alliances (King et al., 2013). King & Janik (2013) provide the possibility that such copying was used to address the signature owner as we address family or friends using their names.

As described above, the form of individuality in contact calls is different among species. To elucidate the evolutionary pathway and adaptive significance of various forms of individuality, we need to increase our understanding of contact calls in other odontocete species. Recently, the existence of individually distinctive contact calls was suggested in beluga whales (*Dephinapterus leucas*) belonging to Monodontidae, and this study focused on contact calls in belugas.

1.2 Beluga whales

Belugas are a circumpolar and annual-migratory species (Richard et al., 2001; Michaud, 2005; Colbeck et al., 2013; Svetochev & Svetocheva, 2013; Hauser et al., 2014). They migrate long distances from overwintering areas of high-latitude polynyas to summering areas in coastal and adjacent offshore waters. Mating is thought to occur from late winter to early summer and calving seems to occur in the summering grounds (Bel'kovitch & Sh'ekotov, 1993; Michaud, 2005; Meschersky et al., 2013). Details of beluga society are still elusive, especially group composition in winter. According to Michaud (2005), their group composition in a summering area in St. Lawrence Estuary, Canada, was primarily classified into two types, fission-fusion matrilineal groups including adult females, calves, and juveniles; and long-lasting smaller groups of adult males. Juvenile groups and huge mixed groups of up to one thousand individuals were occasionally observed. Colbeck et al. (2013) reported that matrilineal groups of Hudson Bay, Canada, maintain relationships with their

group members during migration. However, they tend to associate with non-relatives during summer, excluding mothers and calves who remain together. Moreover, several groups of related belugas may intermingle in the same general summering area, as is evidenced by the lack of associations other than mother–calf pairs when sampled in the summering areas. This would require that matrilineal group members regain contact before departing for migration.

Beluga groups in the White Sea and the Amur Estuary, Russia, were also divided into two types; females with calves and juveniles, and groups of adult males, during summer (Bel'kovitch & Sh'ekotov, 1993). Small groups containing 2–8 individuals were often observed in these areas and they stay together for a long time, but they temporarily form large groups of 15–20 individuals. During migration, the belugas unite into larger formations of up to several hundred of individuals. The high mobility and long-term associations in a fluid social structure of both Canadian and Russian belugas suggest that this species may possess strongly recognizable individuality in contact calls.

Belugas are often called "sea canaries" and make a variety of calls using tonal and pulsed components (Sjare & Smith, 1986a; Bel'kovitch & Sh'ekotov, 1993; Recchia, 1994; Karlsen et al., 2002; van Parijs et al., 2003; Belikov & Bel'kovich, 2006; Belikov & Bel'kovich, 2007; Belikov & Bel'kovich, 2008; Vergara et al., 2010; Panova et al., 2012; Chmelnitsky & Ferguson, 2012; Alekseeva et al., 2013). A few studies into contact calls suggest that belugas use pulse trains for contact. van Parijs et al. (2003) collected vocalizations from temporarily captured belugas in Svalbard, Norway. A mother–calf pair produced pulse trains, and the mother often moved her head toward the calf during the production. A subadult female also produced only pulse trains. These findings suggest that belugas have pulse-type contact calls especially between mothers and dependent young. Vergara & Barrett-Lennard (2008) and Vergara et al. (2010) recorded vocalizations from captive belugas at Vancouver Aquarium, Canada. They reported that one type of pulse train, "Type A" calls, served as contact calls

especially between mothers and calves. The Type A call was the most frequently produced call type in isolation contexts. Moreover, a mother predominantly produced Type A calls the day after the birth of two calves and the death of a calf on a different occasion, as well as whenever she needed to regain or maintain contact with her calf. In addition, there were vocal exchanges of Type A calls between the mother–calf pair. Vergara et al. (2010) additionally made recordings from temporarily restrained belugas in the Nelson River Estuary and social groups in the St. Lawrence Estuary, Canada. The temporarily restrained belugas in the Nelson River Estuary produced only Type A calls. Further, in the St. Lawrence Estuary recordings, a mother swimming around her dead calf and pushing it along also produced a series of Type A calls. The Type A calls were divided into five variants based on difference in pulse repetition rate and energy distribution (Vergara et al., 2010). Although the five Type A variants did not carry individual identity, the possibility that each variant could exhibit identity coding on the basis of some parameters, even if each particular variant per se does not have individual distinctiveness, remains unexplored. Morisaka et al. (2013) suggested the existence of individual identity in contact calls of belugas. The belugas kept at the Port of Nagoya Public Aquarium use one type of pulsed sound, “PS1” calls, for vocal exchange; the pulse repetition pattern of PS1 was different among three adults. Thus, it was speculated that PS1 might have a functional role as contact calls and contain high degrees of individuality as do the signature whistles of bottlenose dolphins. However, the study analyzed only a small amount of data. Therefore, further investigation of PS1 should be performed to understand the form of individuality in the contact calls of belugas.

1.3 Objective and overview of thesis research

Investigation of contact calls in a variety of social odontocete species is important to uncover the evolutionary process and adaptive significance of various forms of individuality in contact

calls. In this study, I increase our knowledge of contact calls in belugas. The function of PS1 as a contact call and the form of individuality in PS1 are presented.

First, vocalizations of isolated belugas were examined in the Port of Nagoya Public Aquarium, and the group-cohesion function and individual distinctiveness of PS1 calls are represented (chapter 2). Second, PS1 calls were collected from other captive belugas kept at Shimane Aquarium and I verified the function and individuality of PS1 calls suggested in chapter 2 were common features in belugas (chapter 3). As a next step, playback experiments would be needed to elucidate whether belugas use PS1 for individual recognition and which acoustic parameter is the recognition cue. However, conventional speakers cannot reproduce broadband pulse trains such as PS1. Therefore, in chapter 4, a broadband transmitting system was established for playback experiments of PS1 calls. A trial of the PS1 playback experiment was also represented. Chapter 5 summarizes this study and describes considerations for future playback experiments. Further, the evolutionary pathway of odontocete contact calls is discussed.

2. Function and individuality in PS1 calls of belugas

2.1 Introduction

Individually specific contact calls in belugas have been reported at the Port of Nagoya Public Aquarium by Morisaka et al. (2013). However, the study analyzed only a small amount of data taken over a short time period. Vocalizations were collected in various contexts; therefore, the data may include context effects on the call rate or acoustic parameters. Further, the study focused only on the temporal domain for comparison of individuals and spectral characteristics were not described. Thus, it would be useful to investigate PS1 calls of belugas in the aquarium to evaluate their usage as contact calls and explore individual distinctiveness in various acoustic parameters.

The production of individually distinctive contact calls can be reinforced in isolation contexts where isolated and non-isolated individuals call to inform one another of their presence; such contexts have been used to identify signature whistles in bottlenose dolphins. Caldwell et al (1990) revealed that signature whistles of most of the captive dolphins represented more than 90% of all whistles produced in such isolation contexts. Signature whistles produced by temporarily captured mothers and calves also accounted for 73%–92% of all whistles (Sayigh et al., 1990). Subsequent studies suggested context-dependent usage of signature whistles: Janik et al. (1994) showed that signature whistles were most common (80% of all whistles) in isolation but their frequency was reduced to approximately 50% in a trained discrimination task where a dolphin chose one of two objects presented simultaneously. Janik & Slater (1998) reported that isolated and non-isolated captive dolphins produced signature whistles most frequently (32%–92% of all whistles) when one individual swam separately from the other members in a pool, but they produced almost only non-signature whistles (98% of all whistles) when all individuals swam together in the same pool. Watwood et al. (2005) demonstrated that signature whistles comprised 59%–100% of all whistles in temporarily restrained dolphins,

while the rate was reduced to an average of 39% in free-swimming contexts. Further, allied males produced signature whistles more frequently (56% on average) of all whistles when they were voluntarily separated from their partners than when they were together (30% on average) and during consortships with females (25% on average).

The frequency and individual distinctiveness of PS1 calls in isolation were therefore important indicators for suggesting whether PS1 calls have a function in group cohesion and individual advertisement. Thus, in the present study, an isolation context was provided for each beluga with the exception of the mother and calf (Mishima et al., 2015). PS1 calls from a calf were also collected in a context where he and his mother or another subadult were separated from the other belugas. I examined the frequency of PS1 calls and individuality in both temporal and spectral parameters.

2.2 Materials and methods

2.2.1 Facility and subjects

Recordings were performed at the Port of Nagoya Public Aquarium, Aichi, Japan between September 2013 and May 2014. The facility held five belugas, one adult male (NM1); two adult females (NF1, NF2); one subadult female (NF3); and one male calf (NM2). Figure 2.1 shows the genealogy of the beluga group. NM1 came from the White Sea, Russia in 2001 and his age was estimated to be 36 years old. NF1 and NF2 came from the Russian Far East in 2001 and their ages were estimated to be 19 and 15 years old, respectively. NF3 and NM2 were born in the aquarium and were half-siblings. NF3 is the daughter of NM1 and NF1, and was six years old. NM2 is the son of NM1 and NF2, and was 13 months old. All ages given are correct as of September 2013.

A schematic view of the beluga pool is represented in Fig. 2.2. The main pool (308 m^2 surface area and 6.3 m deep) is connected to two sub-pools: a holding pool (52 m^2 surface area and

5.1 m deep) and a medical pool (53 m^2 surface area and 5.0 m deep). The three pools were separated by metal lattices but were linked acoustically.

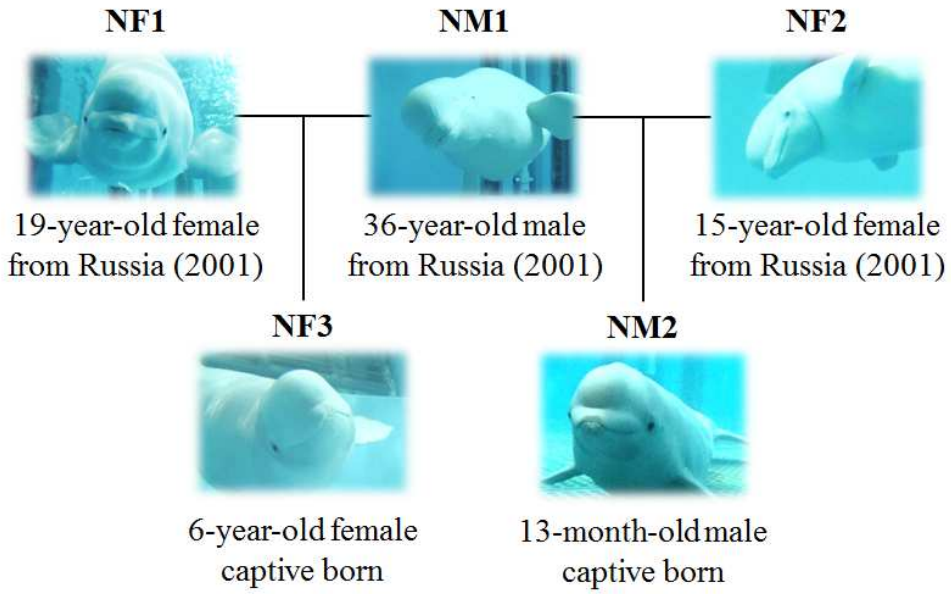


Fig. 2.1 Genealogy of the belugas kept at the Port of Nagoya Public Aquarium as of September 2013.

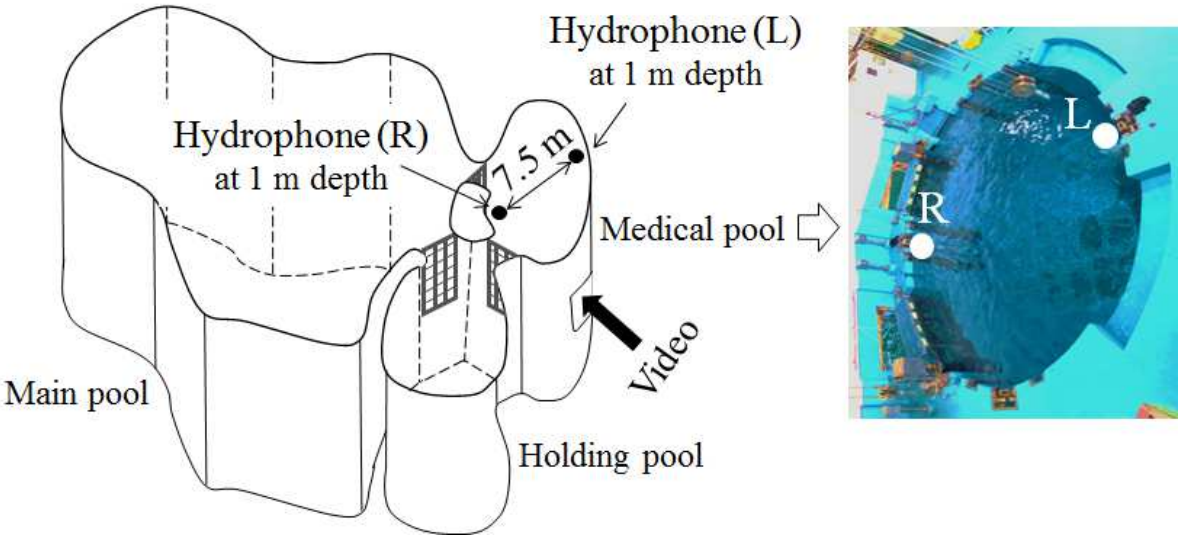


Fig. 2.2 Schematic view of the beluga pool in the Port of Nagoya Public Aquarium.

2.2.2 Data collection

Each of the three belugas, except the mother NF2 and the suckling calf NM2, were isolated and kept alone in the medical pool for recordings. NF2 and NM2 could not be individually separated; therefore, they were kept together in the medical pool. In May 2014, NF2 and NM2 were separated, because NF2 came into estrous. During this time, the three adults were kept together in the main pool and NM2 and NF3 were kept together in the medical pool. It was the first long-term separation between the mother–calf pair; therefore, I predicted that they would frequently produce PS1 calls. There were thus five isolation patterns: “NM1,” “NF1,” “NF3,” “NF2 & NM2,” and “NF3 & NM2.” Recordings were conducted between 08:00 and 18:00, excluding feeding and training times. Each recording continued for 30 min, with the exception of two 20-min sessions. A total of 46 recording sessions over 22 h and 40 min were performed (Table 2.1).

Table 2.1 Number of sessions and total recording time for each isolation pattern.

Isolation pattern	NM1	NF1	NF3	NF2&NM2	NF3&NM2
Total recording time [min.] (Number of sessions)	560 (19)	180 (6)	270 (9)	290 (10)	60 (2)

Acoustic recordings were performed using two TC 4013 hydrophones (Reson Inc., Denmark) covered with polyvinyl chloride (PVC) pipes. They were installed at the right and left sides of the medical pool at 1 m depth and spaced 7.5 m apart from one another (Figure 2.2). These hydrophones exhibit a flat frequency response from 1 Hz to 170 kHz (-211 ± 3 dB re 1 µPa/V at 1 m). The hydrophone signals passed through the analog bandpass filter from 1 kHz to 200 kHz, and were amplified by 50 dB using an Aquafeeler III preamplifier (AquaSound Inc., Japan) with a flat frequency response to 200 kHz (-3 dB). The outputs were connected to two separate channels of an EZ7510 data recorder (NF Co., Japan), which digitally converted the

analog signals from each channel sampling at 500 kHz and 16 bits. Observations were made from an underwater window of the medical pool using an iVIS HF R11 video camera (Canon Inc., Japan).

2.2.3 Call classification

The classification of calls was required in order to estimate the predominant call type in isolation contexts. Belugas produce various complex calls using tonal and pulsed components, although a standardized categorization method has not been established. Previous studies tend to classify using narrow categories and those categories differed among studies (Sjare & Smith, 1986a; Bel'kovitch & Sh'ekotov, 1993; Recchia, 1994; Karlsen et al., 2002; van Parijs et al., 2003; Belikov & Bel'kovich, 2006; Belikov & Bel'kovich, 2007; Belikov & Bel'kovich, 2008; Vergara et al., 2010; Panova et al., 2012; Chmelnitsky & Ferguson, 2012; Alekseeva et al., 2013).

In the present study, echolocation clicks, which have high directivity (Au et al., 1987) and have IPIs generally longer than 20 ms (Morisaka et al., 2013), were excluded from the data and remaining calls were divided into five broader categories. I classified the calls based on visual and aural inspection using Audacity version 2.0.5 (The Audacity Team). Spectrograms were generated with 1024-point FFT, a Hamming window function. I identified four commonly produced call types: a) one type of pulsed call "PS1," b) one type of combined call "C1," c) short calls "S," d) whistles "W," e) the others "O" category (Fig. 2.3). The acoustical definition of each call type was as described below (Mishima et al., 2015):

a) PS1: One type of pulsed call. To human ears, the fixed pulse train sounds like a ratchet or a door creaking and it is easy for humans to discriminate as a PS1 call (Morisaka et al., 2013). Energy distributes broadband from less than 1 kHz up to at least 170 kHz, and this call continues for more than 150 ms. I added a type of combined call, comprising a mixture of

pulsed and tonal components, irregularly to this category. The pulsed components resemble PS1, and the tonal components overlap in both temporal and frequency domains.

b) C1: One type of combined call. This call consists of two components: one is composed of high-frequency broadband pulses, and the other of low-frequency narrowband tones or low-frequency narrowband pulses with different pulse repetition patterns. The two components occur concurrently, but their frequency components do not overlap.

c) S: Short calls. Low-frequency and tonal calls with and without some discrete harmonics. The duration is less than 150 ms, and several calls are repeated within 100 ms.

d) W: Whistles. Low-frequency and tonal calls with and without some discrete harmonics. The duration is more than 150 ms.

e) O: Others. This group includes pulse trains other than PS1 such as pulsed tones (Watkins, 1967). This category also includes combined calls other than C1 such as graded calls with transitions from pulses to whistles (Murray et al., 1998), and noisy calls.

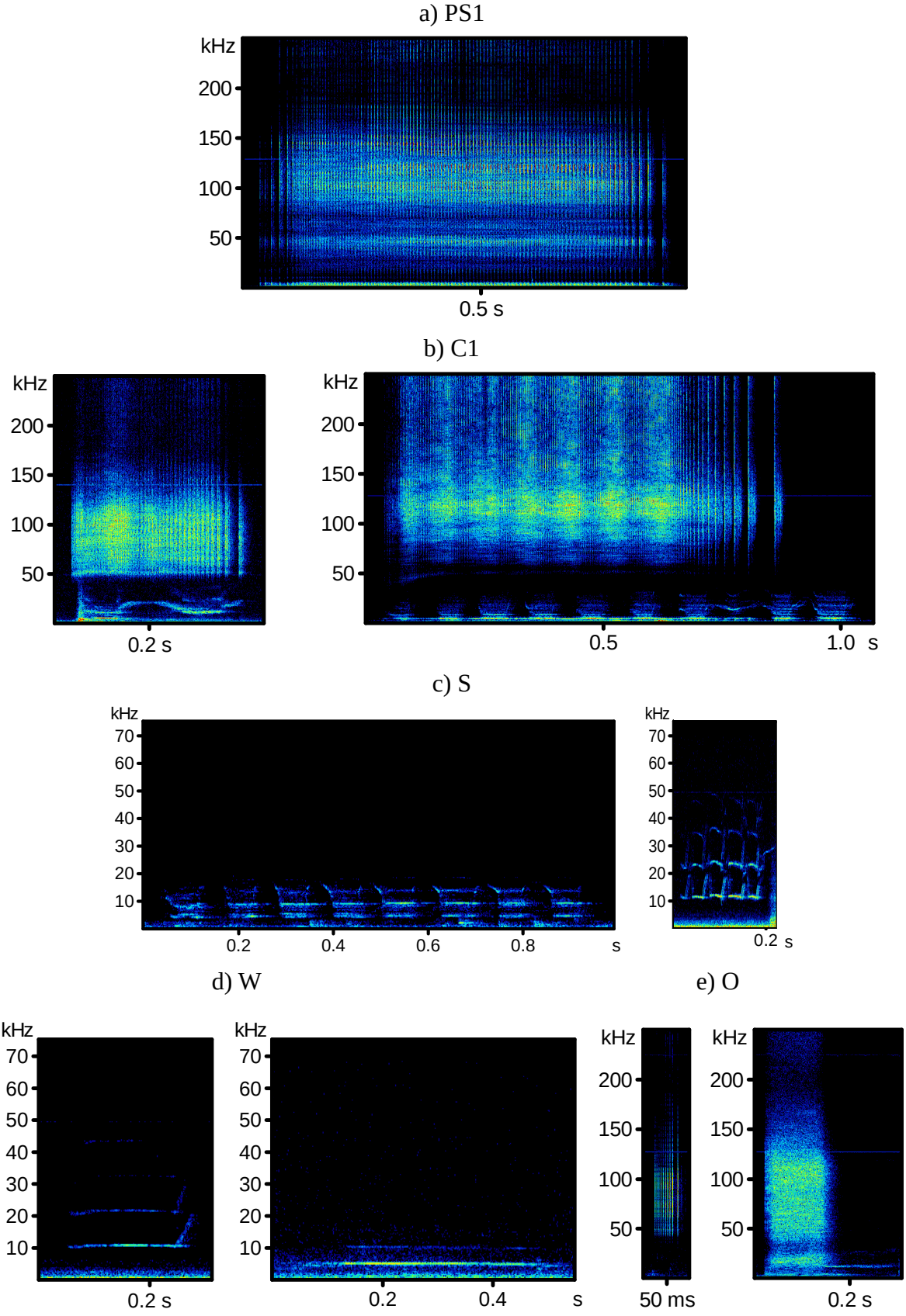


Fig. 2.3 Examples of classified call types: a) PS1, b) C1, c) S, d) W, and e) O. The horizontal and vertical scales vary among call types.

2.2.4 Caller identification

Sound arrival time differences recorded by the two hydrophones were calculated using custom written MATLAB software (2015) to discriminate between the PS1 calls from isolated and non-isolated belugas. The onset of PS1 was set to the threshold that was nearly three times greater than the noise level. The sound speed was approximately calculated as 1504 m/s using the Medwin equation at a salinity of 30.0 ppt and temperature of 16.1°C (Medwin 1975). The hydrophones were separated from one another by 7.5 m. Thus, the time difference was calculated to be between -5 and 5 ms approximately (Fig. 2.4).

Calls from beluga calves, including NM2, frequently co-occur with continuous emission of small bubbles referred as “bubble streams” (Vergara & Barrett-Lennard, 2008; Hill et al., 2011) (Fig. 2.5), which is similar to the whistles of bottlenose dolphin calves (McCowan & Reiss, 1995a; Killebrew et al., 2001; Morisaka et al., 2005a, 2005b). Therefore, concurrent bubble streams were also used to identify PS1 calls from NM2 in NF2 & NM2 and NF3 & NM2 isolation events.

In cases where there was ambiguity in identifying whether the PS1 production was that of an isolated animal, the call was counted but excluded from the isolated individual calls.

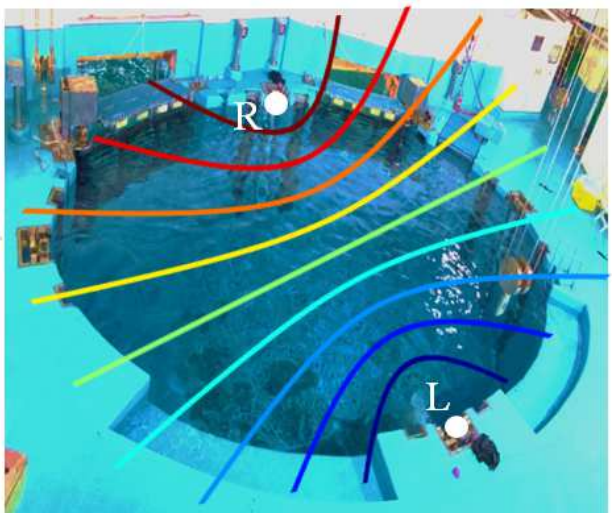


Fig. 2.4 Sound arrival time differences between the two hydrophones (R and L). There was no time difference on the green line. The interval between adjacent lines was 1 ms.

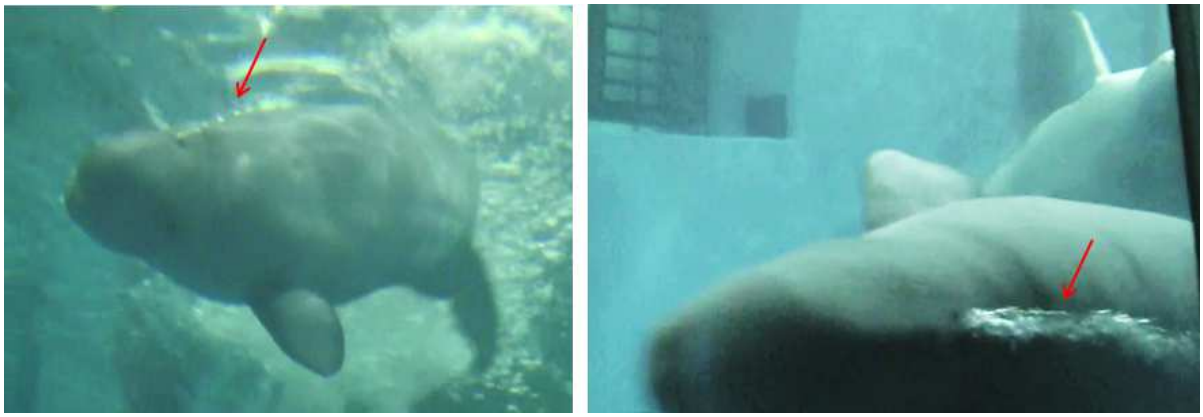


Fig. 2.5　Bubble streams co-occurred with PS1 production of NM2.

2.3 PS1 analyses

2.3.1 Acoustic parameters

Acoustic characteristics of PS1 were analyzed using Avisoft SASLab Pro version 5.2 (Avisoft Bioacoustics, Germany). If at least one of the PS1 calls from the two hydrophones had a good signal-to-noise ratio, the call with the best signal-to-noise ratio of the two sets of PS1 data was used in the analysis. Pulses composing PS1 were automatically detected by the "Pulse Train Analysis" function. In cases where reflecting pulses had been counted or where direct pulses with lower amplitude had not been counted, the miscounted pulses were corrected manually.

Five temporal parameters were extracted for individual comparison: the number of pulses (N_p); duration (DUR); pulse repetition rate (PRR); average inter-pulse intervals (IPIs) of pulse nos. 11–20 (IPI 1); and average IPIs of pulse nos. 11–20 from the last (IPI 2) (Fig. 2.6). N_p is the number of pulses composing PS1. DUR is time length from the peak of the first pulse to that of the last pulse in PS1. PRR is N_p divided by DUR. IPIs were indicated by the time differences from the peak of the preceding pulse to that of the following pulse and averages of the two parts of PS1, IPI 1 and IPI 2, were calculated. Univariate statistical comparison of these parameters was made using Kruskal–Wallis test or one-way ANOVA. Change in IPIs as a function of time, termed "IPI contours," were also depicted and compared visually.

Spectral analysis was performed using the “averaged power spectrum” function in the software. Power spectra were calculated at three pulse locations within PS1, the third pulse, the middle pulse, and the third pulse from the last. The power spectra of pulses were quantified for the 1.5 ms of data containing each direct pulse, and were calculated by a 256-point FFT with Hamming window. The spectra were then smoothed using a five-point window. To examine noise effect on the PS1 spectra, noise spectra were calculated using non-call windows before the onset of the PS1 calls. As a result, the noise level below 5 kHz was consistently high and considered to affect the PS1 spectra. The frequency range above 6 kHz was therefore used for further analyses, and the maximum source level (SL) in the range was set as zero to compare relative spectra. Four spectral parameters were extracted from the spectra: peak frequency (F_p); 10 dB bandwidth (10 BW), which is the frequency band at a level of -10 dB from the peak; the lower frequency of the 10 BW (F_l); and upper frequency of the 10 BW (F_u) (Fig. 2.7).

First, it was investigated whether there was spectral change along with the pulse order within PS1 for each individual. The relative power spectra were averaged at each pulse location and then the averaged spectra among the three pulse locations for each animal were compared using the four parameters. Second, the power spectra of the middle pulses were compared among individuals. Visual comparisons were made of the middle pulse spectra, and univariate statistical comparison of the parameters was carried out using Kruskal–Wallis test or one-way ANOVA. The temporal and spectral parameters are summarized in Table 2.2. All statistical analyses were performed using R software version 3.1.0 (The R Foundation for Statistical Computing).

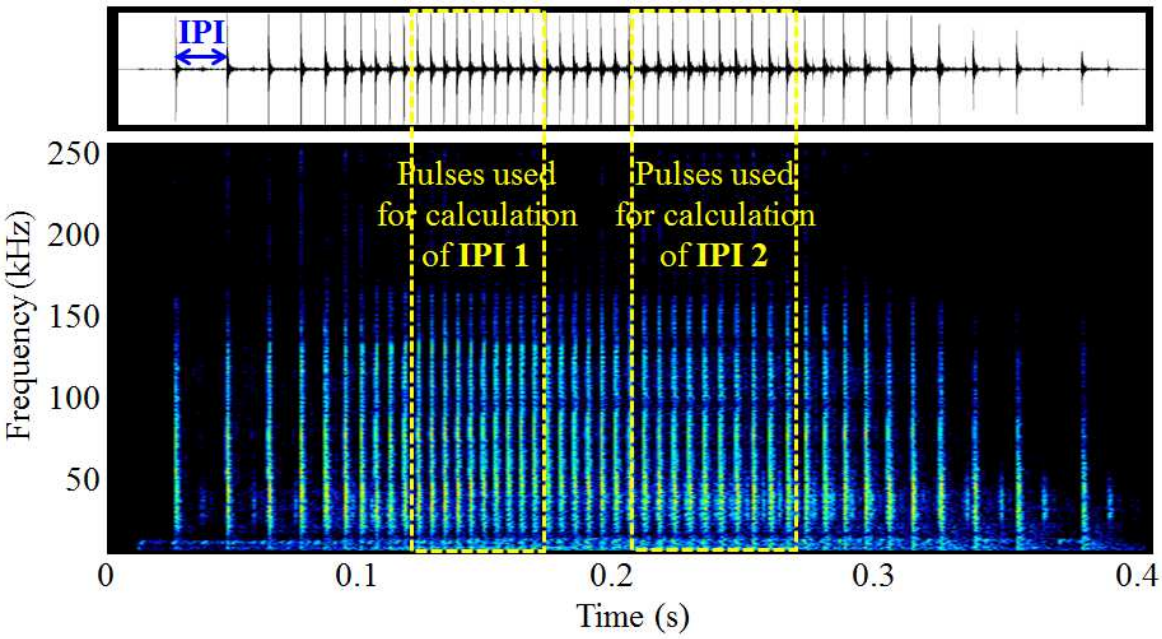


Fig. 2.6 Example of average IPI measurements. IPI 1 is the average IPI of pulse nos. 11–20, and IPI 2 is the average IPI of pulse nos. 11–20 from the last.

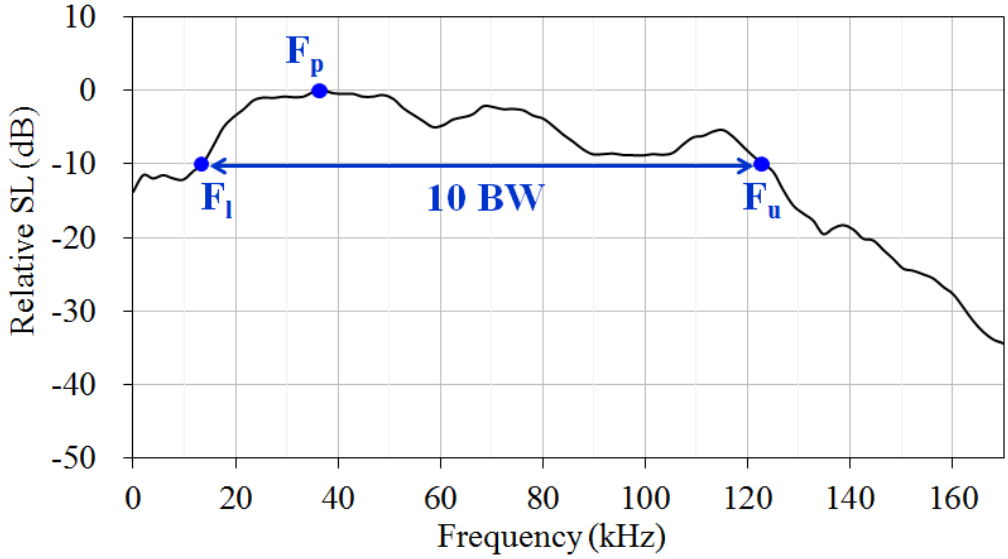


Fig. 2.7 Spectral parameters used for individual comparison: peak frequency (F_p); 10 dB bandwidth (10 BW), which is the frequency band at a level of -10 dB from the peak; the lower frequency of the 10 BW (F_l); and upper frequency of the 10 BW (F_u).

Table 2.2 Temporal and spectral parameters used for individual comparison of PS1 calls.

Parameters	Explanation
N_p	Number of pulses
DUR	Duration
PRR	Pulse repetition rate
IPI 1	Average IPIs of pulse nos. 11–20
IPI 2	Average IPIs of pulse nos. 11–20 from the last
F_p	Peak frequency
10 BW	10 dB bandwidth
F_l	Lower frequency of the 10 BW
F_u	Upper frequency of the 10 BW

2.3.2 Potential for individual coding

Acoustic cues encoding individual identity need to show high inter-individual variability and high intra-individual stereotypy to support individual recognition. Potential for individual coding (PIC) is the ratio of inter-individual variation to intra-individual variation (Robisson et al., 1993). PIC was used to identify which parameters have possibility to carry individual information in various species, for instance in birds (Robisson et al., 1993; Mathevon et al., 2003; Moscicki et al., 2011), primates (Lemasson et al., 2010; Salmi et al., 2014), and pinnipeds (Charrier et al., 2002; Charrier et al., 2010; Trimble & Charrier, 2011).

Coefficients of variation (CV) were used to describe the inter- and intra-individual variations of each measured parameter. CV between individuals (CV_b) was calculated according to the formula:

$$CV_b = SD/X \times 100 \qquad (2\text{-}1)$$

SD is standard deviation and *X* is the mean calculated for the overall samples. CV within individuals (CV_w) was calculated using the formula for small samples:

$$CV_w = SD/X \times (1 + 1/4n) \times 100 \qquad (2\text{-}2)$$

SD is the standard deviation, *X* is the mean, and *n* is the sample size for an individual. PIC was assessed as follows:

$$PIC = CV_b/meanCV_w \qquad (2\text{-}3)$$

The *mean* CV_w is the mean value of the CV_w for all individuals. Acoustic parameters showing a PIC score greater than 1 may be a useful feature for individual recognition since their inter-individual variability is greater than their intra-individual variability. This study assessed PIC for the five temporal and four spectral parameters of PS1, and evaluated which parameters might encode individuality.

2.3.3 Discriminant function analysis

Discriminant function analysis (DFA) was performed to classify PS1 calls into individuals based on the acoustic variables. Before running DFA, the multicollinearity, multivariate outliers, multivariate normality, and homogeneity of variance-covariance matrices were investigated.

The data set of the nine variables, N_p, DUR, PRR, IPI 1, IPI 2, F_p, 10 BW, F_l, and F_u, were prepared. The variance inflation factors (VIFs) were calculated for each parameter by using the “vif” function in the R package “car” to detect multicollinearity (Fox et al., 2015). There is a danger of multicollinearity when VIFs are greater than 4. The parameters with high VIF scores were excluded from the data set to obtain VIFs less than 4 from the remaining parameters. Potential multivariate outliers on the remaining data set were searched for using robust Mahalanobis distances with a 97.5% quantile (Nordhausen et al., 2008). The detected outliers were excluded, but if most PS1 calls of a particular individual were regarded as

outliers, they were included in the DFA. Multivariate normality was examined by Shapiro–Wilk test using the R package “mvnormtest” (Jarek, 2015), and the homogeneity of variance-covariance matrices was inspected by Box’s M test using the “powerTransform” function in the R package “car” (Fox et al., 2015). Because the data set did not satisfy multivariate normality and homogeneity of variance-covariance matrices, quadratic DFA was performed using the “qda” function in the R package “MASS” (Ripley et al., 2015). The performance of quadratic DFA was quantified using a jack-knife leave-one-out cross-validation. Stepwise DFA using the “stepclass” function in the R package “klaR” was also executed to find the most informative parameters (Weihs et al., 2005).

2.4 Results

2.4.1 Frequency of each call type

A total of 6817 calls, including 2633 PS1 calls, 1202 C1 calls, 793 S calls, 628 W calls, and 1561 O calls were recorded from both the isolated and the remaining belugas in 46 isolation events. PS1 was the predominant call type, and occupied 38% of the total calls followed by 18% of S, 12% of C1, and 9% of W (Fig. 2.8).

The number of each call type was variable among sessions as shown in Fig. 2.9. PS1 did not account for the highest percentage during every separation and in several sessions PS1 did not occur. Session nos. 45 and 46 corresponding to NF3 & NM2 isolation events, or the NF2–NM2 (mother–calf) separation, had the highest call numbers. By excluding the last two sessions, the percentage of each call type was 30% of PS1, 21% of S, 15% of C1, and 10% of W, with PS1 being the major call type.

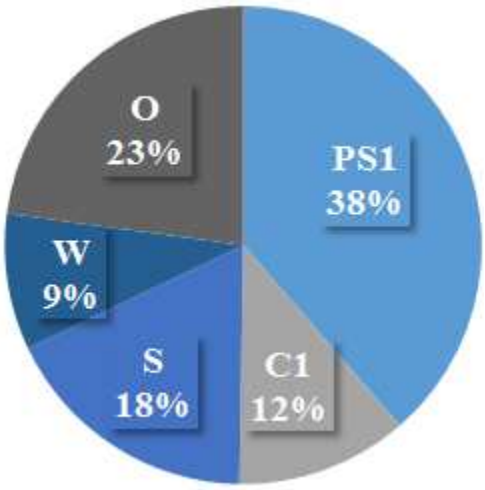


Fig. 2.8 Proportion of each call type in total sessions.

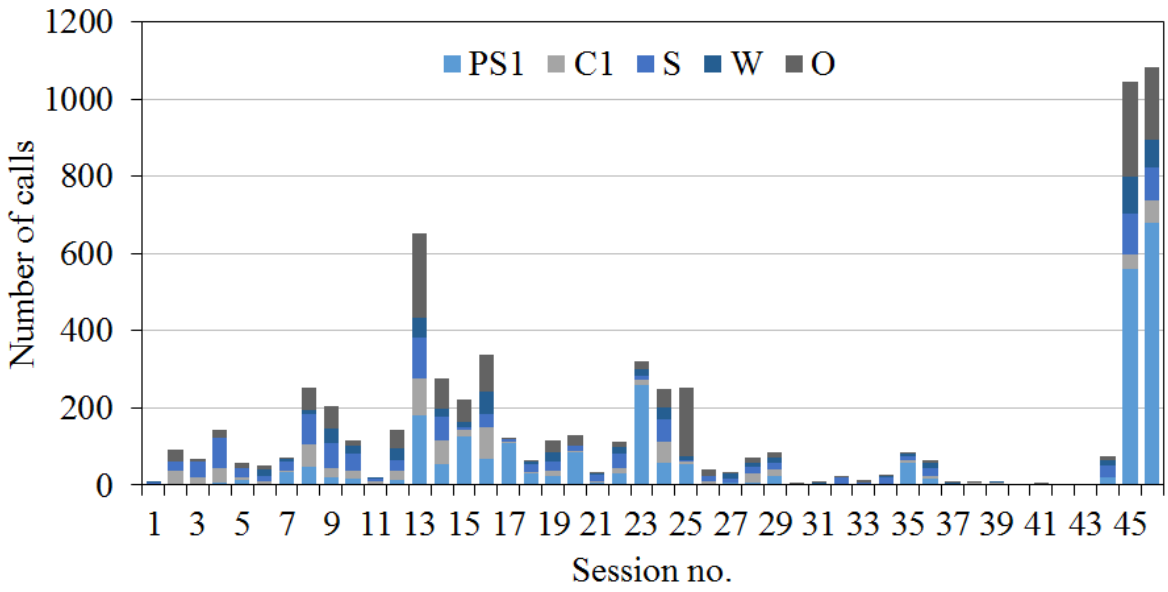


Fig. 2.9 Frequency of each call type per session.

2.4.2 Individuality in PS1

Of the 2633 PS1 calls, 647 calls were identified as the production of the isolated belugas, including 24, 156, 331, 56, and 80 for NM1, NM2, NF1, NF2, and NF3, respectively. NF2 had the least number of isolation events but a high number of PS1 calls were obtained from her.

Of the 647 calls identified from isolated belugas, 187 were further analyzed, including 16, 33, 97, 20, and 21 for NM1, NM2, NF1, NF2, and NF3, respectively. Figure 2.10 presents

waveforms and spectrograms of individual PS1 calls. Most of the PS1 calls from NM1 contained tonal components at around 13 kHz, and this structure was not found in any of the PS1 calls from the other individuals.

Temporal characteristics of PS1 for each beluga are shown in Table 2.3. The Kruskal–Wallis test revealed that N_p, PRR, IPI 1, and IPI 2 were significantly different among individuals ($P < 0.0001$), but that DUR did not differ significantly ($P = 0.316$). PIC values greater than 1 were found in N_p, PRR, and IPI 1, with PRR holding the highest value of 1.57.

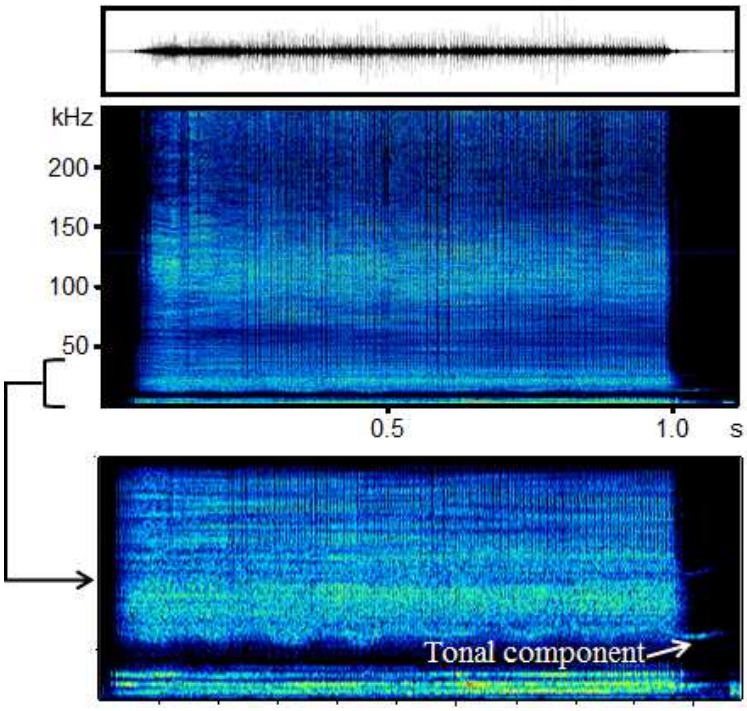
Sixteen examples of IPI contours are depicted for each individual in Fig. 2.11. The subjects, excluding the calf NM2, exhibited individually unique IPI contours which were stereotyped intra-individually and differed inter-individually. NM1 had consistent lower IPIs at the beginning, and his IPI contours were easily discriminated from others. In contrast, the three females (NF1, NF2, and NF3) had similar IPI contours, but they contained slight differences. The IPIs of NF1 and NF2 decreased at approximately the same time from the initial pulses, but the slopes of IPI contours of NF1 were gentler than those of NF2. On the other hand, the slopes of IPI contours in NF1 and NF3 resembled each other, but a decrease in the IPIs of NF3 occurred more rapidly from the initial pulses than those of NF1. No stereotypy was found in the IPI contours of NM2 and the calf's IPIs greatly fluctuated over the duration.

Power spectra were averaged at each pulse location and the averaged spectra were compared among the three pulse locations for each individual. Spectral parameters calculated in the range above 11 kHz were used because the averaged power spectra of all belugas showed similar patterns with an energy peak at 6 kHz below 10 kHz. No individual exhibited high degrees of difference in F_p, 10 BW, F_l and F_u depending on pulse locations (Table 2.4). This suggests that there was no distinctive spectral change along with the pulse order in PS1 calls. Power spectra of middle pulses were thus selected as representatives of the pulses comprising PS1 calls and were used for individual comparison.

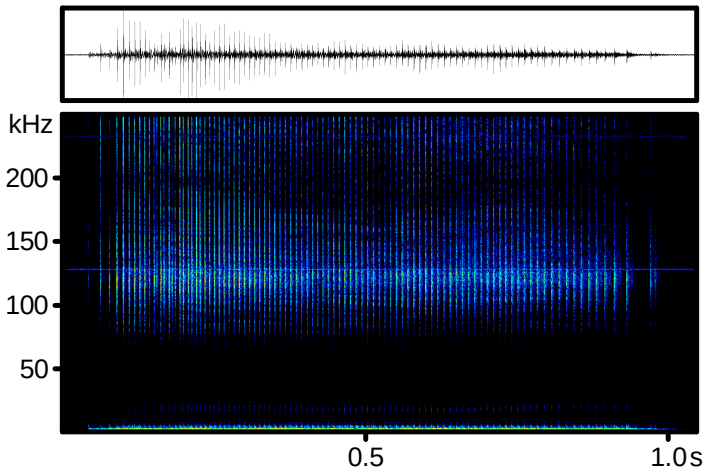
Figure 2.12 presents sixteen examples of the power spectra of the middle pulses. Visual comparisons suggest that, unlike IPI contours, the power spectra did not have obvious individual distinctiveness and consistence. Characteristics in spectral parameters are shown in Table 2.5. A one-way ANOVA revealed that F_p, F_l, and F_u were significantly different among individuals ($P < 0.0001$), although 10 BW did not differ significantly ($P = 0.696$). All spectral parameters had a PIC score greater than 1.

For DFA, multicollinearity was investigated. The parameters, N_p, 10 BW, F_l, and F_u had high VIF values. Excluding them from the data set, the remaining five parameters had VIFs less than 2, and the danger of multicollinearity was eliminated. In DFA, the number of variables had to be lower than one third of the least number of individual samples (McGarigal et al., 2000). NM1 had only 16 samples, therefore five variables could be used for the DFA and the number of the remaining variables fulfilled it. Potential multivariate outliers were found in the remaining data set (Fig. 2.13). Since all of the PS1 calls from NM1 and most of the PS1 calls from NM2 were regarded as outliers, they were included in the DFA, with the exception of the outstanding one from NF1. The samples from NF1 needed to be reduced to decrease disparity in sample size and increase the effectiveness of the DFA (McGarigal et al., 2000), thus 33 samples were randomly selected, which was the same number used for NM2, and sample size fell into the range of 16-33. Because the data set did not satisfy multivariate normality (Shapiro–Wilk test, $P < 0.0001$) and homogeneity of variance-covariance matrices (Box's M test, $P < 0.0001$), quadratic DFA was performed. The quadratic DFA based on five variables, DUR, PRR, IPI 1, IPI 2, and F_p, resulted in correct classification rates of 87.5%, 78.8%, 81.8%, 70.0%, and 85.7% for NM1, NM2, NF1, NF2, and NF3, respectively, with an overall correct classification rate of 80.5% (Table 2.6). The most powerful discriminator detected by stepwise DFA was IPI 1, followed by F_p.

a) NM1


kHz
200
150
100
50
0.5
1.0
s
Tonal component

b) NM2


kHz
200
150
100
50
0.5
1.0s

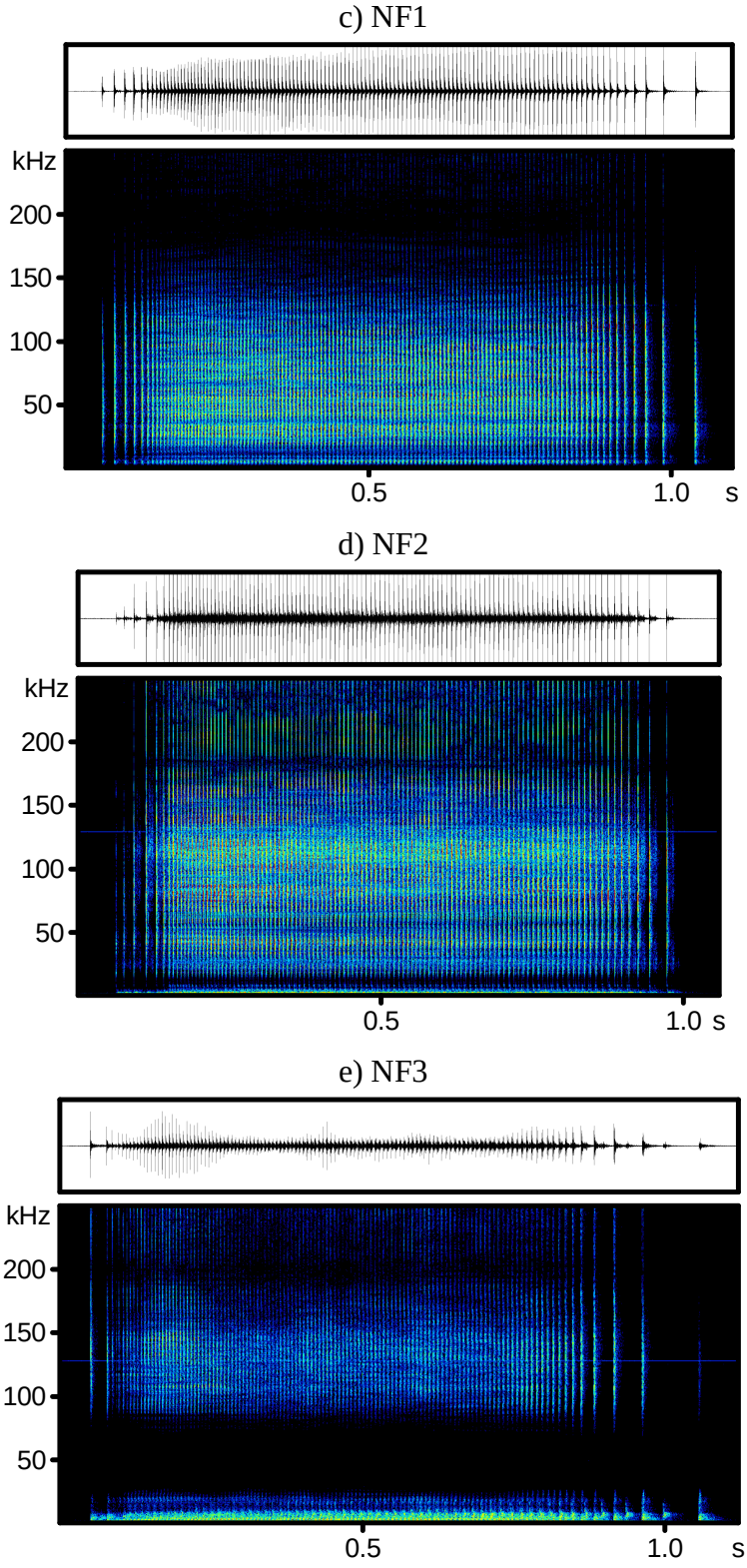


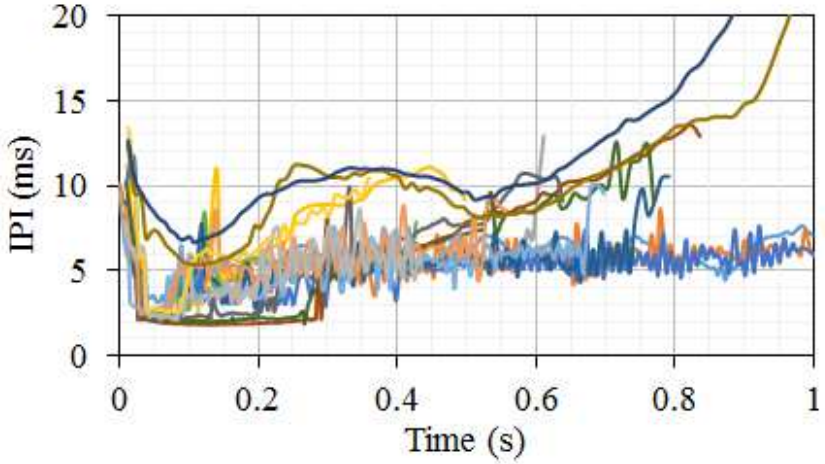
Fig. 2.10 Examples of PS1 calls from five belugas: a) NM1, b) NM2, c) NF1, d) NF2, and e) NF3. The top graphs represent waveforms, and the bottom graphs represent specrograms (FFT size: 1024 points; window: Hamming; overlap: 50%). The PS1 call from NM1 contained tonal component at around 13 kHz.

Table 2.3 Temporal characteristics of PS1 calls.

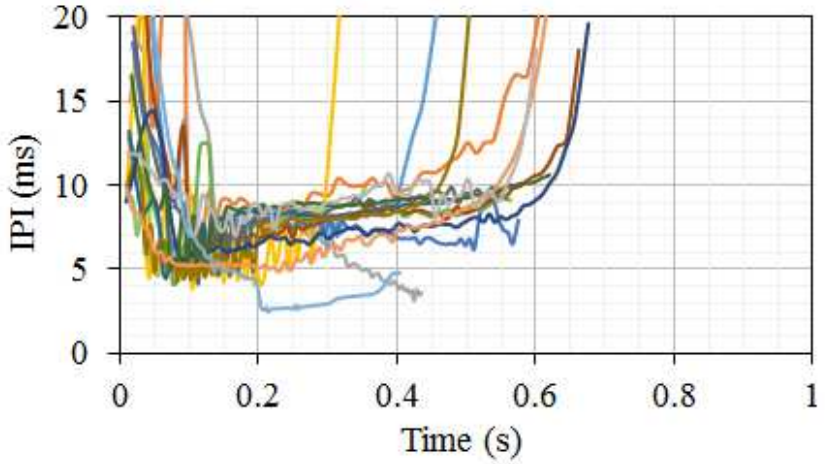
ID	n	N_p			DUR (s)			PRR (pulses/s)			IPI 1 (ms)			IPI 2 (ms)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
NM1	16	144.2 ± 48.4	234	64	0.77 ± 0.23	1.13	0.36	190.8 ± 41.8	248.9	90.1	2.92 ± 1.32	6.96	2.05	7.76 ± 2.38	13.06	5.67
NM2	33	80.2 ± 26.0	180	43	0.68 ± 0.18	1.13	0.36	119.7 ± 25.8	182.1	40.9	6.85 ± 1.18	10.8	4.96	8.24 ± 1.61	11.07	4.42
NF1	97	83.5 ± 31.9	154	19	0.76 ± 0.24	1.48	0.25	110.1 ± 27.7	220.9	32.6	5.69 ± 1.55	19.58	4.48	7.12 ± 0.61	8.42	4.91
NF2	20	100.5 ± 23.0	145	46	0.80 ± 0.22	1.18	0.36	128.3 ± 11.3	152.4	111.2	5.67 ± 0.40	6.4	4.92	7.91 ± 0.86	8.93	5.9
NF3	21	82.6 ± 40.3	144	32	0.79 ± 0.36	1.36	0.29	104.1 ± 15.3	137	68.2	5.96 ± 0.25	6.53	5.48	7.6 ± 0.98	9.38	6.14
Kruskal–Wallis test		$\chi^2 = 26.21$, $P < 0.0001$			$\chi^2 = 4.73$, $P = 0.32$			$\chi^2 = 45.30$, $P < 0.0001$			$\chi^2 = 76.65$, $P < 0.0001$			$\chi^2 = 26.20$, $P < 0.0001$		
CV_b		41.58			32.76			29.13			28.18			16.37		
Mean CV_w		35.52			32.66			18.60			20.41			16.66		
PIC		1.17			1.00			1.57			1.38			0.98		

Note: N_p, number of pulses; DUR, duration; PRR, pules repetition rate; IPI 1, average inter-pulse interval of pulse nos. 11–20; IPI 2, average inter-pulse interval of pulse nos. 11–20 from the last; CV_b, coefficient of variation between individuals; mean CV_w, mean of the coefficients of variation within individuals; PIC, potential for individual coding.

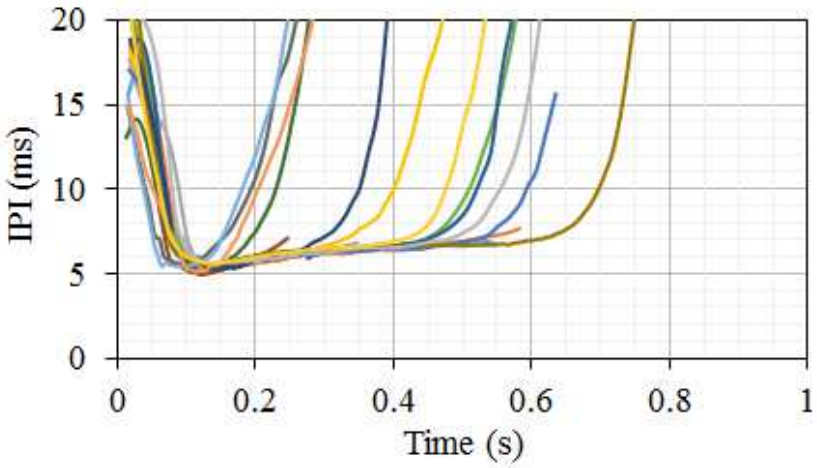
a) NM1


20
15
10
5
0
IPI (ms)
0
0.2
0.4
0.6
0.8
1
Time (s)

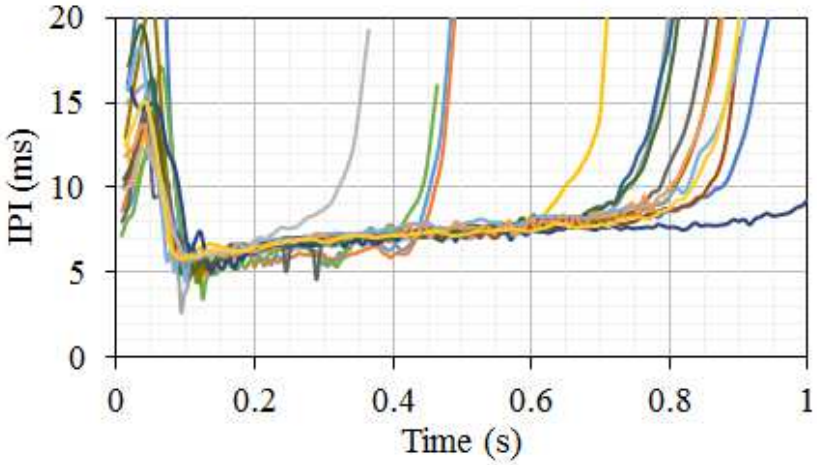
b) NM2


20
15
10
5
0
IPI (ms)
0
0.2
0.4
0.6
0.8
1
Time (s)

c) NF1



d) NF2



e) NF3

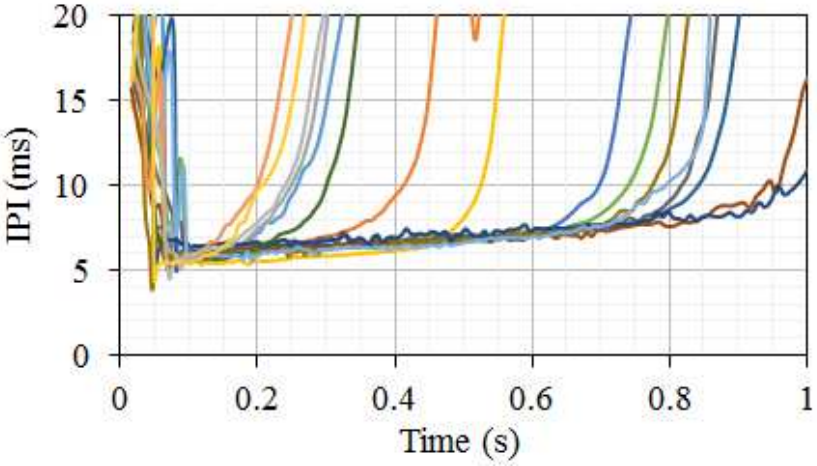


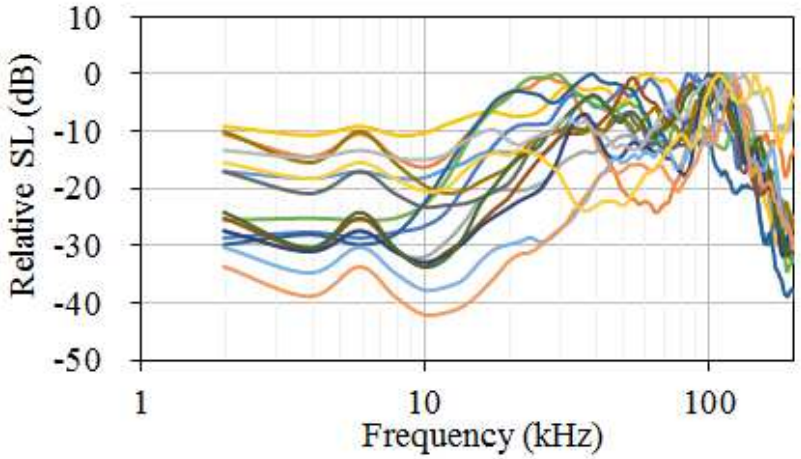
Fig. 2.11 IPI contours of PS1 calls from five belugas (n = 16): a) NM1, b) NM2, c) NF1, d) NF2, and e) NF3. Sixteen examples each from NM2, NF1, NF2, and NF3 were randomly selected to match the number of depicted IPI contours of NM1 from which the smallest samples were collected.

Table 2.4 Characteristics of the averaged power spectra at three pulse locations along the third pulse, the middle pulse, and the third from the last of PS1 calls.

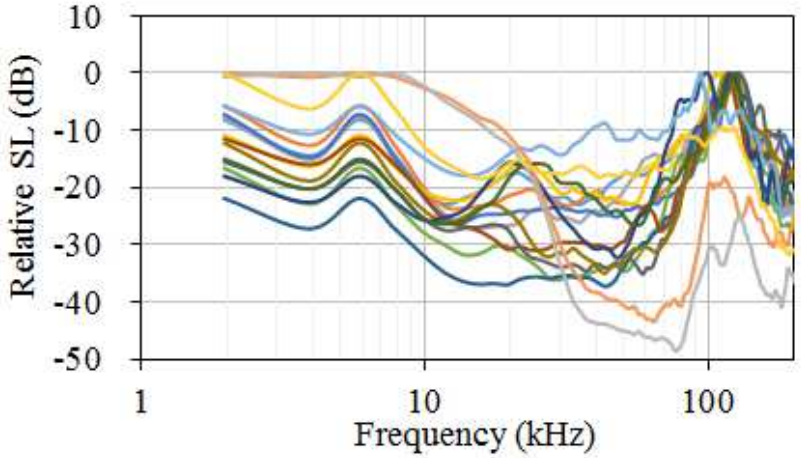
	ID	NM1	NM2	NF1	NF2	NF3
	n	16	33	97	20	21
F_p (kHz)	3rd	109	115	29	113	109
	Middle	113	117	29	117	107
	3rd from the last	109	117	29	115	109
10 BW [F_l - F_u] (kHz)	3rd	133 [14-147]	72 [80-152]	113 [12-125]	72 [76-148]	88 [70-158]
	Middle	126 [22-148]	64 [86-150]	111 [12-123]	70 [78-148]	82 [72-154]
	3rd from the last	119 [22-141]	64 [82-146]	74 [12-86]	72 [76-148]	88 [68-156]

Note: F_p, peak frequency; 10 BW, 10 dB bandwidth; F_l, the lower frequency of the 10 BW; F_u, the upper frequency of the 10 BW.

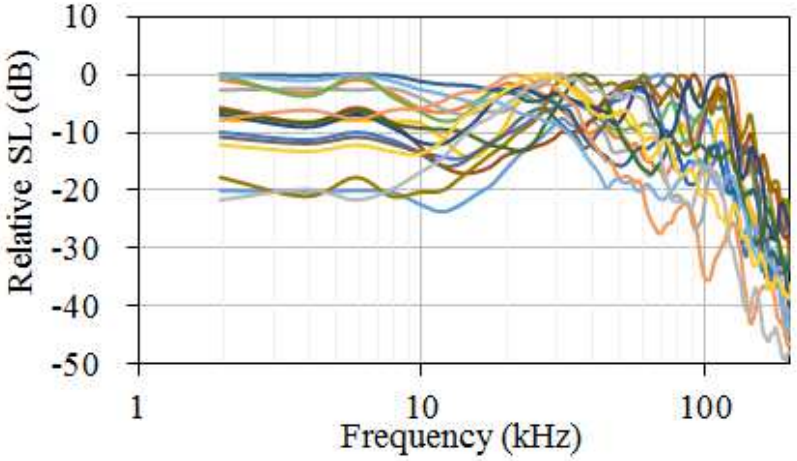
a) NM1



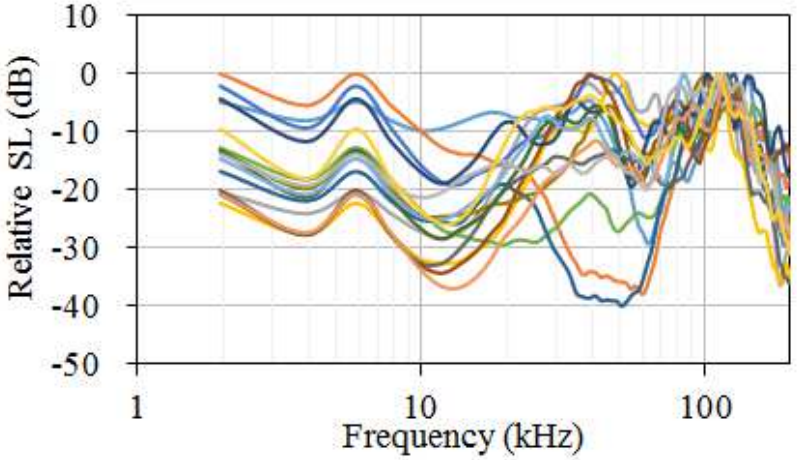
b) NM2



c) NF1



d) NF2



e) NF3

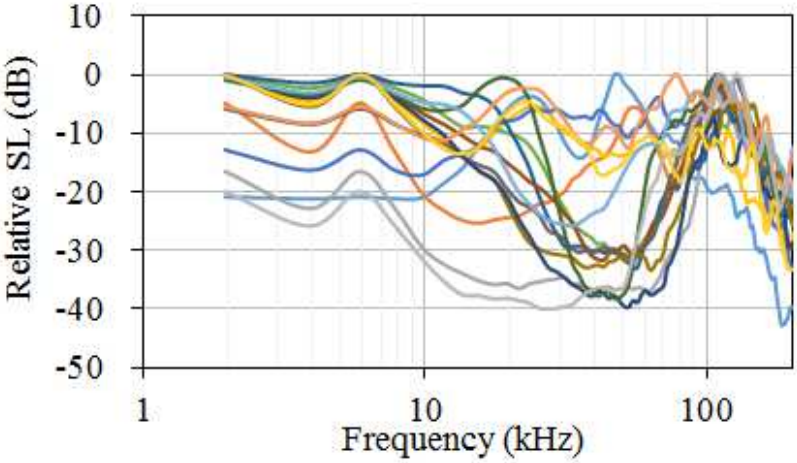


Fig. 2.12 Power spectra calculated at the middle pulse location within PS1 calls from five belugas (n = 16): a) NM1, b) NM2, c) NF1, d) NF2, and e) NF3 (FFT size: 256 points; window: Hamming; smoothing: 5 points). Sixteen examples each from NM2, NF1, NF2, and NF3 were randomly selected to agree with the number of depicted power spectra of NM1 from which the smallest samples were collected.

Table 2.5 Spectral characteristics of the middle pulses within PS1 calls.

ID	n	F_p (kHz)			10 BW (kHz)			F_l (kHz)			F_u (kHz)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
NM1	16	98.4 ± 31.0	146.5	29.3	61.3 ± 28.1	117.2	23.4	65.2 ± 33.6	109..4	11.7	126.5 ± 28.2	169.9	62.5
NM2	33	108.0 ± 31.6	134.8	11.7	51.3 ± 25.1	134.8	5.9	79.4 ± 26.8	105.5	11.7	130.7 ± 37.1	168.0	17.6
NF1	97	46.2 ± 28.8	119.1	11.7	56.9 ± 30.4	119.1	0.0	23.7 ± 19.9	95.7	11.7	80.6 ± 34.6	138.7	11.7
NF2	20	104.1 ± 23.3	127.0	39.1	60.2 ± 23.7	123.0	21.5	75.9 ± 27.8	107.4	21.5	136.0 ± 25.3	164.1	50.8
NF3	21	91.1 ± 41.3	132.8	11.7	60.1 ± 23.6	113.3	7.8	63.5 ± 31.3	105.5	11.7	123.6 ± 42.1	175.8	19.5
One-way ANOVA		$F(4,182) = 37.87$, $P < 0.0001$			$F(4,182) = 0.555$, $P = 0.70$			$F(4,182) = 45.10$, $P < 0.0001$			$F(4,182) = 23.01$, $P < 0.0001$		
CV_b		57.01			49.36			74.64			40.83		
Mean CV_w		38.54			45.83			51.59			29.52		
PIC		1.48			1.08			1.45			1.38		

Note: F_p, peak frequency; 10 BW, 10 dB bandwidth; F_l, the lower frequency of the 10 BW; F_u, the upper frequency of the 10 BW; CV_b, coefficient of variation between individuals; mean CV_w, mean of the coefficients of variation within individuals; PIC, potential for individual coding.

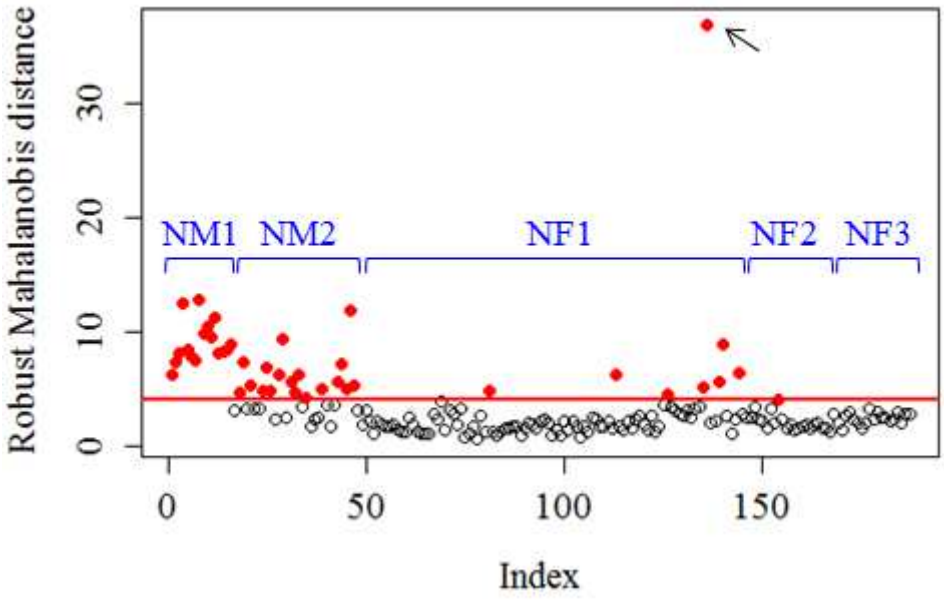


Fig 2.13 Multivariate outliers searched by robust Mahalanobis distances with 97.5% quantile. The samples are represented from the left in the order of NM1, NM2, NF1, NF2, and NF3. Red circles show the outliers. The arrow indicates the outstanding outlier from NF1.

Table 2.6 Results of the quadratic DFA based on five parameters: DUR, PRR, IPI 1 IPI 2 and F_p.

	NM1	NM2	NF1	NF2	NF3	n	Correct rate (%)
NM1	14	2	0	0	0	16	87.5
NM2	1	26	2	4	0	33	78.8
NF1	0	2	27	0	4	33	81.8
NF2	0	2	2	14	2	20	70.0
NF3	0	0	3	0	18	21	85.7

2.5 Discussion

This chapter investigated frequency and individual distinctiveness of PS1 calls in isolation in order to elucidate whether PS1 calls served as a function in group cohesion and which parameter encoded individuality. PS1 was the most frequently produced call type in the total isolation events (Fig. 2.8), which suggests that PS1 was used for group cohesion. The PS1 call rate was variable among sessions and no PS1 calling was found in some sessions (Fig. 2.9). Janik & Slater (1998) stated that, “Signature whistles were the most common whistles in the isolation context but did not occur during every separation.” The rate of PS1 production was possibly related to various factors, such as activity state, presence of trainers at the poolside, and the habituation through separation noted in signature whistles in bottlenose dolphins (Esch et al., 2009). The belugas, excluding the calf, were practically trained to be isolated in the experimental pool, and they might have become accustomed to being segregated from their pool mates. The time from separation to the start of the recordings varied among sessions, and it also seemed to cause different degrees of habituation. Several factors might influence their motivational state, and may have introduced variability to PS1 call rate. The remarkably high numbers of PS1 calls in session nos. 45 and 46 were likely due to the first long separation between the mother NF2 and the calf NM2 (Fig. 2.9). The calf dramatically produced PS1 calls at the metal lattice once he was segregated from his mother, while he only produced three PS1 calls in all of the NF2 & NM2 sessions when separated with his mother. This was a strong indication that PS1 calls had a function in mother–calf cohesion. Vergara & Barrett-Lennard (2008) and Vergara et al. (2010) supported this possibility, and reported a high number of mother–calf contact calls whenever a mother needed to regain or maintain contact with her calf. NF2 and NM2 needed to exchange PS1 calls frequently to contact each other, and it might have activated the vocalizations of other individuals. However, the possibility that the high call rate was associated with the reproductive season could not be ruled out. The two sessions

were made in May when three adults were reproductively active, and season might be related with increase in vocal activity.

Many samples were collected from the adult female NF1 despite the smallest number of sessions (Table 2.1). As NF1 was a strong character, and was the oldest and largest female, the high call rate might be related to the social rank. She appeared to be the top-ranking among females based on previously utilized criteria (Recchia, 1994), although dominance was not specifically measured. Dominant belugas possibly produce PS1 calls at higher rates to instigate and control the movement of other group members, as seen in black-billed gulls (*Larus bulleri*) (Evans, 1982) and green woodhoopoes (*Phoeniculus purpureus*) (Radford, 2004).

PS1 spectrograms revealed that only PS1 calls from the adult male NM1 had tonal components at around 13 kHz (Fig. 2.10). Of the five Type A call variants (A1-A5) described by Vergara et al. (2010), these PS1 calls are similar to A1, which was produced by an adult female and her two offspring. A1 has an average PRR of 94.6 pulses/s, 1.2–1.9 s in duration, and it consistently contains a tonal component at 14.6 kHz. The A4 call overlaps with A1 in PRR and duration but lacks the tonal component, and is similar to PS1 calls produced by individuals other than NM1 in the present study. Therefore, some of the Type A calls may be regarded as PS1 calls. A larger sample size is needed to find out the role of the common tonal components in PS1 and Type A calls.

Individual difference existed in the IPI contours of PS1 calls. NM1 had apparently specific IPI contours, and NF1, NF2, and NF3 had similar, but distinct IPI contours (Fig. 2.11). The belugas might recognize these slight temporal differences, since they have a high temporal resolution of around 1400 Hz in cut-off frequency (Klishin et al., 2000).

The consistent patterns in the initial part of IPI contours were possibly more informative (Fig. 2.11). This was supported by the PIC results showing that IPI 1 had a PIC value greater than

1, but the PIC value of IPI 2 was lower than 1 (Table 2.3). It was also inferred by vocal exchange pattern. Responses occurred within 1 s in 88% of PS1 exchanges by two different individuals, and most of those were overlapping exchanges, or instances when the second caller responded before the termination of the first caller's PS1 (Morisaka et al., 2013). This suggests that belugas recognize and reply to the initial part of PS1 calls. Likewise, Vergara et al. (2010) described that Type A calls were used in antiphonal call matching exchanges, and they also contained overlapping exchanges.

NM2 produced temporally fluctuating PS1 calls and they co-occurred with bubble streams (Fig. 2.5; Fig. 2.11). NM2 was aged 21 months in NF3 & NM2 session nos. 45 and 46 when several PS1 calls were collected from NM2. It was inferred that he was still in the course of vocal production learning (Janik & Slater, 2000) and/or morphological development. This hypothesis was supported by the following facts: male beluga offspring develop the similar type of pulse train described here until they are at least one year old, and the vocal development process continues past his first year of life (Vergara & Barrett-Lennard, 2008); strong mother–calf bonds last until calves are at least three years old (Krasnova et al., 2014) and wild beluga calves live in fluid matrilineal groups (Michaud, 2005); and captive male belugas reach maturity at nine years old or older (Robeck et al., 2005; Brodie et al., 2013). If this assumption is true, belugas develop stereotyped individual calls later than bottlenose dolphins, which develop them during the first year of life (Sayigh et al., 1990).

Duration lacked individual distinctiveness. Univariate statistical analyses revealed that DUR did not differ significantly among individuals although other temporal parameters were significantly different (Table 2.3). The PIC measure of DUR also indicated that inter-individual variability was less than intra-individual variability. Similarly, duration was not consistent in signature whistles. In the signature whistles of bottlenose dolphins, each of the repeated basic contours is called a loop. The number of loops, loop duration, and the inter-

loop interval, which are related to whistle duration, were affected by motivational state (Esch et al., 2009). Each signature whistle type of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) also had a high degree of variation in duration (Gridley et al., 2014). Therefore, duration does not seem to transmit individuality in either pulse-type or whistle-type contact calls.

Each beluga had similar power spectra among the three pulse locations; the third pulse, the middle pulse, and the third from the last (Table 2.4). This indicates that the frequency properties of pulses were relatively fixed within PS1 regardless of pulse location. Thus, belugas may not use change in spectral characteristics within PS1 for individual identification.

Visual inspection did not find apparent individual differences in power spectra of the middle pulses, especially below 10 kHz (Fig. 2.12). Information in the frequency range from 10 kHz to 110 kHz may be effectively utilized by captive belugas, since they had high hearing sensitivity in that range (Awbrey et al., 1988; Klishin et al., 2000; Mooney et al., 2008). Univariate statistical analyses revealed that F_p, F_l, and F_u in the range were significantly different, and all spectral parameters, F_p, 10 BW, F_l and F_u, had PIC values greater than 1 (Table 2.5). However, visual comparison of power spectra in the frequency band did not elucidate apparent individual distinctiveness and found variability in intra-individual spectra (Fig. 2.12). The results of the visual investigations may be counter-intuitive, as human observers have proven to perform better than computers at classifying vocalizations (Janik, 1999). Directivity may have influenced the intra-individual variability, since the angle between the belugas' head and the hydrophones was not considered in calculating spectra. In the case of broadband echolocation clicks, the beam pattern was highly directional to reduce clutters, and clicks produced by belugas had high directivity with a -3 dB beam width of 6.5° in both the horizontal and vertical planes (Au et al., 1987). The sound field of clicks varied in accordance with frequency (Starkhammar et al., 2011). Further, dolphins are able to steer

beams of clicks (Moore et al., 2008). Although the directivity of PS1 calls was unknown, it appears to be lower than that of echolocation clicks used to broadcast the caller's information or message. In practice, power spectra tended to be similar between PS1 calls recorded on the right and left hydrophones in the present study. However, the possibility remains that directivity may account for some of the spectral variations. A wide frequency band was missing for most power spectra from the subadult NF3 and all of the power spectra from the calf NM2 (Fig. 2.12). This could be due to the fact that NF3 and NM2 tended to produce PS1 calls at lower amplitude than the adults.

The results of the quadratic DFA supported individuality in PS1 calls. It classified PS1 calls into individuals with an overall correct classification rate of 80.5%. The stepwise DFA revealed that the most powerful discriminator was IPI 1 followed by F_p. It was unclear at this stage whether the temporal and spectral variables were associated with signatures or were just by-products of voice cue attributed to differences in body size or sex. However, the pulse repetition pattern had a high potential for being a signature, if belugas encoded individual information independent of voice feature in contact calls. It was because IPI contours were stereotyped intra-individually and different inter-individually (Fig. 2.11), PRR had the highest PIC score (Table 2.3), and IPI 1 was the most informative variable in the DFA. In contrast, clear individuality was not found in power spectral shapes, especially in frequencies lower than 10 kHz where sounds effectively propagate in the whales' environment (Fig 2.12). Although spectral parameters calculated in the frequency range above 10 kHz had a PIC value greater than 1 (Table 2.5), and F_p was the second most informative parameter in the DFA, the high frequency components are inappropriate as the carrier for the signature since they are affected by transmission loss and unstable in their environment.

To summarize, this chapter demonstrated that PS1 calls served a function in group cohesion and individuality existed in various temporal and spectral parameters, while IPI contours

might encode strongly recognizable individual information akin to the signature whistles of bottlenose dolphins.

Chapter 3: Verification of the function and individuality in PS1 calls of belugas

3.1 Introduction

As described in chapter 2, one captive population use individually distinctive PS1 as a contact call (Morisaka et al., 2013; Mishima et al., 2015). Here, vocalizations were collected from another captive beluga population to examine whether the PS1 function and individuality suggested in chapter 2 were common features in beluga whales. Because the belugas could not be put unexpectedly in an isolation context, the role of the PS1 call in group cohesion could not be directly confirmed. However, if the subjects use PS1 as a contact call, PS1 should be produced in affiliative contexts and used for vocal exchange (Masataka & Biben, 1987; Sugiura, 1993; Kureta, 2000; Miller et al., 2004; Miller & Wang, 2006; Yosida et al., 2007; Carter et al., 2008; Schulz et al., 2008; Kondo et al., 2010; Nakahara & Miyazaki, 2011; Morisaka et al., 2013). Therefore, the affiliative function of PS1 was explored by assessing the relationship between PS1 bouts and directed behavior. Frequency distribution of inter-PS1 intervals was constructed to reveal whether PS1 was used for vocal exchange. Individual distinctiveness in PS1 was also analyzed as described in chapter 2. Further, it was examined whether there was increase of PS1 production rate in a visual reunion after a long separation that could imply the possibility of PS1 calls functioning in individual advertisement.

3.2 Materials and methods

3.2.1 Facility and subjects

Data were collected from the belugas kept at Shimane Aquarium, Shimane, Japan from October 2014 to March 2015. There were seven belugas, three female adults (SF1, SF2, and SF3); two male adults (SM1 and SM2); one male subadult (SM3); and one female calf (SF4). Figure 3.1 represents the genealogy of the beluga group. All adult belugas were captured at

the Amur River, Russia and estimated to be 16–18 years old. SF1, SF2, and SM2 came to the aquarium in 1999, and SF3 and SM1 in 2003. SM3 and SF4 were siblings born in captivity and their parents are SF1 and SM2. SM3 was five years old and SF4 was two months old. All ages given are correct as of October 2014.

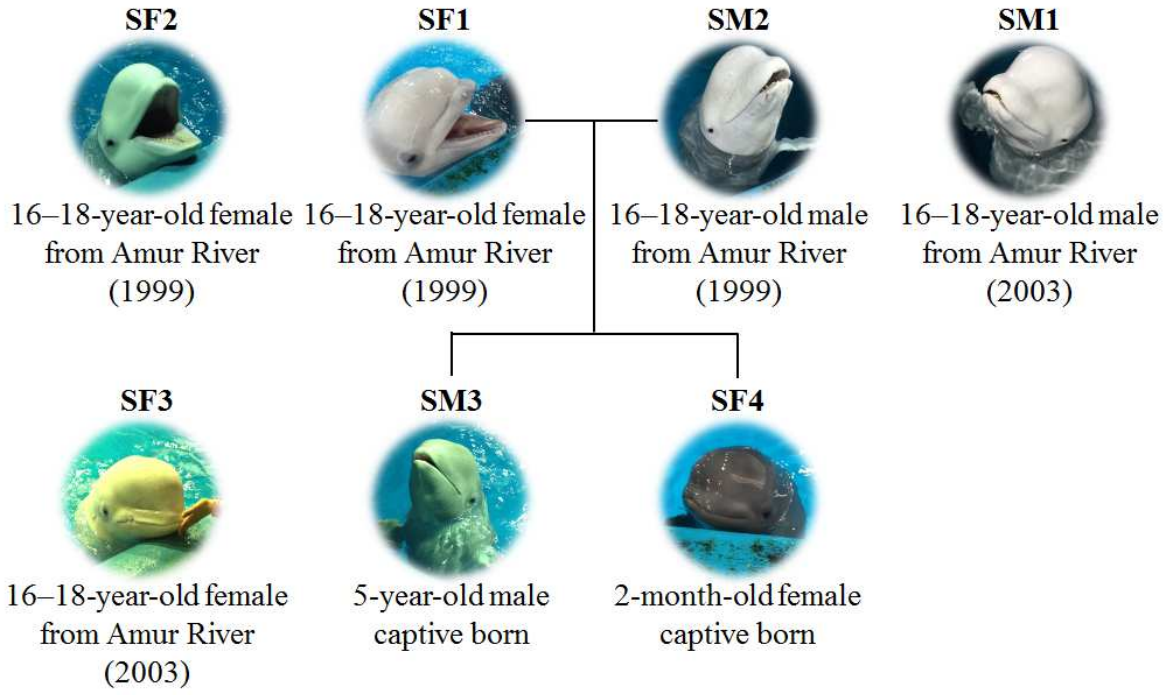


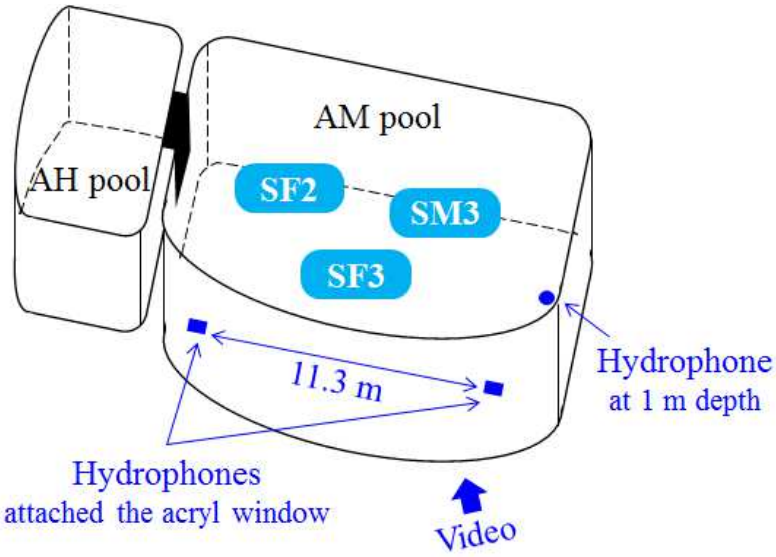
Fig. 3.1 Genealogy of the belugas kept at Shimane Aquarium as of October 2014.

There were two beluga pools, A and B. Pool A was composed of two sub-pools, a main pool (AM pool) with a depth of 5 m, and a holding pool (AH pool) with a depth of 4 m (Fig. 3.2 a). Three belugas, SF2, SF3, and SM3 were kept in the AM pool together while the AH pool held no animals.

Pool B was composed of three sub-pools, a main pool (BM pool) with a depth of 4.5 m, and two holding pools (BH1 and BH2 pools) each with a depth of 4 m (Fig. 3.2 b). Pool B contained four belugas, SF1, SF4, SM1, and SM2. The mother–calf pair (SF1 and SF4) was held in the BM pool. Each holding pool contained either SM1 or SM2. There was a metal

lattice between BH1 and BH2 pools, blocking beluga movement between pools but maintaining visual and acoustical contact between individuals SM1 and SM2. On the other hand, there were gates between BH1 and BM pools, and between BH2 and BM pools. However, these gates prevented visual contact, and although sounds could pass through the gates, they were attenuated. Therefore, communication between the males and the mother–calf pair depended on acoustic contact.

a) Pool A



b) Pool B

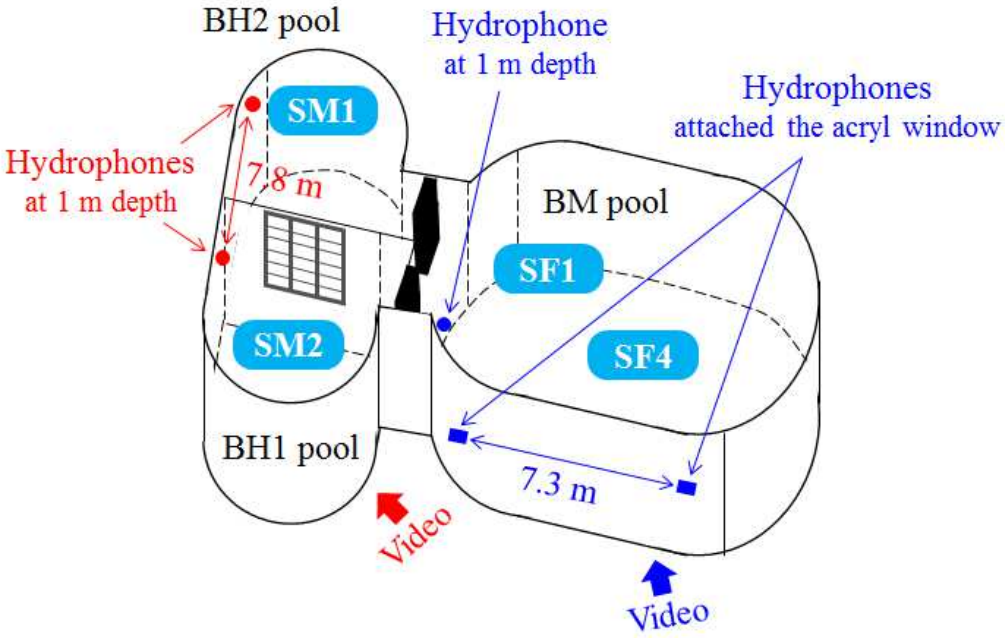


Fig. 3.2 Schematic view of the beluga pools in Shimane Aquarium. Pool A a) and Pool B b), the systems represented in blue and red were used for BM and BH recordings, respectively.

3.2.2 Data collection

There were three recording patterns for normal sessions: “AM recordings” for SF2, SF3, and SM3 in AM pool; “BM recordings” for SF1 and SF4 in BM pool; and “BH recordings” for SM1 and SM2 in BH pools. AM recordings were conducted in a normal free swimming context. BM and BH recordings were made in the normal separated context: SM1 and SM2 were isolated in each BH pool, and SF1 and SF4 were segregated from the males together in BM pool. Each recording continued for 30 min, and a total of 28 normal sessions over 14 h were performed (Table 3.1).

To increase the number of PS1 samples in SM2, an additional 30-min recording session was conducted when there were high vocal activity levels (Table 3.1).

Further, a special session of “gate-open recording” was carried out in pool B (Table 3.1). In this special session, the gates between BM pool and BH pools were opened, with the metal lattice remaining in place, meaning individuals could not move between the sub-pools. This was a visual reunion after a long separation for the males and SF1, and a first opportunity for the males to see the calf SF4.

All sessions were performed between 09:00 and 17:00, excluding feeding and training times, and except during gate-opening, nobody entered the pool sides during recording.

Table 3.1 Number of sessions and total recording time for each recording pattern.

Session name	Normal			Additional	Special
Recording pool	AM	BM	BH	BH	B
Housed individual	SF2 SF3 SM3	SF1 SF4	SM1 SM2	SM1 SM2	SM1 SM2 SF1 SF4
Total recording time (min) (Number of sessions)	360 (12)	300 (10)	180 (6)	30 (1)	36 (1)

AM and BM recordings used three hydrophones (Fig. 3.2). One was a TC 4013 underwater hydrophone (Reson Inc., Denmark) which exhibits a flat frequency response from 1 Hz to 170 kHz (-211 ± 3 dB re 1 μPa/V at 1 m), and was housed in PVC pipes and placed at a depth of 1 m. The others were AQH-100DTP touch panel hydrophones (AquaSound Inc., Japan), which exhibit a flat frequency response from 20 Hz to 100 kHz (above -212 dB re 1 μPa/V at 1 m). The elements of the touch panel hydrophones were covered by acryl resin and attached to the acryl observation window by using grease. The hydrophone can efficiently collect sounds generated underwater behind acryl windows because the acoustic impedance of acryl materials is near to that of the acoustic matching layer between seawater and the element of the hydrophone. The two touch panel hydrophones were attached at 2 m in depth and spaced 11.3 m and 7.3 m apart from each other in the AM and BM pools, respectively. The underwater hydrophone was used for analysis of acoustic parameters and the two touch panel hydrophones were used to identify callers by measuring time difference of sound arrival. BH recordings used two underwater hydrophones submerged at a depth of 1 m and separated by 7.8 m from each other (Fig. 3.2 b). They were used for the analyses of acoustic parameters and identification of callers. During the special session, vocalizations were recorded using all hydrophones used in BM and BH recordings.

The sound from underwater hydrophones was analog high-pass filtered at 1 kHz, and amplified by 32 dB using VP1000 preamplifiers (Reson Inc., Denmark), with a flat frequency response to 1 MHz (-3 dB). The sound from in-air hydrophones was analog band-pass filtered from 1 kHz to 200 kHz, and amplified by 50 dB using Aquafeeler III preamplifiers (AquaSound Inc., Japan), with a flat frequency response to 200 kHz (-3 dB). All analog data were collected by EZ7510 data recorders (NF Co., Japan), which digitized up to two channels of sound sampling at 500 kHz and 16 bits. Observations were made from an underwater window using a GZ-V675-R video camera (JVC Co., Japan).

3.2.3 Call classification and caller identification

Calls were classified using the same five categories as described in subsection 2.2.3, composed of one type of pulsed call “PS1,” one type of combined call “C1,” short calls “S,” whistles “W,” and others “O.” Although belugas in AM pool frequently produced PS1 calls in air as well as underwater, all in-air production instances were excluded from the data.

PS1 callers were identified using arrival time differences at the two separated hydrophones. The time differences in waveforms were inspected visually or by custom written MATLAB software as described in subsection 2.2.4. The video and touch panel hydrophone arrangement in AM pool enabled for discrimination between PS1 callers. Likewise, it was possible to discriminate the PS1 calls from SF1 and SF4 in BM recordings. The hydrophones installed in BM pool also recorded vocalizations from the males held in BH pools although the vocalizations were of poor acoustic quality. The male calls were counted but those PS1 callers could not be identified. In BH recordings, PS1 calls from SM1 and SM2 could be discriminated through the video and underwater hydrophone arrangement. BH recordings also collected vocalizations from the mother–calf pair in BM pool. PS1 calls with poor acoustic quality were considered to be produced by the mother SF1 since no PS1 calls were heard from the calf SF4 during the experimental period. In any cases where there was ambiguity in the identification of PS1 callers, the calls were counted but excluded from the caller-identified category.

3.3 PS1 analyses

3.3.1 Bout criteria interval

PS1 production appears to occur in bouts. Investigation of the relationship between PS1 bouts and directed behavior tells us whether PS1 is an affiliative signal. Therefore, it is necessary to define a PS1 bout, and thus bout criteria interval (BCI) was calculated using the two-process

exponential model (Sibly et al., 1990).

In the two-process model, interval events are generated by one of two random processes, a fast process operating within bouts, and a slow process generating new bouts. The two-process model was expressed as

$$y = \log_e\left(N_f\lambda_f e^{-\lambda_f t} + N_s\lambda_s e^{-\lambda_s t}\right) = -\lambda_f t + \log_e N_f\lambda_f - \lambda_s t + \log_e N_s\lambda_s \qquad (3\text{-}1)$$

N_f is the total number of intervals within bouts and N_s is the total number of bouts. λ_f is probability per unit time that interval events will occur within a bout, and λ_s is probability per unit time that a new bout will begin. $N_f\lambda_f e^{-\lambda_f t}$ and $N_s\lambda_s e^{-\lambda_s t}$ represent frequency of the interval length between t and t+1 for fast process and for slow process, respectively. When the two-process model is fitted, the regression of the $\log_e$ frequency distribution follows two straight lines, a steep line with slope of $-\lambda_f$ and y-intercept of $\log_e N_f\lambda_f$ and a gradual line with slope of $-\lambda_s$ and $\log_e N_s\lambda_s$. In this case, PS1 production is split into bouts.

As a first step for fitting the model, all intervals between successive PS1 calls throughout each 30-min session were measured. The widely used criterion of inter-call interval is the latency period from the end of the preceding call to the beginning of the following call. However, the data contained many overlapping exchanges, and it was easier to identify the beginning than the end of those overlapping PS1 calls. Therefore, inter-PS1 interval in this study was defined as the duration from the peak of the first pulse in the preceding PS1 to the peak of the first pulse in the following PS1 (Fig. 3.3). In cases where intervals could not be measured correctly and in-air PS1 production was inserted in underwater PS1 sequences, the samples were excluded from the interval data.

The $\log_e$ frequency distribution of the inter-PS1 intervals was constructed. The bin width can be chosen in any convenient way; therefore, it was determined by the Freedman–Diaconis rule.

Two regression lines were fitted by eye, and the values of slopes and y-intercepts expressed in the regression lines were used in calculating the initial parameter estimates, λ_{f0}, λ_{s0}, N_{f0}, and N_{s0} for a non-linear curve-fitting procedure (NLIN) of equation (3-1). The output parameter estimates from NLIN, λ_f, λ_s, N_f, and N_s were used for calculation of BCI.

There are two methods to calculate BCI: one minimizes the total time misassigned (BCI1, Fagen & Young, 1978) and the other minimizes the total number of events misassigned (BCI2, Slater & Lester, 1982).

$$\mathrm{BCI1} = \frac{1}{\lambda_f - \lambda_s} log_e \frac{N_f}{N_s} \qquad (3\text{-}2)$$

$$\mathrm{BCI2} = \frac{1}{\lambda_f - \lambda_s} log_e \frac{N_f \lambda_f}{N_s \lambda_s} \qquad (3\text{-}3)$$

The number of misassigned events with BCI = t are calculated as

$$N_f e^{-\lambda_f t} + N_s(1 - e^{-\lambda_s t}) \qquad (3\text{-}4)$$

Using equation (3-4) (Slater & Lester, 1982), I evaluated the performance of BCI1 and BCI2, and chose the BCI that had the smaller number of misassigned events. PS1 bouts were extracted based on the selected BCI. PS1 bout duration was defined as the time length from the onset of the first PS1 to the offset of the last PS1 in a PS1 bout. All statistical analyses were performed using R software version 3.1.0. (The R Foundation for Statistical Computing).

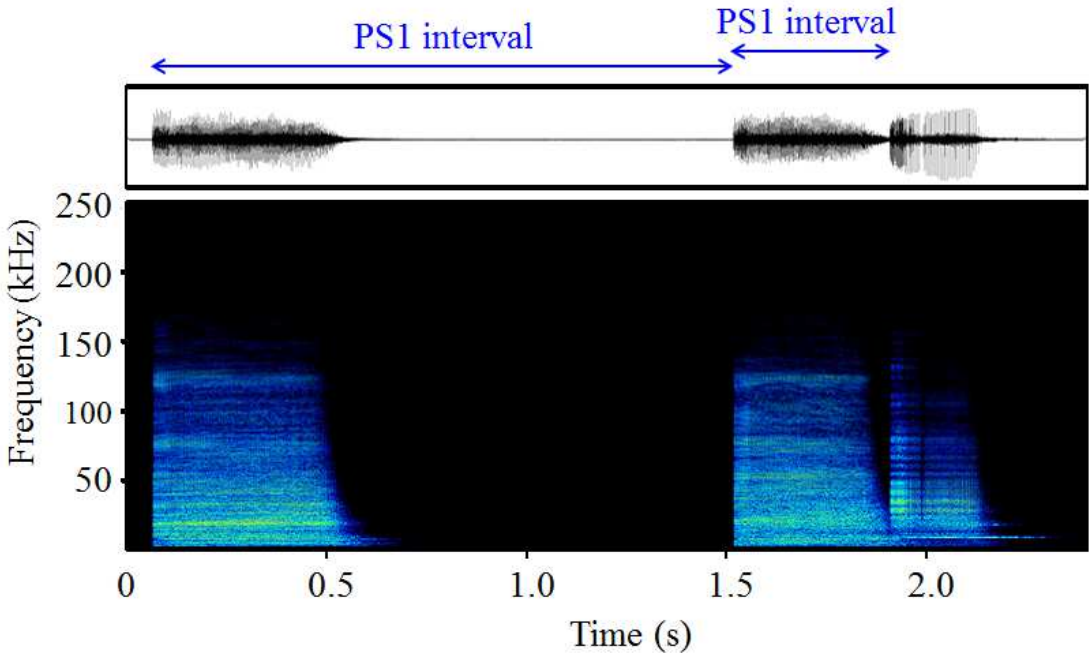


Fig. 3.3 Interval measurements of PS1 sequences.

3.3.2 Relation to directed behavior

The relation between PS1 bouts and directed behavior was analyzed using the AM data to evaluate the affiliative function of PS1. The belugas there were all in the same space and contained in the video frame, which enabled for the observation of interactive behavior. In the free swimming context, the salient directed behaviors were aggression and submission, and other directed behaviors such as sexual interaction were not observed. Thus, if PS1 bouts were not involved with aggressive/submissive behavior, PS1 was regarded as an affiliative call.

Recchia (1994) defined the aggressive behavior of captive belugas as hit, bite, bite threat, closed-mouth bite threat, charge, slow charge, chase, jaw clap, mouth open, directed look, head jerk, face to face, stare, and melon extension. In addition, Hill (2009) proposed bubble bursts directed towards an animal as an aggressive action. Submissive behavior was also defined, such as flee, close flee, flinch, look away, roll away, avoid, and lie passive (Recchia, 1994). Aggressive/submissive behaviors were investigated based on the above criteria. The

behavior often occurred in successive combinations; therefore, a chain of aggressive/submissive behavior was regarded as one aggressive/submissive event. Each event was clearly independent. The relation between the PS1 bouts and the aggressive/submissive events were examined.

3.3.3 Exchange pattern

Vocal exchange is characterized as a call sequence when the preceding call is followed by the call-back of another individual within a particular temporal window. Calls produced within the temporal window are considered as responses. Callers wait for replies from other individuals during the temporal window, and they repeat a call unless they hear replies. In this manner, a temporal rule to regulate vocal exchange exists in contact calls of several species, including birds such as large-billed crows (*Corvus macrorhynchos*) (Kondo et al., 2010), terrestrial mammals such as squirrel monkeys (*Saimiri sciureus*) (Masataka & Biben, 1987), Japanese macaques (*Macaca fuscata*) (Sugiura, 1993), cotton-top tamarins (*Saguinus oedipus*) (Kureta, 2000), common marmosets (*Callithrix jacchus*) (Miller & Wang, 2006), naked mole-rat (*Heterocephalus glaber*) (Yosida et al., 2007), white-winged vampire bats (*Diaemus youngi*) (Carter et al., 2008), and marine mammals such as killer whales (Miller et al., 2004), sperm whales (Schulz et al., 2008), bottlenose dolphins (Nakahara & Miyazaki, 2011), and belugas kept at the Port of Nagoya Public Aquarium (Morisaka et al., 2013).

To ascertain the prediction that PS1 calls produced by belugas at Shimane Aquarium were also used for vocal exchange, the intervals between two consecutive PS1 calls by different belugas (between-individual interval: BII) and those by single belugas (within-individual interval: WII) were investigated. BII and WII frequency distributions were constructed for each of the AM and BH recordings, and a temporal pattern was searched. BM recordings were excluded in this analysis because in BM recordings, PS1 callers of BH pools could not be

identified and it was difficult to classify most intervals as either BII or WII.

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 just an incidental result. In the simulation, imaginary individuals produced PS1 calls independently at their own pace. Therefore, simulated BIIs were intervals between independent PS1 calls by different individuals. Simulated BII data were generated using the observed WII data. 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. The order of the selected WII samples was randomized, and then the re-ordered samples were accumulated from the first. The accumulated values were regarded as onset times of PS1 calls produced by an imaginary individual. The manipulation was repeated three times for three imaginary individuals, which was the same as the number of individuals engaged in PS1 production in each of the AM and BH pools. The three 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 sequences by different imaginary individuals. This procedure was repeated to obtain 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.

3.3.4 Individual comparison

Acoustic properties of the individual PS1 calls were analyzed and compared among individuals as described in section 2.3. Temporal parameters were the number of pulses (N_p); duration (DUR); pulse repetition rate (PRR); average inter-pulse interval of pulse nos. 11–20 (IPI 1); and average inter-pulse interval of pulse nos. 11–20 from the last (IPI 2) (Table 2.2). IPI contours were also depicted. Noise spectra were calculated using non-call windows before

the onset of the PS1 calls, and consequently there was no distinct noise above 1 kHz in contrast to the Port of Nagoya Public Aquarium. Thus, relative spectra were compared in the frequency range above 1 kHz. Four spectral parameters were calculated from the spectra: peak frequency of the middle pulse (F_p); 10 dB bandwidth (10 BW), which is the frequency band at a level of -10 dB from the peak; and the lower and upper frequency of the 10 BW (F_l and F_u) (Table 2.2). The acoustic parameters were compared inter-individually by using Kruskal–Wallis test or one-way ANOVA. PIC was also used to compare the dynamic of intra- and inter-individual differences. Quadratic DFA was performed after examination for multicollinearity, multivariate outliers, multivariate normality, and homogeneity of variance-covariance matrices on the data set of the nine variables. The most predominant discriminator was found by stepwise DFA.

3.3.5 Role of individual identity advertisement

Quick & Janik (2012) reported that bottlenose dolphins exchanged signature whistles in contexts in which they encountered and joined conspecifics at sea in order to identify each other. The special session in pool B was conducted to explore whether a similar function existed in PS1. The adult belugas in pool B are already known to each other and acoustic communication has been possible through the closed gates. However, if belugas use PS1 calls for identity advertisement, PS1 call rate may increase by the gate open where both acoustical and visual contact are fully available to verify individual identification. It took 12 min for trainers to enter the pool side, open all of the gates, and leave the pool side. Recordings were thus classified three parts, for 12 min before the gate open, during the 12-min gate open, and for 12 min after the gate open. Changes in PS1 call rate were examined.

3.3.6 Context-dependent changes of acoustic parameters

The nine acoustic parameters of PS1 calls (Table 2.2) were compared between normal sessions and the 12-min gate opening in the special session to explore whether there were differences in the acoustic parameters depending on context. The differences were tested by using Wilcoxon rank-sum test, Student's *t*-test, or Welch's *t*-test, with test selection being dependent on normality and variance homogeneity.

3.4 Results

3.4.1 Frequency of each call type

In AM pool, a total of 2421 calls were collected in 12 sessions over 6 h. PS1 was the predominant call type and it accounted for 40% of total calls followed by 24% of C1, 17% of W, and 5% of S (Fig. 3.4 a). The results of BM and BH recordings were pooled and a total of 3260 calls were recorded in 16 sessions over 8 h. PS1 was again the most frequent call type and occupying 39% of total calls followed by 24% of W, 18% of C1, and 4% of S (Fig. 3.4 b).

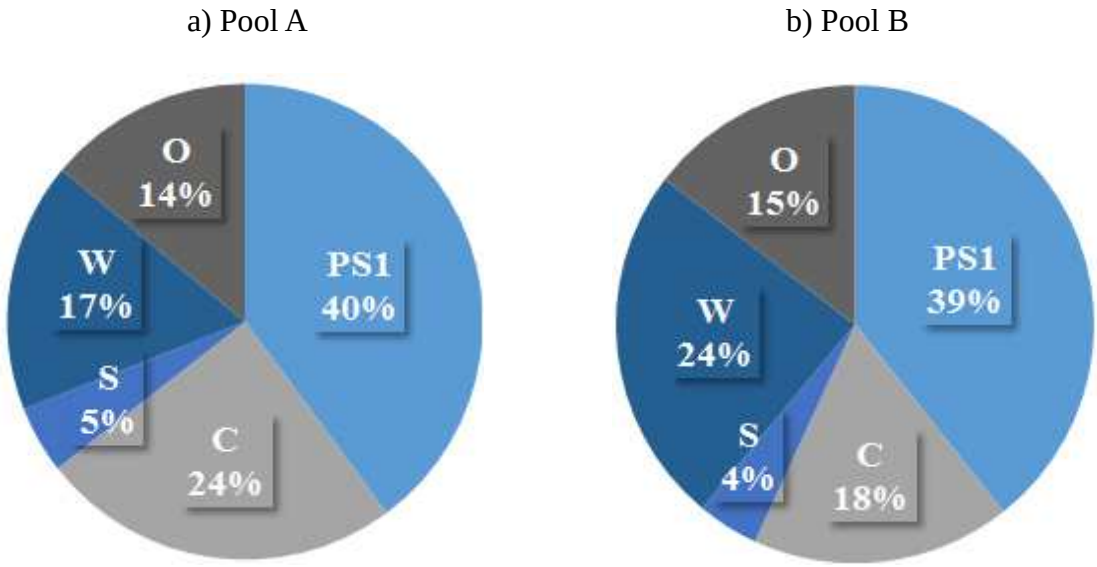


Fig. 3.4 Frequency of each call type in a) pool A and b) pool B.

3.4.2 Affiliative function of PS1

Inter-PS1 intervals in AM recordings were measured to estimate BCI. A total of 559 intervals were obtained. The $\log_e$ frequency distribution of PS1 intervals are shown in Fig. 3.5. The bin width was determined as 0.5 s. The two regression lines for two-process model were calculated as

$$y = -0.671x + 5.850 \ \ (R^2 = 0.84, P < 0.0001) \qquad (3\text{-}5)$$

$$y = -0.067x + 2.213 \ \ (R^2 = 0.79, P < 0.0001) \qquad (3\text{-}6)$$

From the slopes and y-intercepts of equations (3-5) and (3-6), initial parameters of the formula (3-1) were calculated as

$$\lambda_{f0} = 0.6711, \qquad N_{f0} = 517.6, \qquad \lambda_{s0} = 0.0666, \qquad N_{s0} = 137.3$$

NLIN provide estimates of the parameters:

$$\lambda_f = 0.7059, \qquad N_f = 464.3, \qquad \lambda_s = 0.0632, \qquad N_s = 124.0$$

The distribution of PS1 intervals was fitted to the two-process model generated by using the above estimated parameters (Fig. 3.5). To calculate BCI, the final parameter values were applied to the formulas (3-2) and (3-3), and consequently values of BCI1 = 2.1 s and BCI2 = 5.8 s were obtained. Since the number of points misassigned was 124.0 for BCI1 and 45.8 for BCI2, BCI2 was selected (Fig. 3.6). There were a total of 202 PS1 bouts in 12 sessions (16.8 ± 8.3 per session), and the average PS1 bout duration was 5.3 ± 4.6 s.

The most commonly observed aggressive behaviors were stare, directed look, and mouth open, and the most common submissive behavior was flee (Fig. 3.7). A total of 47 aggressive/submissive events were identified in the 12 sessions of AM recordings (3.9 ± 3.8 per session), and the average duration of aggressive/submissive events was 3.6 ± 2.2 s.

Only one out of the 47 aggressive/submissive events overlapped with a PS1 bout. This PS1 bout was the second longest recorded (27.1 s) and took place in exchanges between the individuals SF2 and SF3. On the other hand, the aggressive/submissive event within the PS1

bout occurred between SF3 and SM3 and it ceased in a short period (2.3 s). Therefore, it was concluded that there was no relationship between PS1 bouts and aggressive/submissive events.

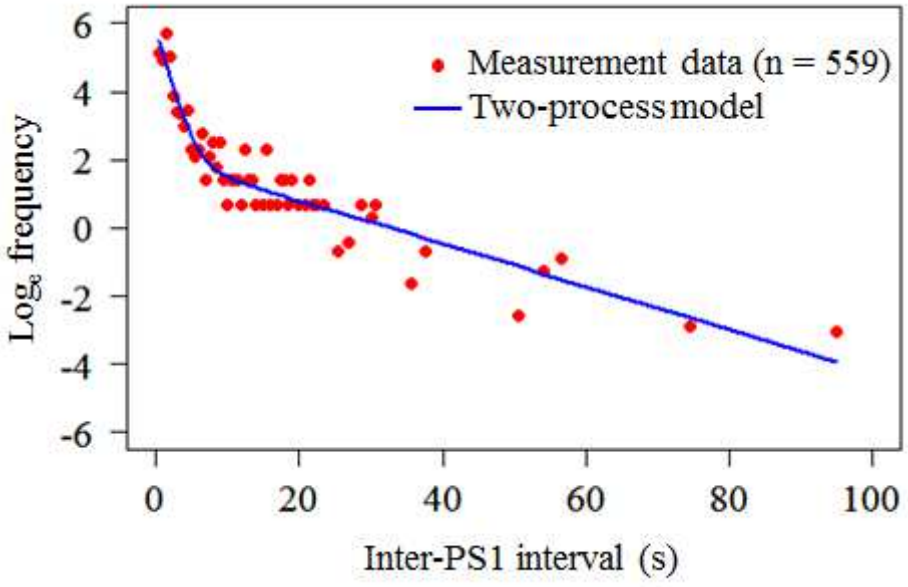


Fig. 3.5 Fitting of the log$_e$ frequency distribution of inter-PS1 intervals to the two-process model.

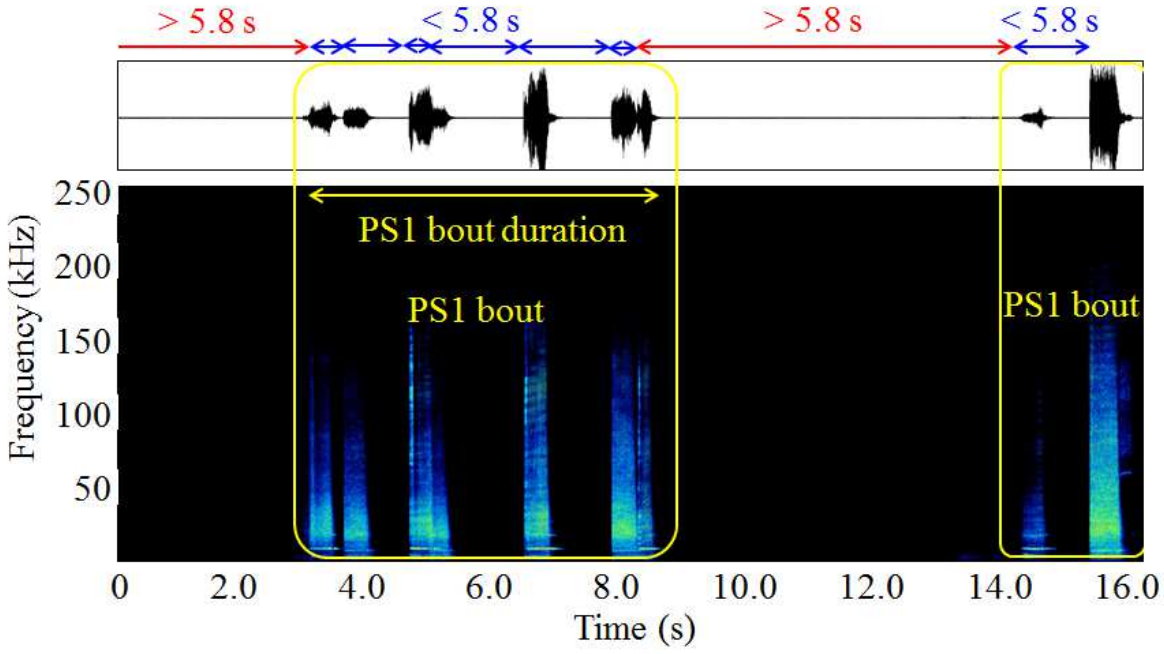


Fig. 3.6 Definition of PS1 bout.

a) Directed look and flee b) Mouth open and flee

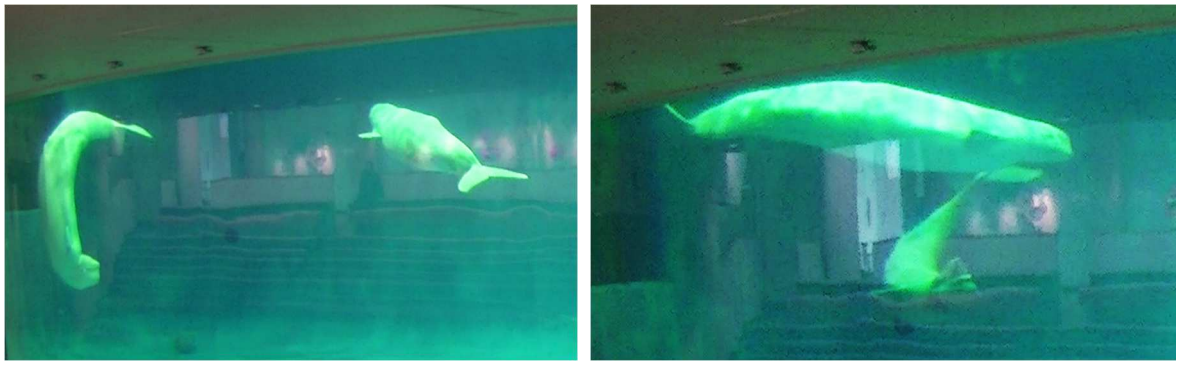


Fig. 3.7 Examples of the observed aggressive and submissive behaviors of belugas, a) directed look and flee and b) mouth open and flee.

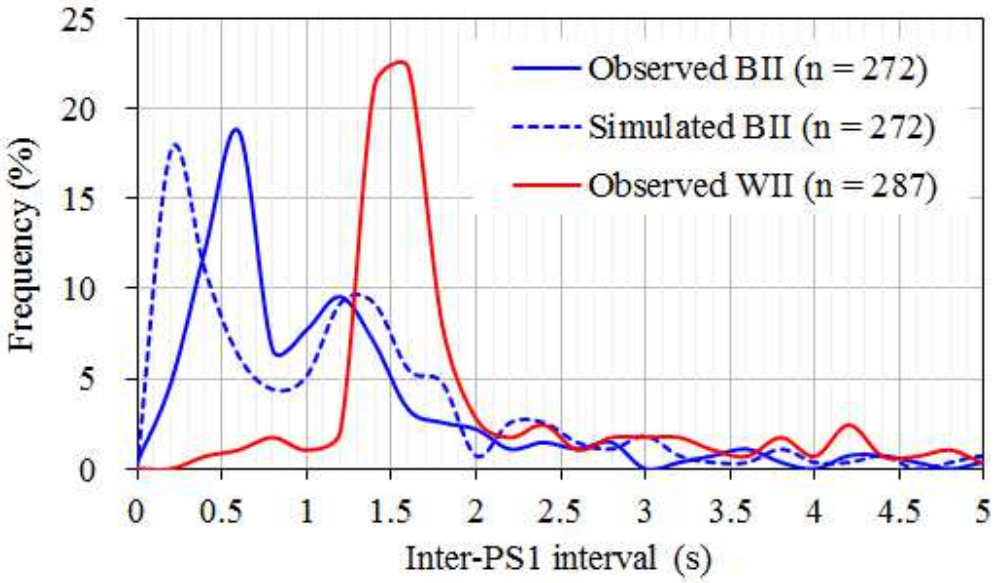
3.4.3 Temporal rule of PS1 exchange

Measured inter-PS1 intervals were divided into BII and WII. A total of 272 BIIs and 287 WIIs were collected from the AM recordings. In the BH recordings, a total of 237 BIIs and 170 WIIs could be measured. Figure 3.8 represents the frequency distribution of BII and WII. In AM recordings, the WII distribution had a sharp peak at an interval of 1.6 s, while the BII distribution had two peaks, with a sharp peak at 0.6 s and a small peak at 1.2 s. Further, 95.5% of the WIIs occurred after 1 s, whereas 50.4% of the BIIs occurred within 1 s. It was found that 20.2% of the BIIs were intervals of overlapping PS1 sequences. In BH recordings, the WII distribution showed two small peaks at intervals of 1.8 s and 3.2 s, while BII distribution had a sharp peak at 0.4 s. All of the WIIs occurred after 1 s, whereas 51.1% of the BIIs occurred within 1 s, and 47.3% of the BIIs were intervals of overlapping PS1 sequences.

For an imaginary individual of BII simulation, I selected 8 and 10 samples from the WII data pool of AM and BH recordings, respectively. This manipulation was repeated three times to generate an imaginary group composed of three individuals. Then, the simulated BIIs were computed by measuring the intervals between consecutive PS1 calls by different imaginary individuals. This process was repeated 24 times to collect 272 samples for AM data. Similarly, it was repeated 21 times to collect 243 samples for BH data.

The simulated BII distribution of AM recordings had two peaks, with a sharp peak at 0.2 s and a small peak at 1.2–1.4 s (Fig. 3.8). 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. Two-sample Kolmogorov–Smirnov test revealed that the simulated and observed BIIs had significantly different frequency distributions in both the AM and BH recordings ($D = 0.143$, $P < 0.01$; $D = 0.447$, $P < 0.01$, respectively).

a) AM recordings



b) BH recordings

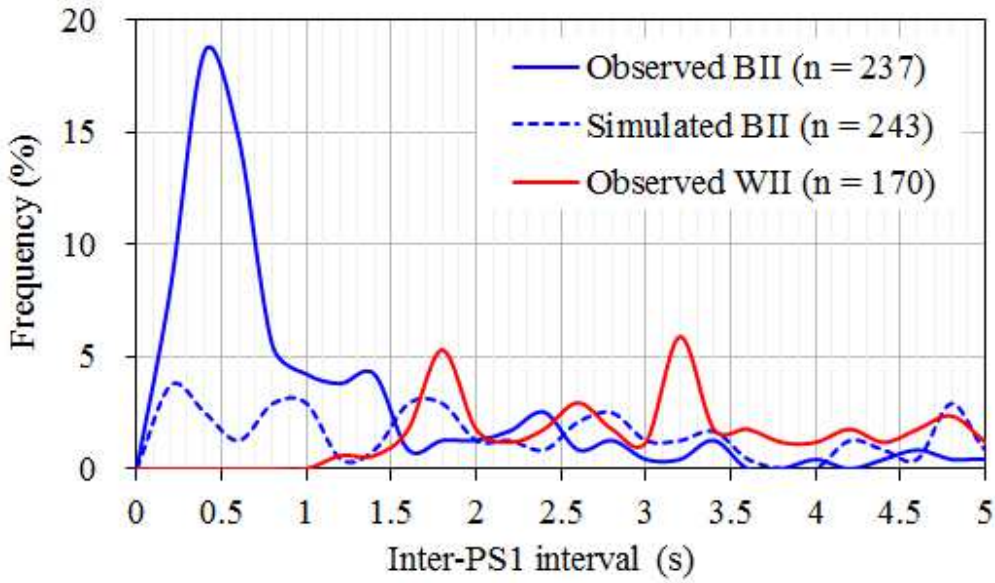


Fig. 3.8 Frequency distribution of observed BIIs, simulated BIIs, and observed WIIs for a) AM recordings and b) BH recordings.

3.4.4 Individuality in PS1

Of the 963 PS1 calls in AM recordings, a total of 867 calls were identified callers, including 489, 211, and 167 from SF2, SF3, and SM3, respectively. A total of 220 PS1 calls in BM recordings were identified as calls from SF1, and the newborn SF4 did not produce any PS1 calls. All of the 422 PS1 calls in BH recordings were identified callers, including 243, 108, and 71 from SM1, SM2, and SF1 respectively, but the SF1 samples in BH recordings are excluded from this analysis because they had poor acoustic quality. The most vocal belugas were SF2 in Pool A and SM1 in Pool B.

Characteristics of PS1 calls from each beluga were investigated spectrographically and by listening: males produced various PS1 calls, while females had only one PS1 call type. When PS1 calls produced by males were classified based on IPI contour, both SM1 and SM2 had five PS1 variants, while SM3 had three variants (Fig. 3.9). PS1-V4 of SM1 was a slightly shifted configuration of his PS-V1. Likewise, PS1-V2 of SM2 was a slightly shifted configuration of his PS1-V1. The PS1-V2 of SM3 was a repeated structure of his PS1-V1. The PS1-V4 of SM1 and PS1-V4 of SM2 resembled one another. The predominant PS1 call type of each male, PS1-V1 calls, were used for further analysis (Fig. 3.10).

Acoustic parameters were extracted from PS1 calls with good signal-to-noise ratio. The sample numbers were 26, 12, 31, 53, 100, and 21 for SM1, SM2, SM3, SF1, SF2, and SF3, respectively. The PS1 sample numbers for SM2 were increased to 30 through the additional recording.

Spectrogram examples of PS1 calls are shown in Fig. 3.11. All PS1 calls contained a tonal component or a secondary pulsed component, and this component is composed of low-frequency, narrowband pulses with different pulse repetition patterns. Those components were not modulated over the call duration. SM1 and SF1 had a secondary pulsed component and the dominant frequencies were 13.8 ± 0.1 kHz for SM1 and 11.3 ± 0.3 kHz for SF1. Others

had a tonal component and their dominant frequencies were 9.3 ± 0.0 kHz for SM2, 9.1 ± 0.1 kHz for SM3, 7.8 ± 0.1 kHz for SF2, and 9.2 ± 0.1 kHz for SF3.

Temporal characteristics of PS1 for each beluga are summarized in Tables 3.2. Kruskal–Wallis test revealed that all parameters were significantly different among individuals ($P < 0.0001$). Furthermore, all parameters had PIC > 1 and were more variable among individuals than within an individual. Pulse repetition pattern was most likely to convey individual information because PRR, IPI 1, and IPI 2 had high PIC values above 3.

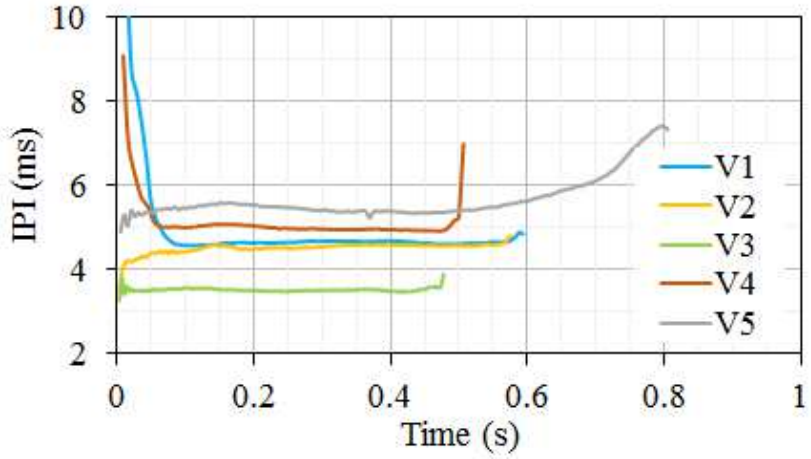
Figure 3.12 shows 21 examples of IPI contours for each individual. They were stereotyped intra-individually and differed inter-individually, and individuals were easily discriminated from each other. The IPI contours of SM1 started with a high value whereas those of others were started with a low value. There tended to be abrupt change in IPIs at the beginning of the IPI contours.

Averaged power spectra of middle pulses within PS1 were calculated and 21 samples for each individual are shown in Fig. 3.13. When they were compared visually, there was no obvious individual distinctiveness and consistency. The spectral characteristics are summarized in Table 3.3. Univariate statistical analyses demonstrated that all parameters differed significantly among individuals ($P < 0.0001$). While all spectral parameters had PIC > 1, those PIC values were lower than those of temporal parameters.

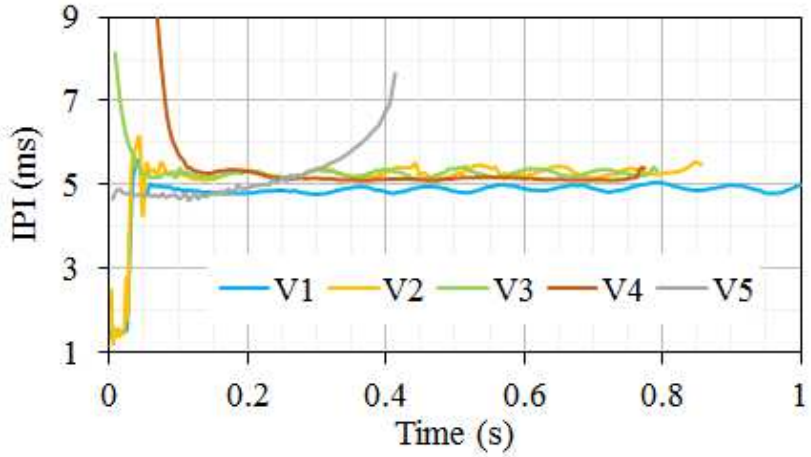
DFA was performed using the measured parameters. Multicollinearity was found in N_p, 10 BW, F_l, and F_u, and these parameters were excluded from the data set. The remaining variables, DUR, PRR, IPI 1, IPI 2, and F_p, which were the same as those in the Port of Nagoya Public Aquarium, had a VIF of less than 4. Potential multivariate outliers in the remaining data set are shown in Fig. 3.14. Since most of the PS1 calls from SM2 and all of the PS1 calls from SM3 and SF3 were regarded as outliers, all outliers were included in the DFA with the exception of three outstanding calls of SM3. To decrease disparity in sample size and increase

the effectiveness of the DFA (McGarigal et al., 2000), 53 samples were randomly selected from SF2, which was the same number used for SF1, and sample size fell into the range of 21–53. Because the data set did not satisfy multivariate normality (Shapiro–Wilk test, $P < 0.0001$) and homogeneity of variance-covariance matrices (Box's M test, $P < 0.0001$), quadratic DFA was performed. The quadratic DFA based on the five variables resulted in correct classification rates of 96.2%, 96.7%, 89.2%, 100%, 100%, and 71.4% for SM1, SM2, SM3, SF1, SF2, and SF3, respectively, with an overall correct classification rate of 94.8% (Table 3.4). The stepwise DFA revealed that the most powerful discriminator was IPI 2, followed by IPI 1.

a) SM1



b) SM2



c) SM3

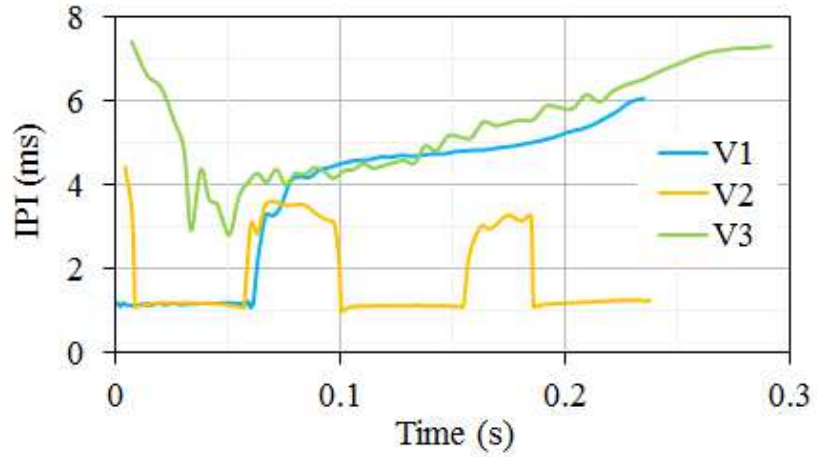


Fig. 3.9 IPI-contour examples of PS1 variants of each male, a) SM1, b) SM2, and c) SM3.

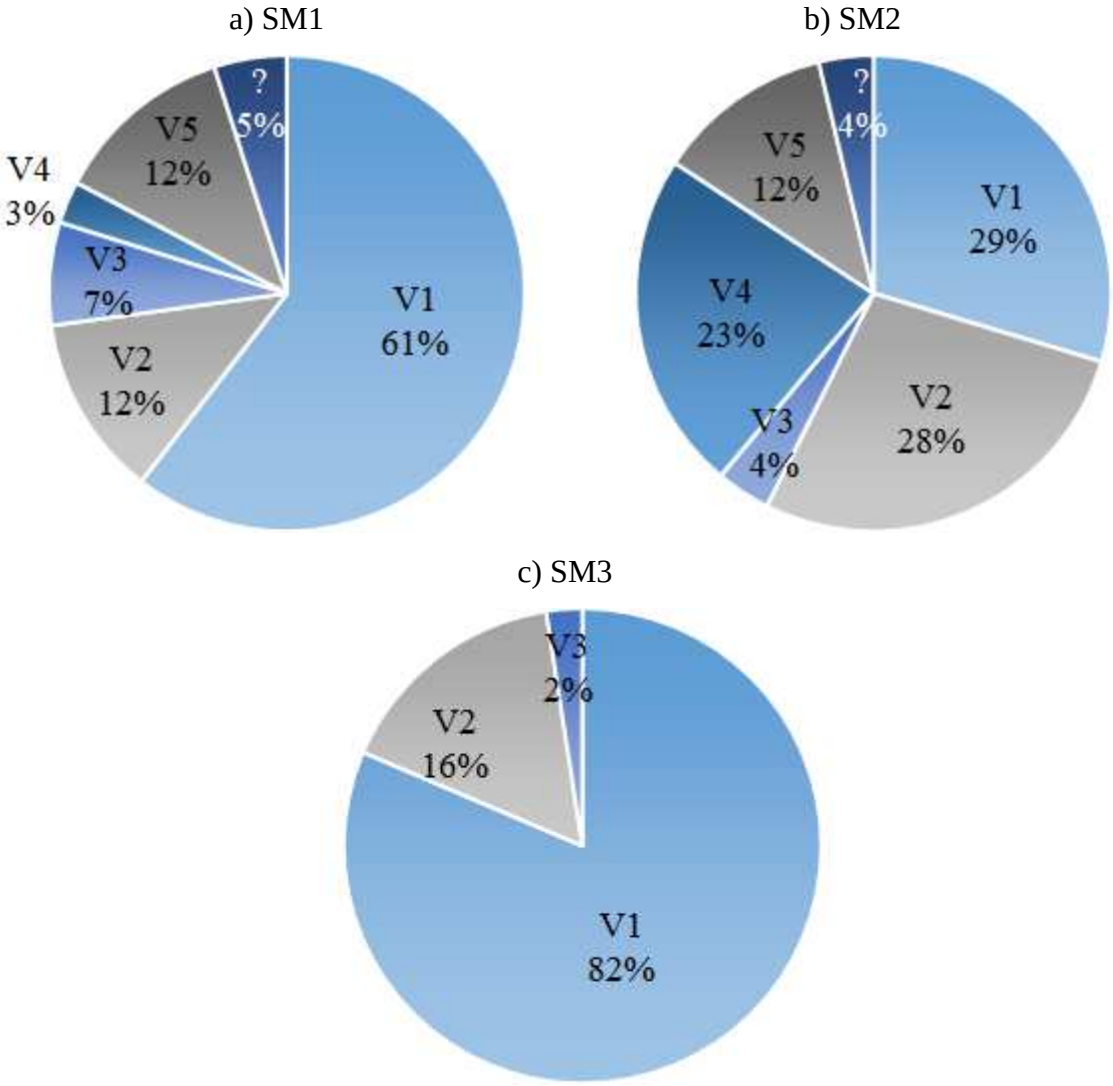
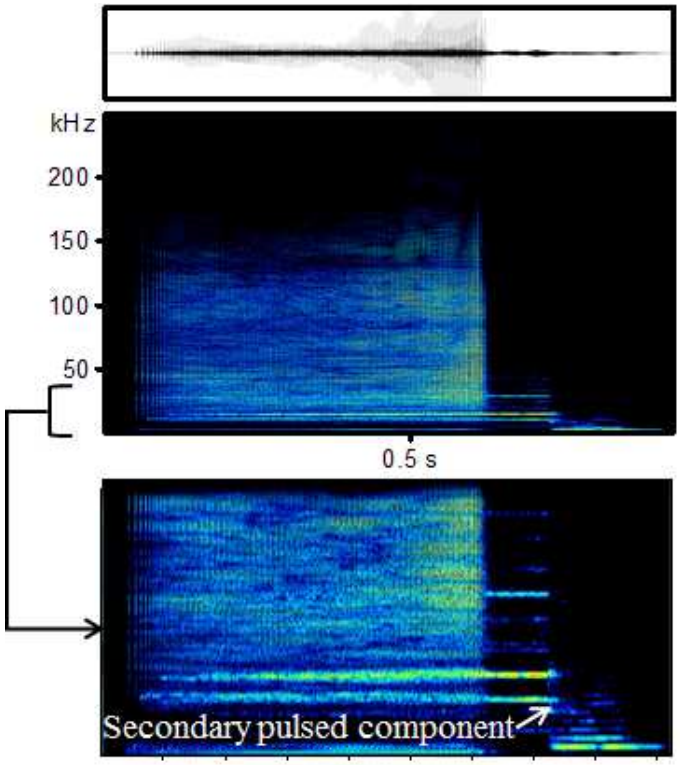
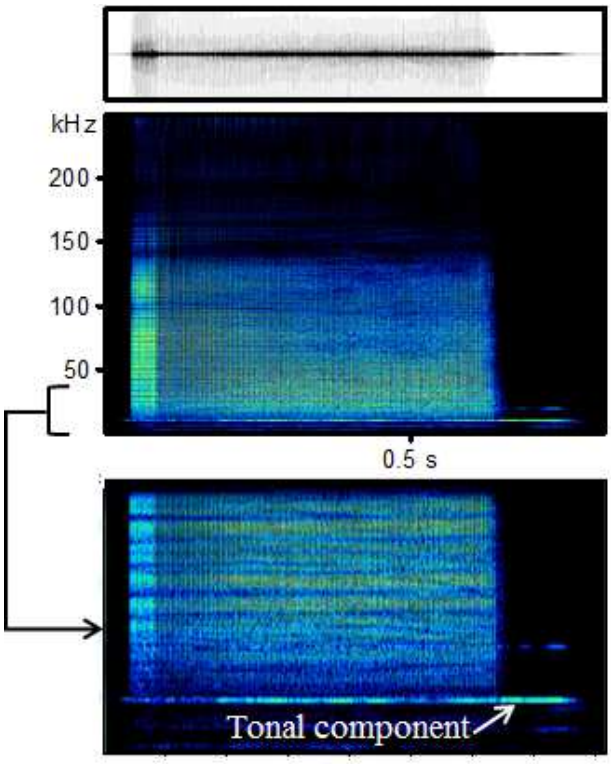


Fig. 3.10 Proportion of each PS1 variant in three males: a) SM1, b) SM2, and c) SM3.

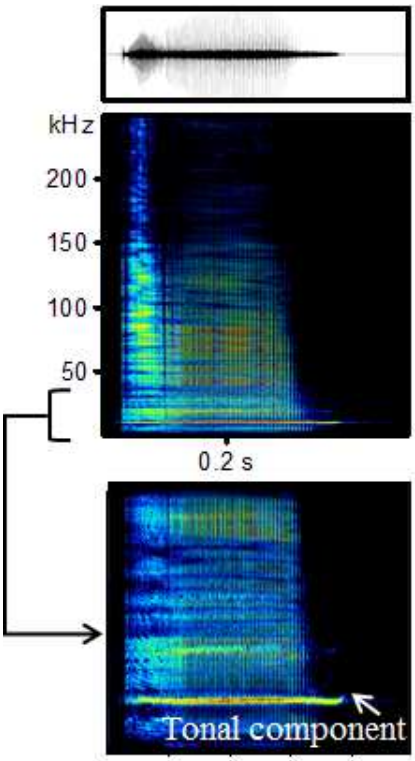
a) SM1



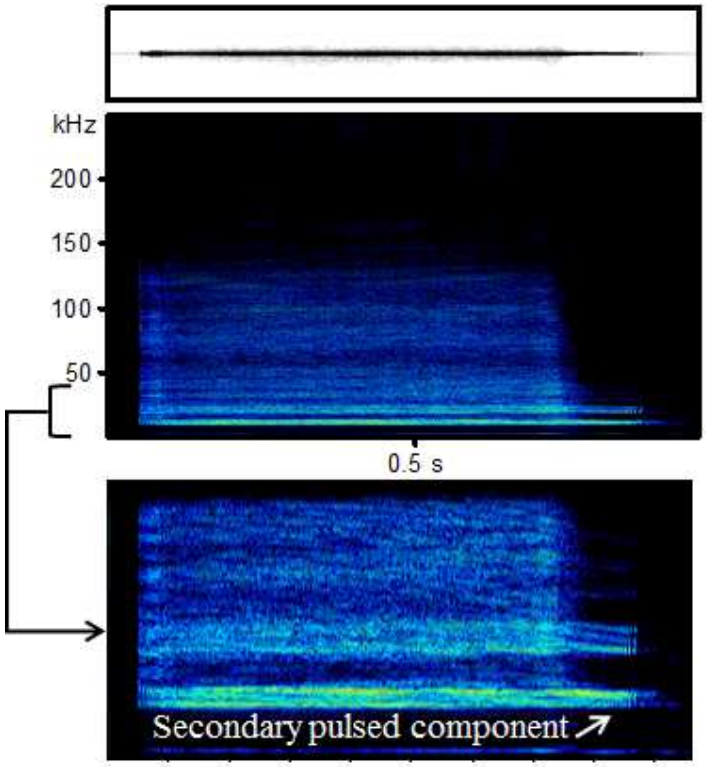
b) SM2



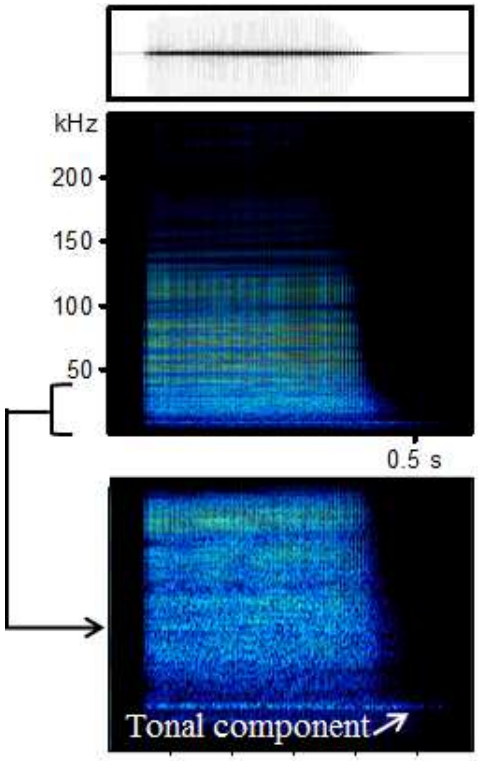
c) SM3



d) SF1



e) SF2



f) SF3

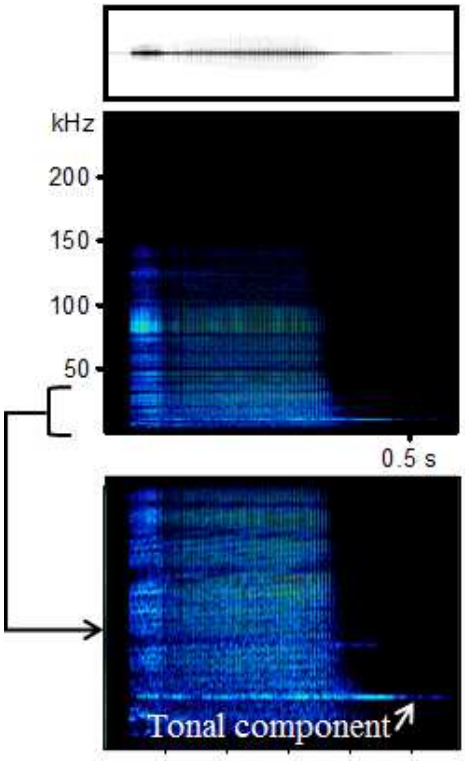


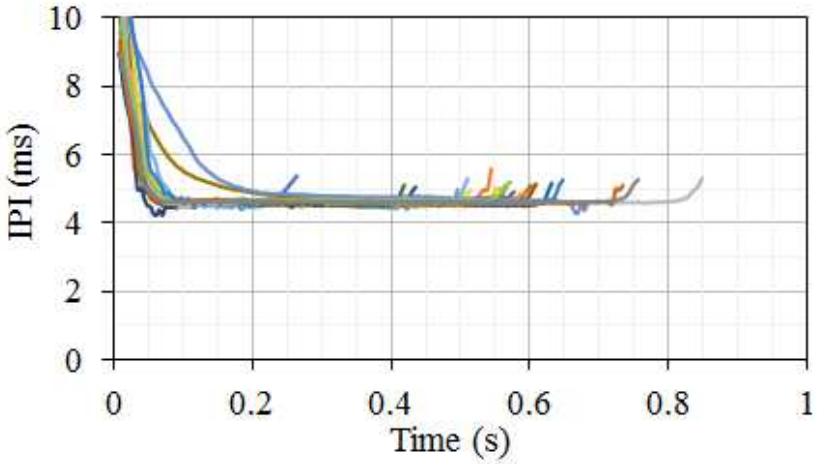
Fig. 3.11 Examples of PS1 calls from six belugas: a) SM1, b) SM2, c) SM3, d) SF1, e) SF2, and f) SF3. The top graphs represent waveforms, and the bottom graphs represent spectrograms (FFT size: 1024 points; window: Hamming; overlap: 50%).

Table 3.2 Temporal characteristics of PS1 calls.

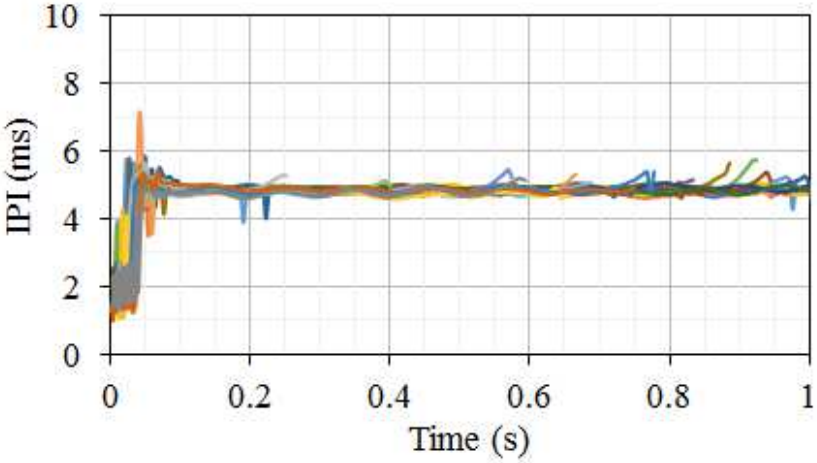
ID	n	N_p			DUR (s)			PRR (pulses/s)			IPI 1 (ms)			IPI 2 (ms)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
SM1	26	121.0 ± 28.4	180	53	0.58 ± 0.13	0.85	0.26	210.0 ± 5.3	215.9	196.3	4.73 ± 0.35	6.12	4.39	4.59 ± 0.06	4.79	4.48
SM2	30	168.1 ± 55.7	260	69	0.75 ± 0.27	1.19	0.25	227.3 ± 13.2	275.1	210.9	2.31 ± 1.08	5.19	1.42	4.83 ± 0.07	4.95	4.71
SM3	31	107.1 ± 16.2	148	78	0.31 ± 0.05	0.40	0.16	353.3 ± 55.6	525.5	249.9	1.18 ± 0.13	1.72	1.04	5.16 ± 0.24	5.62	4.35
SF1	53	158.3 ± 19.1	192	104	0.55 ± 0.08	0.70	0.37	289.2 ± 10.0	316.1	267.8	1.34 ± 0.06	1.56	1.22	4.17 ± 0.07	4.38	4.07
SF2	100	79.6 ± 9.3	110	55	0.37 ± 0.04	0.51	0.26	213.6 ± 3.5	223.5	206.2	4.45 ± 0.20	4.82	3.90	5.03 ± 0.09	5.23	4.85
SF3	21	81.1 ± 11.6	101	56	0.28 ± 0.05	0.37	0.21	294.3 ± 27.4	348.3	239.4	1.35 ± 0.21	2.12	1.16	5.26 ± 0.12	5.43	4.89
Kruskal–Wallis test		$\chi^2 = 181.51$, $P < 0.0001$			$\chi^2 = 178.72$, $P < 0.0001$			$\chi^2 = 213.51$, $P < 0.0001$			$\chi^2 = 214.85$, $P < 0.0001$			$\chi^2 = 214.67$, $P < 0.0001$		
CV_b		38.18			40.16			21.37			53.34			7.99		
Mean CV_w		18.44			19.74			6.48			15.10			2.20		
PIC		2.07			2.03			3.30			3.53			3.63		

Note: N_p, number of pulses; DUR, duration; PRR, pules repetition rate; IPI 1, average inter-pulse interval of pulse nos. 11–20; IPI 2, average inter-pulse interval of pulse nos. 11–20 from the last; CV_b, coefficient of variation between individuals; mean CV_w, mean of the coefficients of variation within individuals; PIC, potential for individual coding.

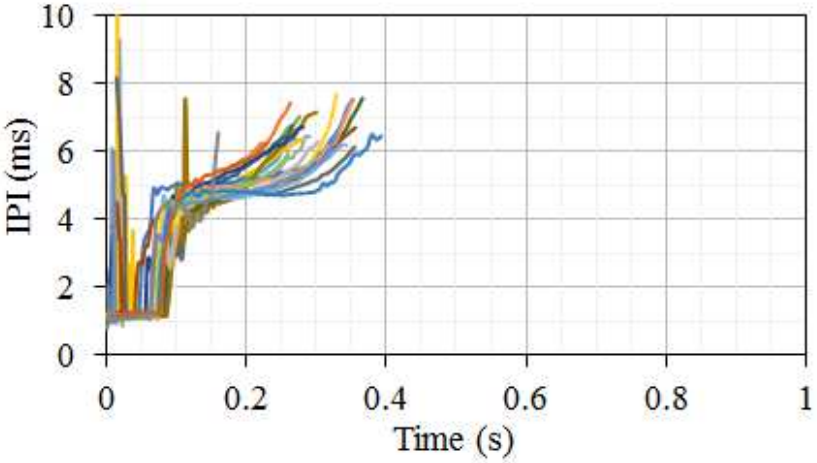
a) SM1



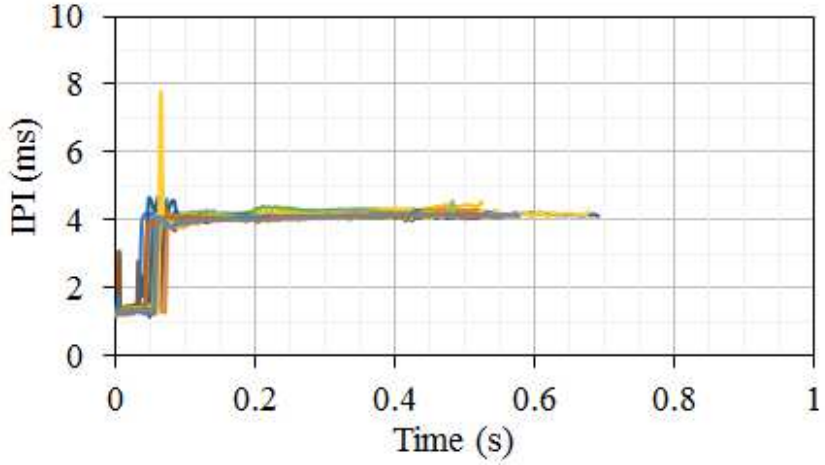
b) SM2



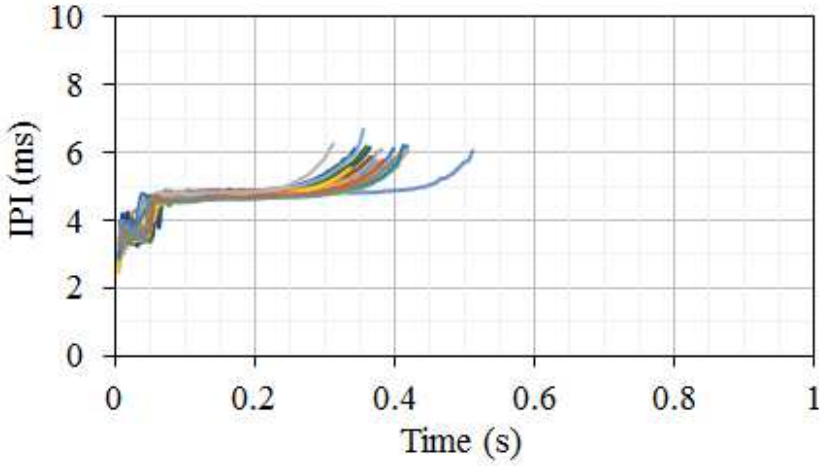
c) SM3



d) SF1



e) SF2



f) SF3

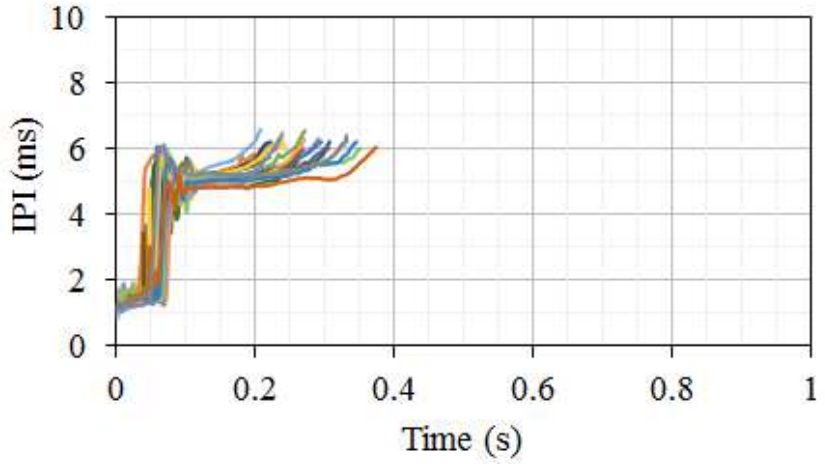
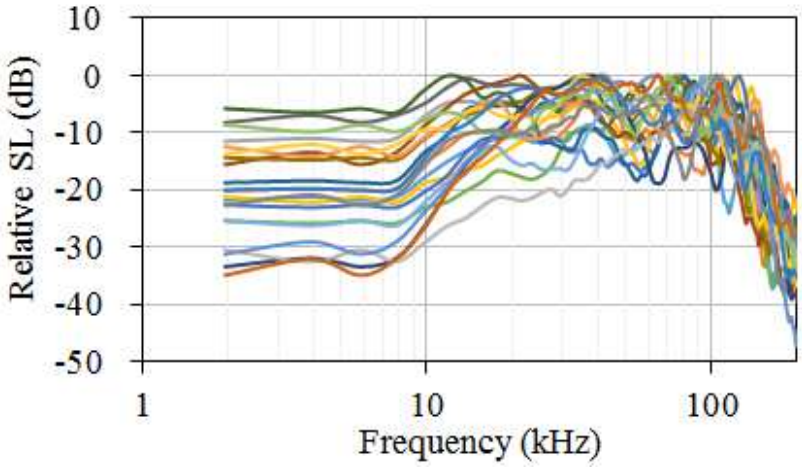
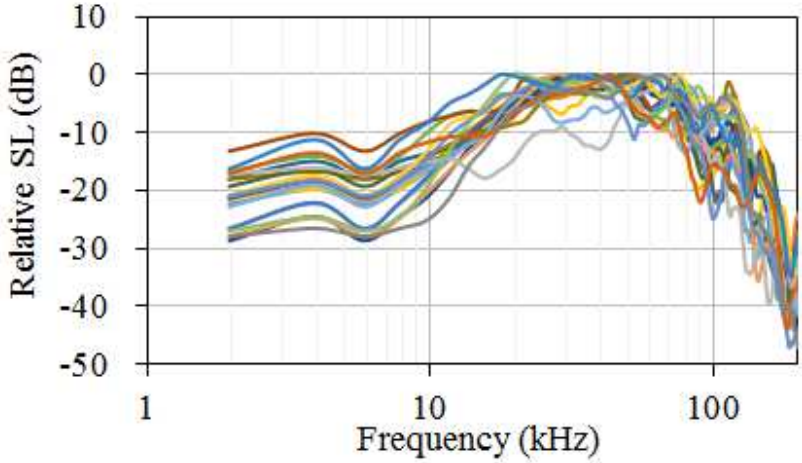


Fig. 3.12 IPI contours of PS1 calls from six belugas (n = 21): a) SM1, b) SM2, c) SM3, d) SF1, e) SF2, and f) SF3. Twenty-one examples each from SM1, SM2, SM3, SF1, and SF2 were randomly selected to match the number of depicted IPI contours of SF3 from which the smallest samples were collected.

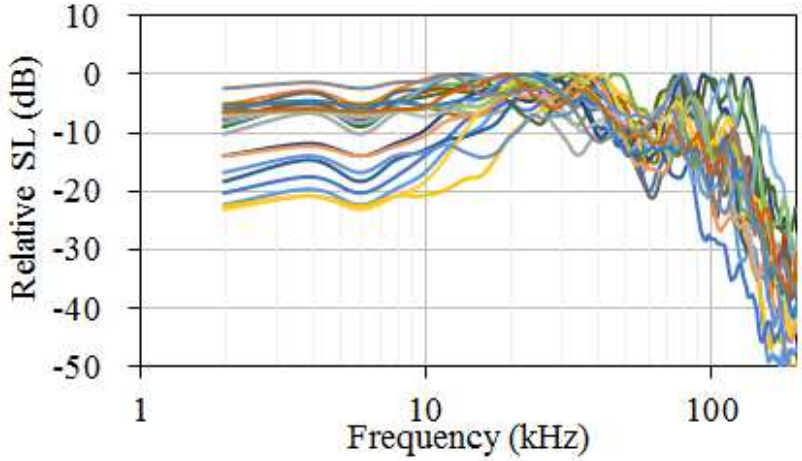
a) SM1



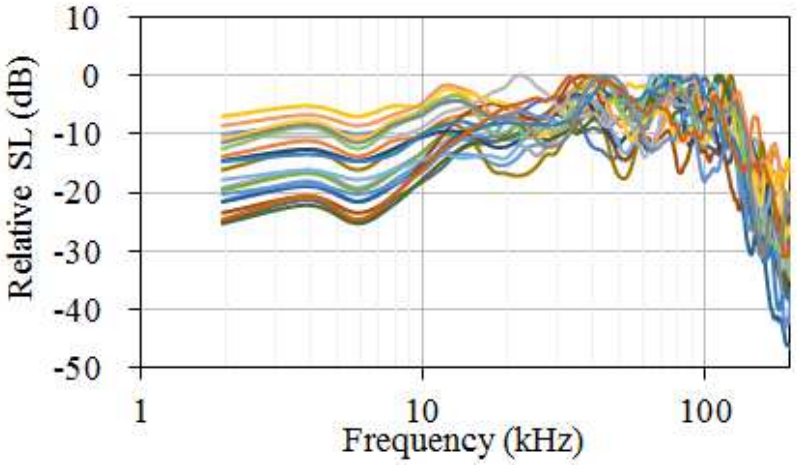
b) SM2



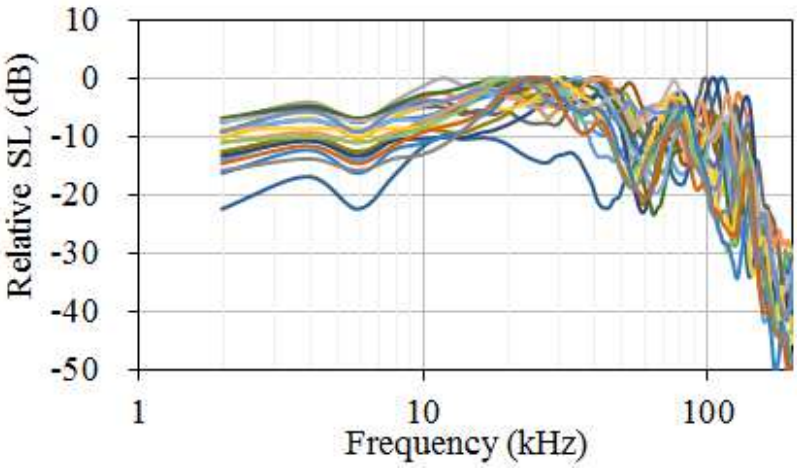
c) SM3



d) SF1



e) SF2



f) SF3

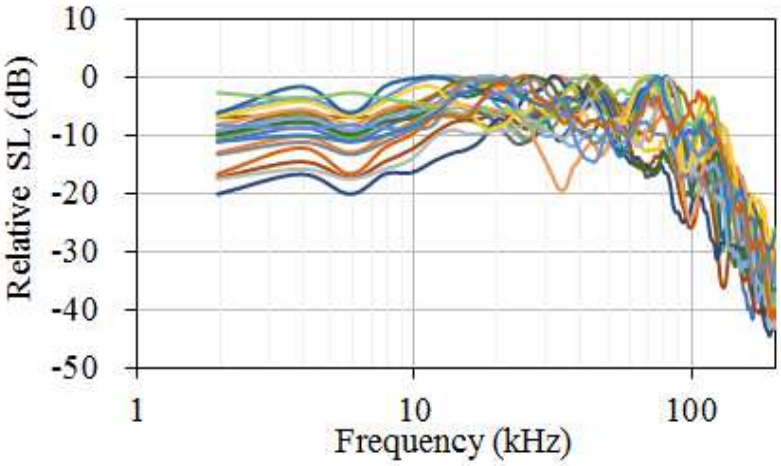


Fig. 3.13 Power spectra calculated at the middle pulse location within PS1 calls from six belugas (n = 16): a) SM1, b) SM2, c) SM3, d) SF1, e) SF2, and f) SF3 (FFT size: 256 points; window: Hamming; smoothing: 5 points). Twenty-one examples each from SM1, SM2, SM3, SF1, and SF2 were randomly selected to match the number of depicted power spectra of SF3 from which the smallest samples were collected.

Table 3.3 Spectral characteristics of the middle pulses within PS1 calls.

ID	n	F_p (kHz)			10 BW (kHz)			F_l (kHz)			F_u (kHz)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
SM1	26	75.3 ± 25.2	111.3	35.2	63.9 ± 29.2	130.9	21.5	40.0 ± 24.6	89.8	2.0	103.8 ± 24.9	142.6	52.7
SM2	30	44.9 ± 17.1	76.2	19.5	77.9 ± 25.0	132.8	33.2	15.5 ± 6.1	44.9	9.8	93.4 ± 24.4	144.5	46.9
SM3	31	37.0 ± 25.6	117.2	11.7	52.7 ± 26.7	136.7	15.6	14.3 ± 19.3	74.2	2.0	67.0 ± 32.5	144.5	25.4
SF1	53	64.6 ± 28.7	123.0	13.7	72.2 ± 29.5	130.9	15.6	27.9 ± 21.2	91.8	2.0	100.1 ± 26.3	140.6	43.0
SF2	100	38.4 ± 21.6	111.3	11.7	48.8 ± 15.6	95.7	21.5	13.4 ± 20.0	93.8	2.0	62.2 ± 20.6	123.0	37.1
SF3	21	47.6 ± 25.0	82.0	11.7	55.6 ± 29.2	123.0	25.4	20.1 ± 22.0	68.4	2.0	75.7 ± 26.7	125.0	29.3
Kruskal-Wallis test		–			$\chi^2 = 43.68$, $P < 0.0001$			$\chi^2 = 71.84$, $P < 0.0001$			–		
One-way ANOVA		$F(5,259) = 34.85$, $P < 0.0001$			–			–			$F(5,259) = 57.60$, $P < 0.0001$		
CV_b		56.24			44.66			108.76			38.19		
Mean CV_w		49.31			42.65			95.74			32.45		
PIC		1.14			1.05			1.14			1.18		

Note: F_p, peak frequency; 10 BW, 10 dB bandwidth; F_l, the lower frequency of the 10 BW; F_u, the upper frequency of the 10 BW; CV_b, coefficient of variation between individuals; mean CV_w, mean of the coefficients of variation within individuals; PIC, potential for individual coding.

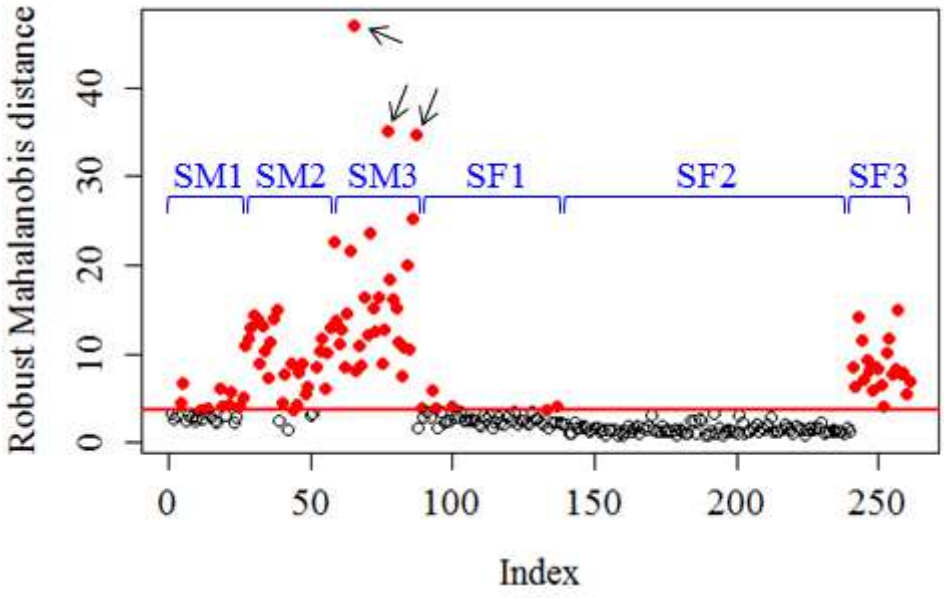


Fig. 3.14 Multivariate outliers searched by robust Mahalanobis distances with 97.5% quantile. The samples are represented from the left in the order of SM1, SM2, SM3, SF1, SF2, and SF3. Red circles show the outliers. The arrows indicate the outstanding outliers from SM3.

Table 3.4 Results of the quadratic DFA based on five parameters: DUR, PRR, IPI 1 IPI 2, and F_p.

	SM1	SM2	SM3	SF1	SF2	SF3	n	Correct rate (%)
SM1	25	1	0	0	0	0	26	96.2
SM2	0	29	0	0	0	1	30	96.7
SM3	0	0	25	0	0	3	28	89.2
SF1	0	0	0	53	0	0	53	100.0
SF2	0	0	0	0	53	0	53	100.0
SF3	0	0	6	0	0	15	21	71.4

3.4.5 Individual identity advertisement using PS1

PS1 call rate was compared among three sections of the special session: 12 min before the gate open, 12 min during the gate open, and 12 min after the gate open. PS1 call rate increased by the gate open and decreased after the gate open (Fig. 3.15). PS1 calls were frequently produced by all of the three adults during the gate open, whereas only SM1 produced PS1 calls for 12 min before the gate open and there were no PS1 calls for 12 min after the gate open. During the special session, SM1 produced three types of PS1: PS1-V1, PS1-V2, and PS1-V5, and SM2 produced only one type, PS1-V1. While SM1 used PS1-V1, PS1-V2 and PS1-V5 before the gate open, he limited to only PS1-V1 during the gate open.

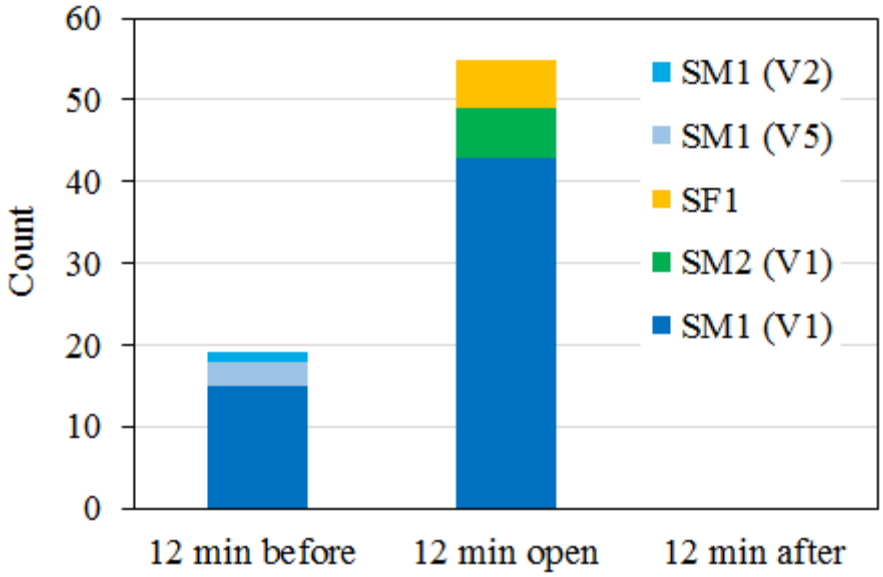


Fig. 3.15 Number of PS1 calls for three sections of the special session: 12 min before the gate open, 12 min during the gate open, and 12 min after the gate open. V1, V2, and V5 indicate PS1 variant type produced by males.

3.4.6 Context-dependent changes in PS1 acoustic parameters

Acoustic parameters of PS1 calls were compared between normal sessions and the 12-min gate open in the special session. Because few samples with good signal-to-noise ratio were collected from SM2 and SF1 during the gate open period, only SM1 samples were used for

comparison. Fourteen samples of PS1-V1 calls from SM1 had good acoustic quality, and their temporal and spectral parameters were compared to the results of normal sessions (Tables 3.5, 3.6). When the differences between the normal and special sessions were tested by Wilcoxon rank-sum test, Student's *t*-test, or Welch's *t*-test, N_p, DUR, and IPI 2 were significantly different ($P < 0.0001$). The few PS1 samples of SM2 and SF1 in the special session also tended to exhibit increased N_p and longer DUR than normal sessions, although IPI 2 values were comparable between them.

Table 3.5 Temporal comparison of PS1 calls from SM1 between normal sessions and the 12-min gate open in the special session.

Session name	n	N_p			DUR (s)			PRR (pulses/s)			IPI 1 (ms)			IPI 2 (ms)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
Normal	26	121.0 ± 28.4	180	53	0.58 ± 0.13	0.85	0.26	210.0 ± 5.3	215.9	196.3	4.73 ± 0.35	6.12	4.39	4.59 ± 0.06	4.79	4.48
Special	14	288.4 ± 39.1	360	205	1.39 ± 0.22	1.89	0.98	208.9 ± 9.5	217.3	176.1	5.25 ± 2.15	13.00	4.49	4.71 ± 0.05	4.78	4.62
Student's *t*-test		$t = -15.11$, $P < 0.0001$			—			—			—			$t = -5.85$, $P < 0.0001$		
Welch's *t*-test		—			$t = -12.27$, $P < 0.0001$			$t = 0.39$, $P = 0.70$			—			—		
Wilcoxon rank sum test		—			—			—			$W = 154$, $P = 0.44$			—		

Note: N_p, number of pulses; DUR, duration; PRR, pules repetition rate; IPI 1, average inter-pulse interval of pulse nos. 11–20; IPI 2, average inter-pulse interval of pulse nos. 11–20 from the last.

Table 3.6 Spectral comparison of PS1 calls from SM1 between normal sessions and the 12-min gate open in the special session.

Session name	n	F_p (kHz)			10 BW (kHz)			F_l (kHz)			F_u (kHz)		
		Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min	Mean ± sd	Max	Min
Normal	26	75.3 ± 25.2	111.3	35.2	63.9 ± 29.2	130.9	21.5	40.0 ± 24.6	89.8	2.0	103.8 ± 24.9	142.6	52.7
Special	14	79.1 ± 22.1	130.9	39.1	75.1 ± 28.0	136.7	25.4	40.6 ± 20.2	76.2	7.8	115.7 ± 21.5	144.5	84.0
Student's *t*-test		$t = -0.46$, $P = 0.65$			$t = -1.14$, $P = 0.26$			$t = -0.08$, $P = 0.94$			$t = 1.46$, $P = 0.15$		

Note: F_p, peak frequency; 10 BW, 10 dB bandwidth; F_l, the lower frequency of the 10 BW; F_u, the upper frequency of the 10 BW.

3.5 Discussion

This chapter demonstrated PS1 function and individual distinctiveness of PS1 in another aquarium to conclude that PS1 characteristics were common in belugas. PS1 was the most frequently produced call type in both pools (Fig. 3.4). PS1 was produced in bouts, and there was no relation between the PS1 bouts and aggressive/submissive events. Thus, PS1 is likely to be an affiliative call.

PS1 was used for vocal exchange. Most of the observed WIIs occurred after 1 s, with a first peak at 1.6–1.8 s, but more than a half of observed BIIs occurred within 1 s, with a sharp peak at 0.4–0.6 s (Fig. 3.8). The frequency distribution of observed BIIs was significantly different from that of simulated BIIs, and it suggested that observed BII distribution indicated a temporal rule regulating vocal exchange, not just an incidental result. These results led to the conclusion that the belugas exchanged PS1 calls in accordance with a response rule where responding individuals called back within approximately 1 s. The previous study on inter-PS1 intervals in the Port of Nagoya Public Aquarium showed that frequency distribution of BII had a sharp peak at -0.5 s, while frequency distribution of WII had a gradual peak at 1 s (Morisaka et al., 2013). The interval criterion was the latency period from the end of a PS1 to the beginning of the next PS1, and was shorter by PS1 duration than my criterion. When the distributions of the previous study were shifted to the right along the temporal axis by 0.85 s—the average PS1 duration of Morisaka et al. (2013) —their distribution resembled that of this study.

The adult female SF2 produced the highest number of PS1 calls in Pool A. Based on previously utilized criteria (Recchia, 1994), she seemed to be top-ranked in Pool A because of possessing the largest body size. Likewise, the most vocal individual in Pool B, adult male SM1, was the largest beluga. In addition, both SF2 and SM1 were strong characters. These coincided with the results of chapter 2; the highest PS1 production rate was noted in the

potential top-ranked NF1 and supported the suggestion that PS1 production rate related to social rank.

In chapter 2, it was suggested that PS1 was acquired by vocal production learning (Janik & Slater, 2000) and/or morphological development as the PS1 calls of 21-month-old NM2 did not have fixed IPI contours. This assumption was supported by the fact that SF4, who is approximately between four and six months old during BM recording sessions, did not produce any PS1 calls.

There were variations in the PS1 calls of males (Fig. 3.9), with the PS1-V1 of each male being the predominant PS1 variant (Fig. 3.10). The PS1-V4 of SM1 was a slightly shifted IPI contour of his PS1-V1, and PS1-V2 of SM2 was a slightly shifted IPI contour of his PS1-V1. Noise did not appear to influence those shifts because there was no distinct noise during the recordings. The shifted PS1 variants might have a different biological meaning from PS1-V1 or could just be variations within PS1-V1 and should have been classified into PS1-V1. The PS1-V2 of SM3 was a repeated pattern of his PS1-V1 and it was similar to the loop structure in signature whistles (Sayigh et al., 1990). In bottlenose dolphins, the number of loops was affected by motivational state (Esch et al., 2009). When he produced the PS1-V2, he swam in the pool but the heads of other individuals were above the surface of the water and they produced PS1 calls in air. Thus, it is speculated that there were unusual sounds in air and they were excited at that time. The excited inner state perhaps caused SM3 to produce loop-structured PS1-V2. Mature males of bottlenose dolphins disperse from their matrilineal group and made stable alliances with other adult males (Connor & Krützen, 2015). Signature whistle convergence and sharing were found among allied males (Smolker & Pepper, 1999; Watwood et al., 2004). Although little is known about beluga male society, adult males appeared to facilitate long-term social bands with other mature males (Michaud, 2005; Colbeck et al., 2013). Thus, the similarity between PS1-V4 of SM1 and PS1-V4 of SM2 was possibly the

result of convergence and/or sharing as seen in bottlenose dolphins. There were other PS1 variants. Bottlenose dolphins also produce several non-signature whistles (Tyack, 1986; Sayigh et al., 1990; Janik et al., 1994; McCowan & Reiss, 1995b; Janik & Slater, 1998; Nakahara & Miyazaki, 2011) and males have a broader whistle repertoire (Tyack, 1986; Sayigh et al., 1990). Although the role of non-signature whistles was unclear, it is reported that most of the whistles produced by allied males in consortship with a female are non-signature whistles (Watwood et al., 2005). Thus, PS1 variants also possibly related to consortship with females, but larger samples from males will be required to uncover the role of PS1 variants.

Only NM1 in the Port of Nagoya Public Aquarium had a tonal component in PS1, while all of the belugas in Shimane Aquarium had tonal or narrowband pulsed components that co-occurred with the main broadband pulse train in PS1 (Fig. 3.11). The overlapping components of SM1, SF1, and SF2 were individually distinctive, while those of others had less individuality. The main pulse trains of PS1 calls in Shimane Aquarium were similar to those of Type A calls described in Vergara et al. (2010). As described in chapter 2, of the five Type A call variants, the overlapping of a tonal component is found in type A1. The A1 call was produced by an adult female and her two offspring. A1 has an average PRR of 94.6 pulses/s, is 1.2–1.9 s in duration, and consistently contains a tonal component at 14.6 kHz. Overlapping of a pulsed component is also found in the A3 call. The A3 call was produced by an adult female and her daughter. It has an average PRR of 306.4 pulses/s and duration of 1.2–1.9 s and has a secondary pulsed component that synchronizes with the main pulse train. Belugas live in pack ice or polynyas in winter, which are extremely noisy and reverberant environments (Brown & Milne, 1967). The biphonation perhaps plays a role in the enhancement of individual recognition in this noisy environment as seen in emperor penguins (*Aptenodytes forsteri*), king penguins (*Aptenodytes patagonicus*) (Aubin et al., 2000) and dhole (*Cuon*

alpinus) (Volodina et al., 2006). Recordings from wild belugas and playback experiments will be needed for the functional interpretation of the overlapping tonal or pulsed components in PS1 and Type A calls.

Individuality was found in various parameters and the quadratic DFA classified PS1 calls into individuals with an overall classification rate of 94.8%. Classification of individuals by PS1 calls was also possible spectrographically and by ear. Thus, as belugas possess enhanced hearing capabilities compared to humans (Fay, 1988; Klishin et al., 2000), it appeared to be very easy for belugas to recognize individual PS1 calls.

Intra-individual consistency and inter-individual difference existed in IPI contours of PS1 calls (Fig. 3.12). In addition, PIC results showed that PRR, IPI 1, and IPI 2 had high PIC scores greater than 3 (Table 3.2). Further, the most informative parameters in DFA were IPI 2 followed by IPI 1. These results suggested that pulse repetition pattern had a high potential as an individual identification media. An abrupt change was found in the initial part of IPI contours (Fig. 3.12). Further, overlapping exchanges often occurred, as seen in the PS1 calls reported by Morisaka et al. (2013) and Type A calls noted by Vergara et al. (2010). In addition, IPI 2 varied dependent on context (Table 3.5). These results supported the hypothesis that the initial part of IPI contours contained sufficient information for individual identification, as described in chapter 2.

Here all temporal parameters were significantly different among individuals (Table 3.2), whereas in the results of chapter 2, duration lacked individual distinctiveness (Table 2.3). However, N_p, DUR, and IPI 2 varied dependent on context in this study (Table 3.5). The PS1 calls in the 12-min gate open in the special session had more N_p and longer DUR than those of normal sessions. The subjects might need those changes to assert individual identity in a context of visual reunion after a long separation. Therefore, there still remains the speculation that duration, which is related to the number of pulses, is affected by motivational state and

does not carry individual information as do signature whistles of bottlenose dolphins (Esch et al., 2009).

Visual inspection did not find clear individuality in power spectra of middle pulses (Fig. 3.13), although statistical analyses revealed that all of the spectral parameters differed significantly among individuals and had PIC values greater than 1 (Table 3.3). These results were similar to those in chapter 2. Given that human observers have proven to perform better than computers at classifying vocalizations (Janik, 1999), spectral cues seemed not to encode apparent individuality. Directivity may explain the intra-individual variability as described in chapter 2.

The PS1 call rate was increased by the 12-min gate open in the special session, which was a context of visual reunion after a long separation for the males and SF1, and the first opportunity for the males to see the calf (Fig. 3.15). In addition, the males produced only the predominant, individually distinctive PS1 variant PS1-V1 at that time. These results suggest that the individualized PS1 calls served as individual advertisement. An alternative explanation is that the unusual situation, or gate open, caused the high PS1 call rate to increase the maintenance of group cohesion.

This chapter exhibited that PS1 served as an affiliative contact call and that while individual difference existed in various temporal and spectral parameters, IPI contours had highly distinctive individuality. These results correspond with those of chapter 2. Therefore, there is a high possibility that the function and individuality of PS1 calls are common features among beluga whales. Playback experiments will elucidate whether belugas use PS1 for individual recognition and which acoustic parameter is the recognition cue.

Chapter 4: Establishment of a broadband transmitting system for PS1 playback experiments

4.1 Introduction

Playback experiments are used to investigate the behavioral or vocal responses of animals to acoustic stimuli reproduced using a speaker. The method can be a powerful tool for revealing animal cognition and has been incorporated in several study areas, including vocal function, or kin and individual recognition in odontocetes (Deecke, 2006).

In bottlenose dolphins, playback experiments have been used since the 1960s. Lang & Smith (1965) broadcast an audio track of an isolated dolphin to another isolated conspecific to investigate whistle function in bottlenose dolphins. The target dolphin frequently responded to a specific whistle type in the recording with a specific type out of his whistle repertoire; therefore, it was suggested that those particular whistle types might be used for localization or identification of other dolphins. Dreher (1966) also elicited different behavioral and vocal responses to presentation of six whistle types. Caldwell et al. (1972) trained a dolphin to react positively to the signature whistle stimuli of four conspecifics and not to react to those of four other similar-aged conspecifics. The dolphin could respond correctly to the randomly presented whistles of the eight different conspecifics. This indicated that dolphins could easily differentiate conspecifics' signature whistles. A later study by Sayigh et al. (1999) also used playback experiments to reveal whether dolphins could distinguish signature whistles of different familiar individuals. The target mothers responded more strongly to the signature whistles of their own calf than to those of a similar-aged non-related calf. Likewise, calves responded more strongly to the signature whistles of their own mother than to those of a familiar, similar-aged female. These results led to the conclusion that signature whistles were used for individual recognition. Subsequently, Janik et al. (2006) examined by using synthesized signature whistles whether dolphins could extract individual identity information

from the frequency modulation patterns even after all voice features have been removed from them. Dolphins reacted more strongly to the synthesized signature whistles of a close-related individual than to those of an unrelated individual, regardless of the level of similarity between each stimulus and their own whistles. This provided compelling evidence that sufficient information on identity was encoded in the frequency modulation patterns of signature whistles. Nakahara & Miyazaki (2011) found that signature whistle exchanges were regulated by a rule in which the respondents called back within 1 s, and they implemented playback experiments to support the findings experimentally. As expected, dolphins mostly responded with their own signature whistles within 1 s from the exposures to the associates' signature whistles. They also reported that in rare cases, dolphins called back copies of the presented associates' signature whistles. The function of occasional signature whistle copying and matching was uncovered by a playback study of King & Janik (2013). Dolphins more often replied with their own signature whistles when exposed to the same whistle type than to other whistle types from familiar and unfamiliar individuals. It suggested that the copying of conspecifics' signature whistles can be used as a label to address particular individuals and that the addressed individuals responded with their own signature whistles. Long-term memory of signature whistles was also proven by playback experiments, in which dolphins presented a stronger response to the signature whistles of former tank mates than to those of strangers, even if the target dolphin and the former tank mate were separated for up to 20 years (Bruk, 2013).

Playback experiments regarding kin recognition are also carried out with other species including sperm whales and killer whales. Rendell & Whitehead (2005) played back codas to sperm whale groups to examine whether the coda playbacks evoked their coda production and whether they showed a differential response to codas from their own clan over codas from other clans. Although they failed to find a consistent reaction, with most coda exposure

eliciting no clear reaction, Filatova et al. (2011), who performed similar playback experiments to killer whale units, found that the recipient units always responded vocally to the playbacks of discrete calls from the same pods, but they never responded vocally to the playbacks of discrete calls from different pods. They concluded that killer whales could discriminate between discrete calls of their own and different pods.

In belugas, Morgan (1979) conducted playback experiments with a captive group at the New York Aquarium and wild belugas in the Saguenay River, Canada, to investigate vocal function of several call types. This provided the possibility that certain specific calls and combinations of calls had particular significance, but obvious functions were not found.

To elucidate PS1 function or kin and individual recognition of PS1 calls, I should perform several playback experiments. However, the reproduction of broadband calls is problematic, and such calls cannot be effectively produced through playback. The reproducible frequency band underwater has been limited to frequencies lower than 20 kHz, as there have been no dedicated broadband underwater speakers, and conventional speakers have transmitting sensitivity only below 20 kHz. These speakers cannot faithfully reproduce broadband pulse-type vocalizations such as PS1 not only in the spectral domain but also in the temporal domain, because there is an inversely proportional relationship between bandwidth and pulse width. Actually, the reproduced sounds in the studies described above were whistles or only the low frequency part of pule trains. Therefore, to conduct playback experiments with PS1 recordings, it is necessary to start by developing a broadband underwater speaker and establishing a broadband transmitting system.

4.2 Requirements for the broadband transmitting system

Requirements for the broadband transmitting system were determined based on the acoustic properties of PS1. The transmitting system should cover the frequency band from <1 kHz to

>170 kHz in which PS1 has consistent energy. In addition, it should have the power to project sounds of at least 130–160 dB (root mean square: RMS) re 1 μPa at 1 m, which is the typical SL of PS1. Further, any equipment should not generate a large amount of noise and distort waveforms (Deecke, 2006).

The SLs of PS1 calls in both aquariums were roughly estimated by using the equation (4-1):

$$SL = 20\log(V) + |M| - G + 20\log R + \alpha R \qquad (4\text{-}1),$$

where V is the output voltage of the preamplifier, M is the receiving sensitivity of the hydrophone, G is the gain of the preamplifier, R is the distance between the hydrophone and the sound source, and α is the absorption coefficient. I used RMS value for V and did not consider directivity for the calculation of V, as the directivity of PS1 was unknown. R was roughly estimated from the video data, and αR was a small value and it was ignored in this study.

4.3 Development of a broadband underwater speaker

4.3.1 Broadband technique of a transducer used for sensing aquatic animals

An underwater speaker is one of the most important components of the transmitting system, which changes the electronic signals into pressure signals, or sounds. Therefore, the principle step should be the development of a broadband underwater speaker.

Echo-sounding systems are used for sensing fishes and planktons using sounds in order to estimate marine resources. This field also desired the development of broadband transducers, because broadband signals improve range resolution, and classification and discrimination of the scattering sources (Stanton, 2009). Some broadband transducers have been developed (Foote, 1998; Mortensen et al., 1999; Imaizumi et al., 2008) and also a successful prototype

broadband transducer spanning the frequency band from 20 kHz to 150 kHz with a transmitting sensitivity of about 160 dB re 1 μPa/V at 1 m (Amakasu et al., 2013). Thus, the broadband technique of the prototype transducer could be usefully applied to a broadband underwater speaker for PS1 playback experiments.

The target frequency band of Amakasu et al. (2013) was from 38 kHz to 120 kHz. They used 10 mm-square resin-coated type of multilayer piezoelectric actuators (NEC/TOKIN Co., Japan; Fig. 4.1 a), which had large generated force with low power consumption and had resonances at 134 kHz and 161 kHz (Fig. 1 described in Amakasu et al., 2013). In this study, the actuators were referred to as high-frequency actuators. A very wideband transmitting sensitivity existed in the high-frequency actuator itself and it covered the frequency range higher than 80 kHz. To increase the sensitivity in the frequency range lower than 80 kHz, the actuators were Langevin structured. In Langevin structure, actuators are sandwiched by two masses one each at the front and rear. This structure generated resonances at lower frequencies than the resonant frequencies of actuator itself. When 30 high-frequency actuators were Langevin structured using 13 mm-thick and 110 mm-diameter acryl disks for front and rear masses, four resonances appeared at 34, 60, 131, and 164 kHz. The 34 kHz and 60 kHz resonances seemed to be attributed to the Langevin structure, while the 131 kHz and 164 kHz resonances were attributed to the actuators themselves. As a result, the transducer covered the frequency band from 20 kHz to 150 kHz.

4.3.2 Application of the broadband technique to an underwater speaker

The frequency band of Amakasu et al. (2013) is still not enough for an underwater speaker used for PS1 playback experiments, because the sensitivity drops down below 20 kHz. Therefore, the addition of another type of the multilayer piezoelectric actuator, metal case type (NEC/TOKIN Co., Japan; Fig. 4.1 b) with a transmitting surface of 35 mm in diameter was

required to compensate for poor sensitivity below 20 kHz (Mishima et al., 2013; Fig. 4.2). The resonant frequencies emerged at 8 kHz and 20 kHz, and the actuator was termed a low-frequency actuator (Fig. 4.3).

Seven high-frequency actuators and one low-frequency actuator were placed in the same acryl disk of 13 mm thickness and 110 mm diameter (Fig. 4.4). The directivity increases as the diameter of a sound source increases. Amakasu et al. (2013) arranged 30 high-frequency actuators equally in the acryl disk and made the entire disk vibrate to realize a sharp directivity with beamwidth of 6.6° at 120 kHz. Although the directivity of PS1 was unknown, it appeared to be lower than that of the echolocation clicks produced to broadcast the caller's message. Therefore, the vibration area of the high-frequency actuators was reduced in this study and was adjusted close to the transmitting surface of the low-frequency actuator. Thus, seven high-frequency actuators were placed approximately within a 35 mm circle on one side of the front acryl disk, and the low-frequency actuator was put on the other side (Fig. 4.4). Two circular slits of 50 mm and 38 mm in diameter were made surrounding the high- and low-frequency actuators, respectively, to simplify the vibration mode as much as possible. Only high-frequency actuators were Langevin structured using another acryl disk for a rear mass, and the low-frequency actuator was fixed in the front disk using bolts. This structure was embedded in a vinyl chloride housing (Fig. 4.5).

a) Resin-coated type b) Metal case type

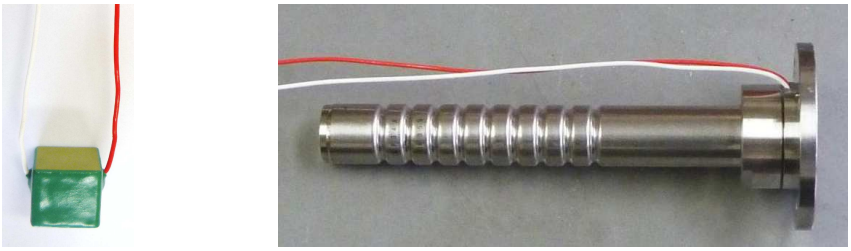


Fig. 4.1 Multilayer piezoelectric actuators; a) resin-coated type and b) metal case type (NEC/TOKIN Co., Japan).

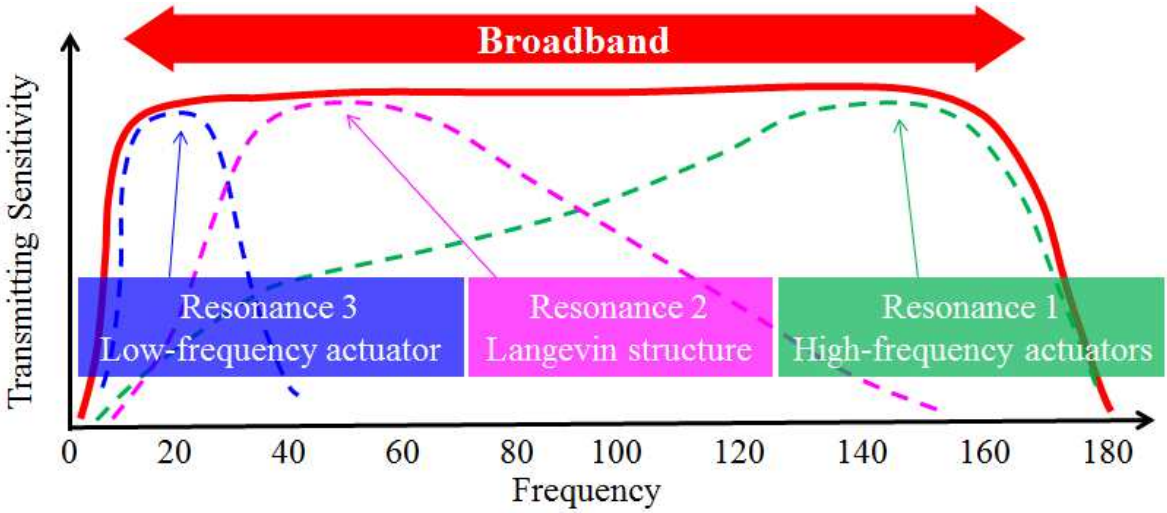


Fig. 4.2 Broadband technique for an underwater speaker.

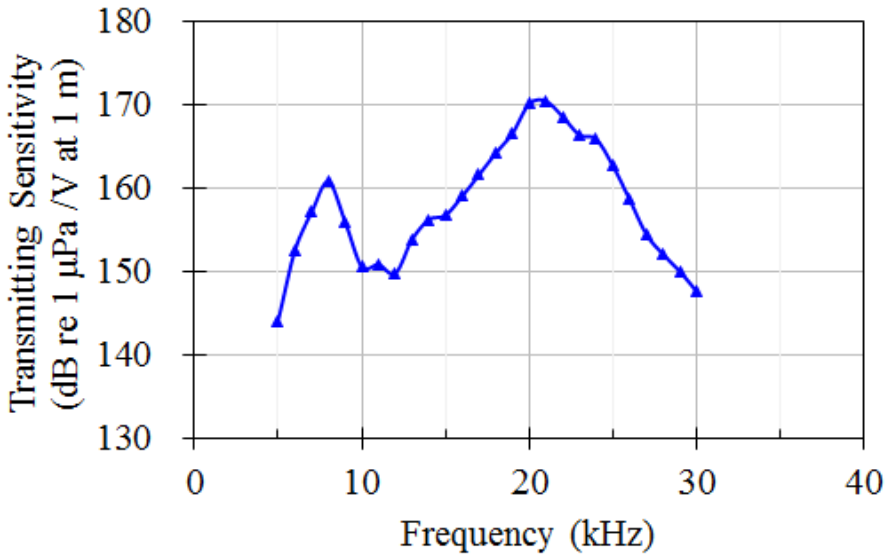


Fig. 4.3 Transmitting sensitivity of the low-frequency actuator.

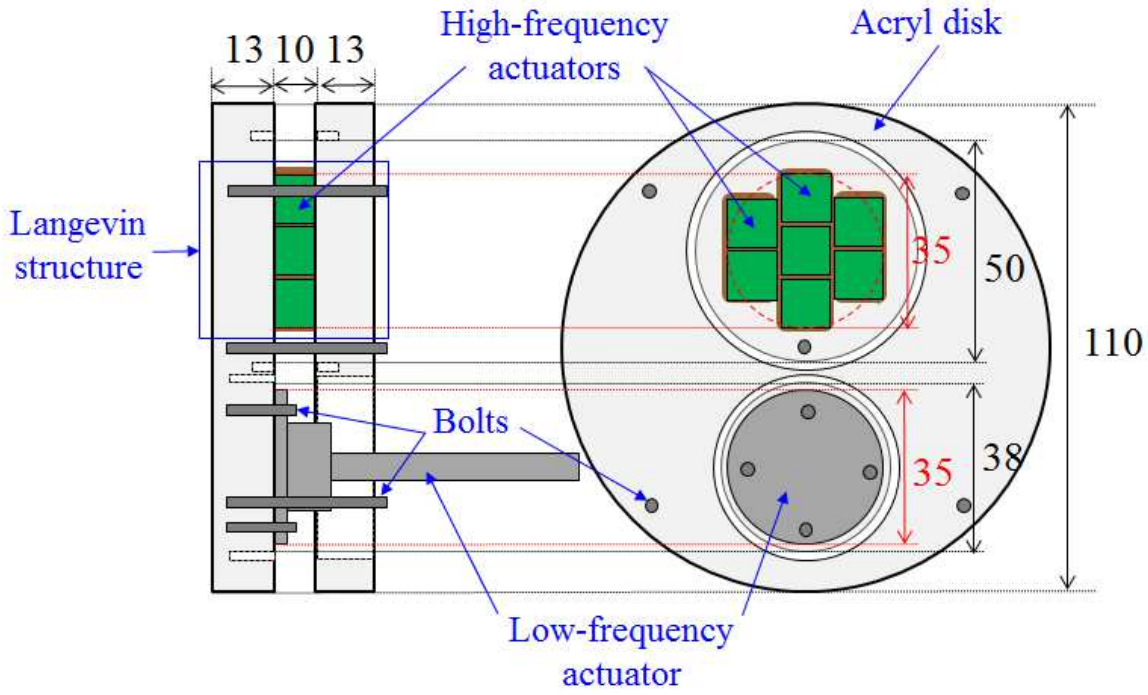


Fig. 4.4 Design of the broadband underwater speaker.

Fig. 4.5 Developed broadband underwater speaker.

4.4 Establishment of a broadband transmitting system

The transmitting system was composed of the developed broadband underwater speaker, a GTO 504 power amplifier (JBL Inc., USA), a NI USB-6351 data acquisition device (National Instruments Co., USA), and the MATLAB software (2015). The power amplifier has flat sensitivity from 10 Hz to 100 kHz within -3 dB. The gradually declining sensitivity above 100 kHz appeared to be useful in suppressing the resonance of the high-frequency actuators at around 130 kHz–140 kHz (Fig. 1 described in Amakasu et al., 2013). The power amplifier can amplify two channels separately and therefore I can regulate the power delivered to each type of actuator of the speaker to provide a uniform output level. The data acquisition device also corresponds to the frequency band and converts analog signals sampling at 500 kHz and 16 bits.

4.5 Evaluation of the broadband transmitting system

4.5.1 Methods

Evaluation tests were conducted at a depth of approximately 30 m in Tateyama Bay on June 22, 2015. The underwater speaker and a TC 4013 hydrophone (Reson Inc. Denmark) which exhibits a flat frequency response from 1 Hz to 170 kHz (Fig. 4.6) were mounted on a stainless steel frame and positioned 1 m apart from each other (Fig. 4.7), and the assembly was submerged horizontally at a depth of 2 m.

An up-chirp signal with frequency modulation from 1 kHz to 180 kHz was used to evaluate the transmitting system (Fig. 4.8). The chirp signal continued for 15 ms and was transmitted three times with a 200 ms interval. Onset and offset time were set to 400 μs; subsequently, the frequency range of approximately 5 kHz to 175 kHz was available for the evaluation. The signals were created by using the MATLAB software (2015).

The diagram of the transmitting and receiving system is represented in Fig. 4.9. The signals

projected by the transmitting system were received by the hydrophone. The output signals from the hydrophones were analog high-pass filtered at 1 kHz, and amplified by 32 dB using VP1000 preamplifiers (Reson Inc., Denmark), with a flat frequency response to 1 MHz (-3 dB). The NI USB-6351 data acquisition device (National Instruments Co., USA) digitized the output signals sampling at 500 kHz and 16 bits. The digital signals were collected and analyzed using the MATLAB software (2015). To avoid confusion, input signals to the DA converter were defined as “transmitting” signal, output signals from the speaker as “reproduced” signal, and output signals from the AD converter as “receiving” signals. Recordings were made six times with the level of the transmitting chirp signals, referred to as transmission level, varying in six steps: 2.5, 5, 10, 20, 40, and 80 mV.

The receiving chirp waveform and the transmitting chirp waveform were cross correlated, and the impulse response of the transmitting and receiving system was computed. The amplitude and phase information were extracted from the impulse response. The amplitude spectrum and the phase spectrum were obtained via fast Fourier transformation (FFT) with a size of 10000 points. SL was calculated by using the equation (4-1) and SL spectrum was constructed. The receiving sensitivity of the hydrophone in each frequency (Fig. 4.6) and the gain of the preamplifier were considered for SL calculation.

To examine the mechanical noise generated by operation of the transmitting system, two recordings were compared; one was an environmental noise recording where all power supply to the transmitting equipment was turned off, and the other was a silence track recording where all transmitting equipment was operational but the signal was not projected

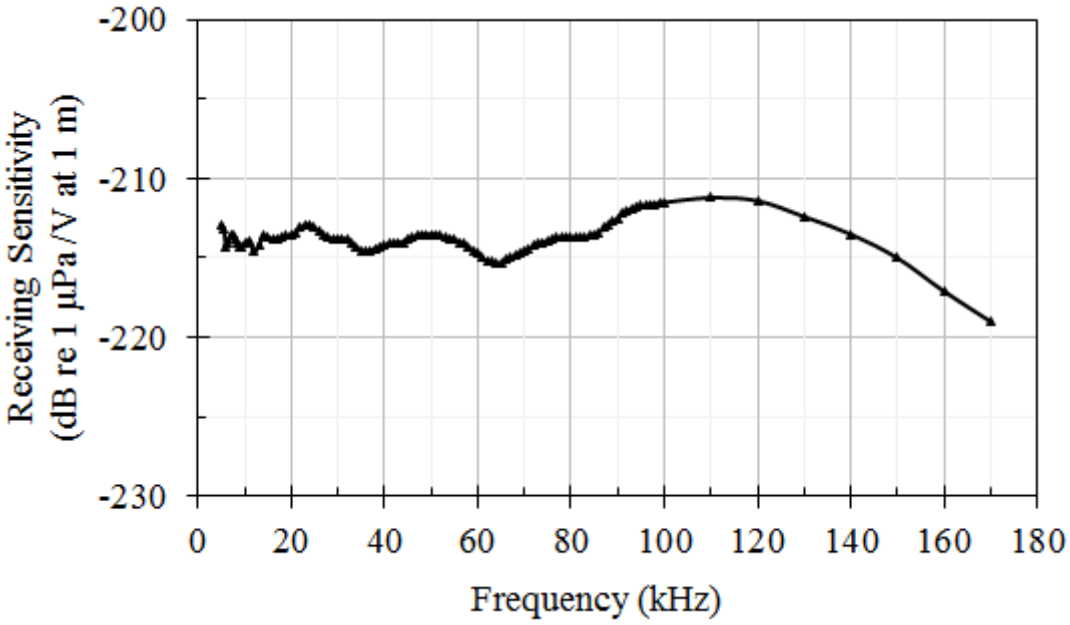


Fig. 4.6 Receiving sensitivity of a TC 4013 hydrophone (Reson Inc., Denmark).

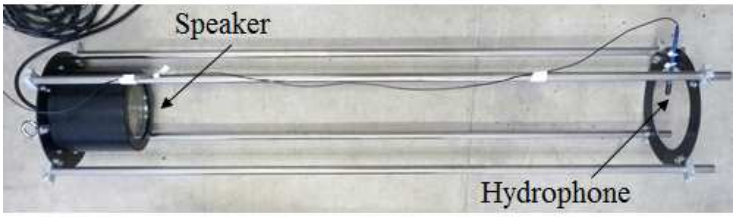


Fig. 4.7 A speaker and a hydrophone attached to a frame. They were separated by a distance of 1 m.

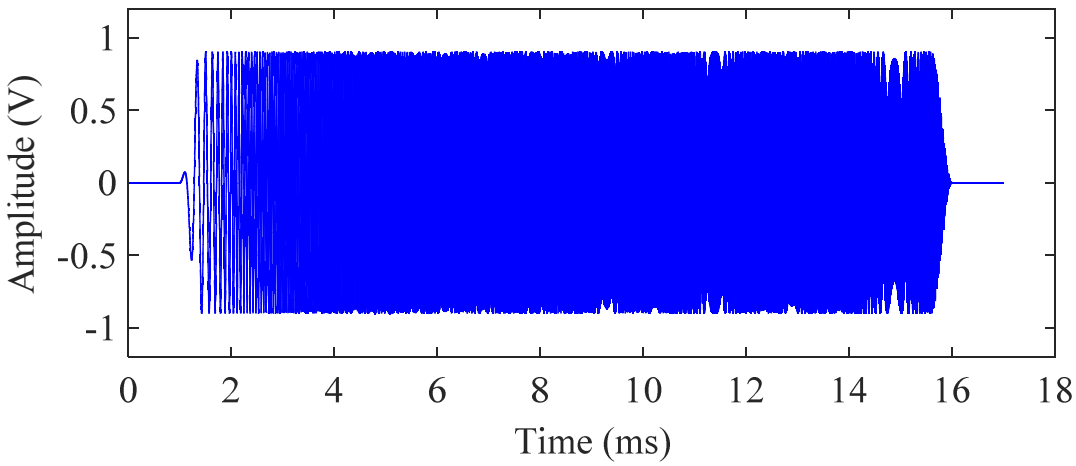


Fig. 4.8 Transmitting chirp signal with frequency modulation from 1 kHz to 180 kHz within 15 ms. Onset and offset were set to 400 µs.

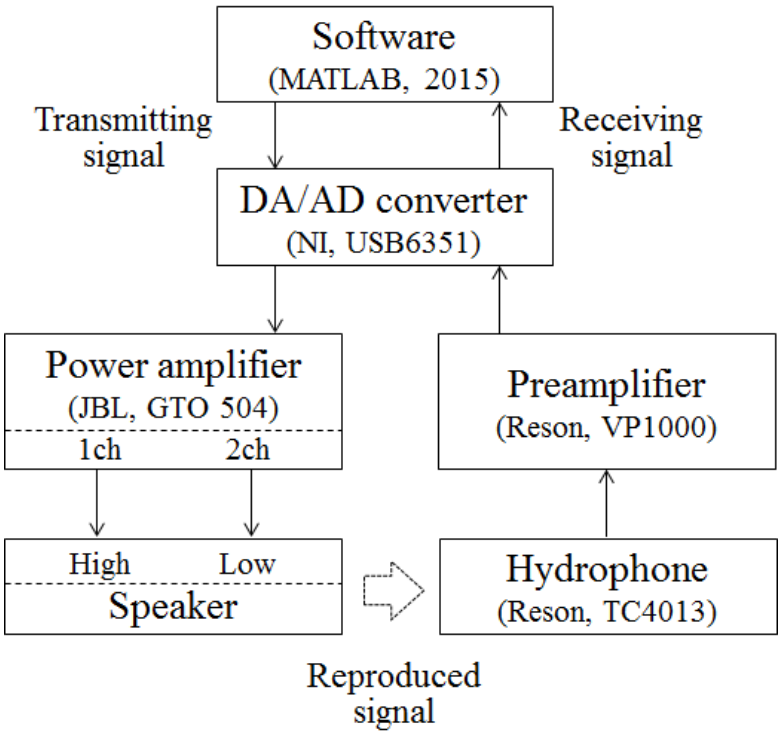


Fig. 4.9 Transmitting and receiving system.

4.5.2 Results

The receiving chirp waveform at a transmission level of 20 mV is shown in Fig. 4.10. Its spectrogram demonstrates that the broadband chirp signal is reliably reproduced (Fig. 4.11). The SL spectra at all of the six transmission levels are shown in Fig. 4.12. It revealed that the transmitting system had flat frequency response in the range of at least 5 kHz–175 kHz with ±12 dB ripple and could project sounds at least 130–170 dB re 1 μPa at 1 m. In the frequency band, the SL appeared to grow proportionally to the transmission level. For instance, Fig. 4.13 shows the linear relationship between the transmission level and the SL at three frequencies: 40, 80, and 120 kHz. Thus, I could hold linearity to calculate transmission levels according to desired SLs, at least in the range of 130–170 dB.

Strong harmonics and constant noise were not found in the receiving chirp spectrogram (Fig. 4.11). In addition, a proportional relationship between the transmission level and the SL suggested that the chirp signals were not saturated (Fig. 4.12, 4.13). Further, noise levels did

not differ between the environmental noise recording and the silence track recording (118.9 dB and 118.7 dB, respectively). In addition, Fig. 4.14 shows spectrographically that there are no distinct noise artifacts in the silence track recording. These results suggest the transmitting system did not generate waveform distortion or evident mechanical noise from any constituent equipment.

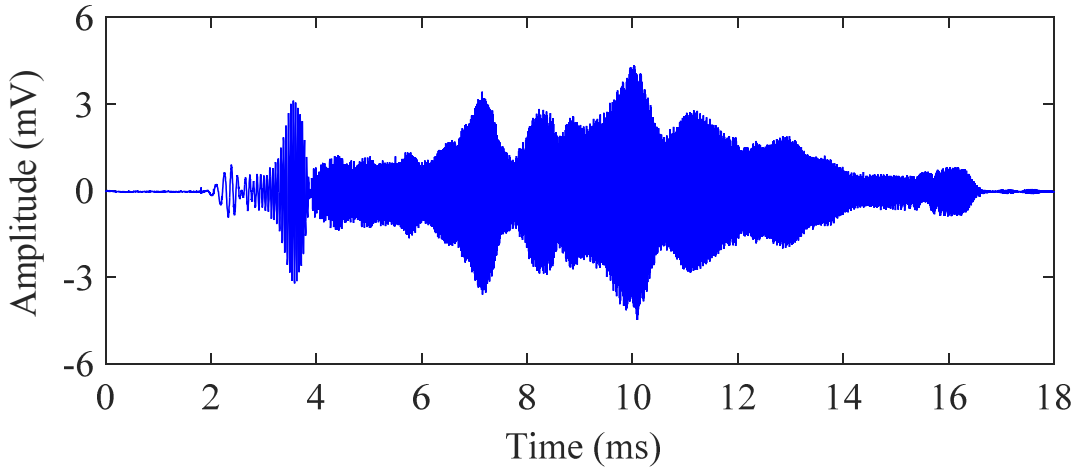


Fig. 4.10 Receiving chirp waveform at a transmission level of 20 mV.

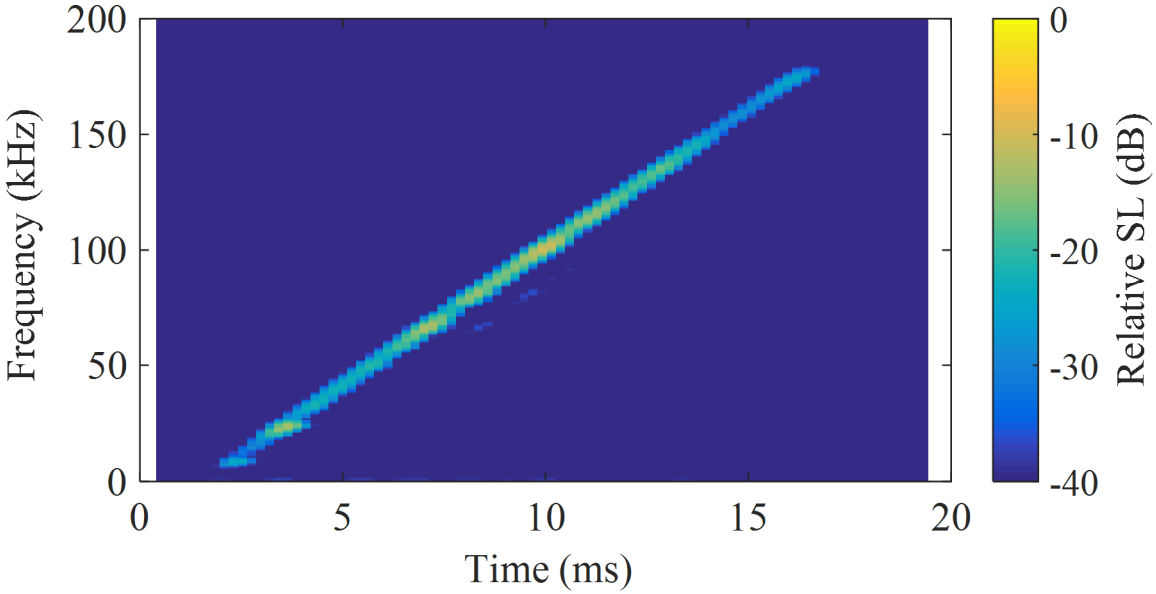


Fig. 4.11 Receiving chirp spectrogram at a transmission level of 20 mV (FFT size: 512 points; window: Hamming; overlap: 400 points).

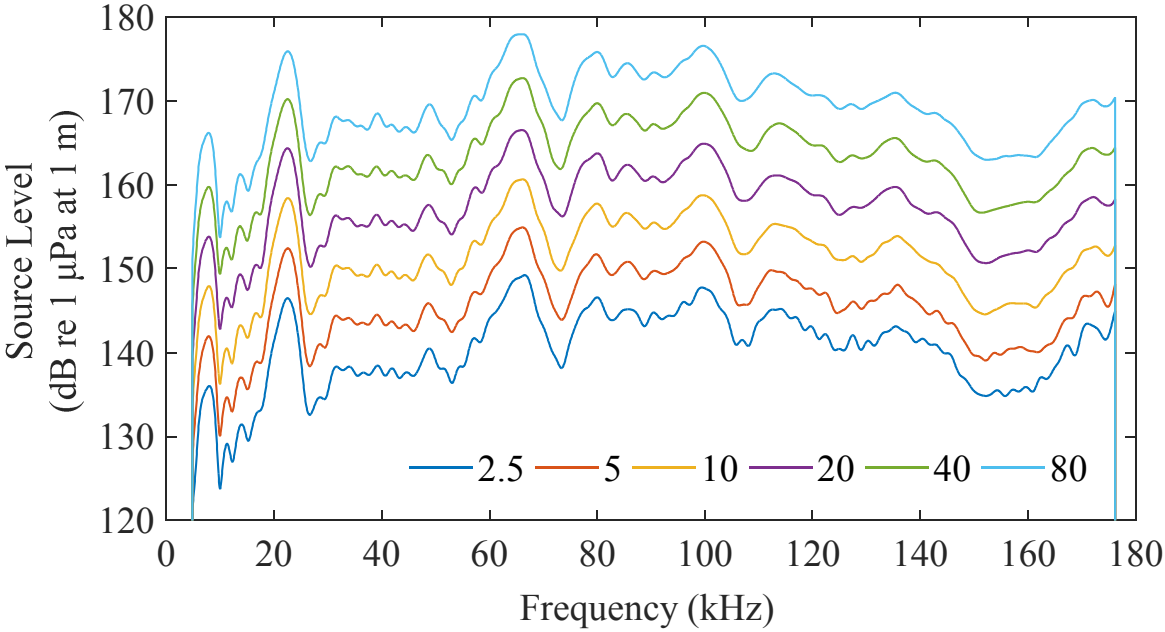


Fig. 4.12 SL spectra of the receiving chirp signals at the six transmission level from 2.5 mV to 80 mV.

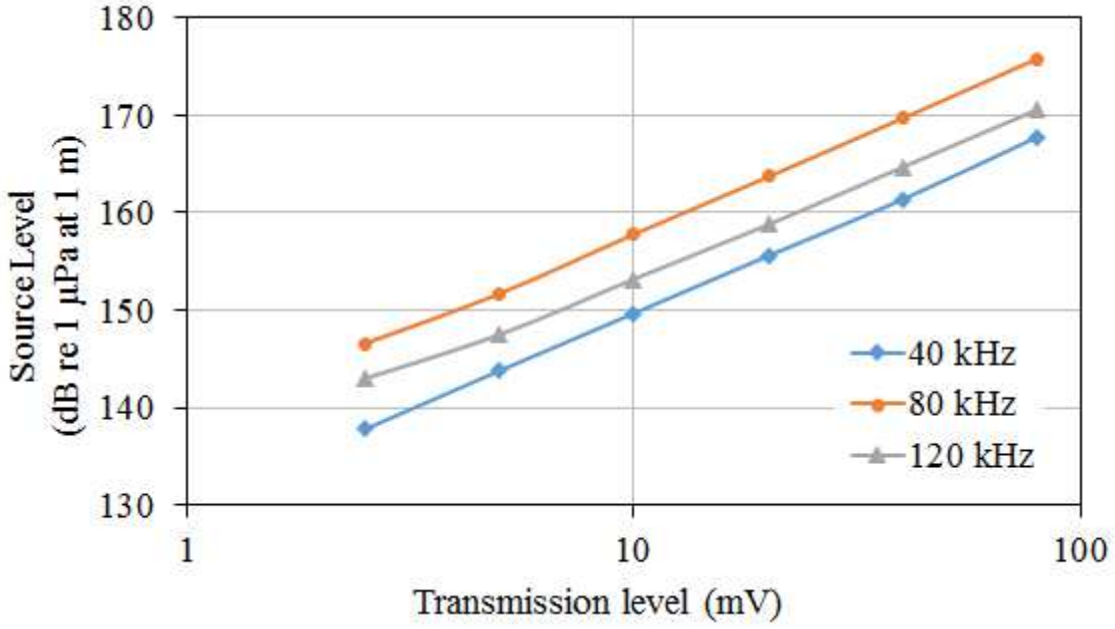


Fig. 4.13 Linear relationship between transmission level and SL at 40, 80, and 120 kHz.

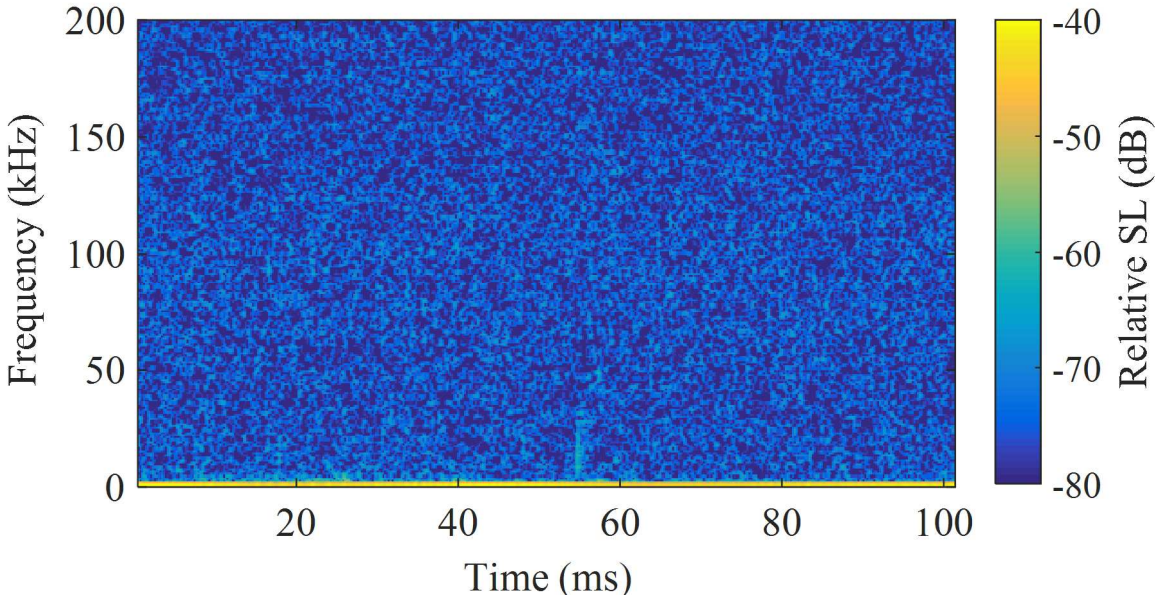


Fig. 4.14 Spectrogram of the silence track recording in which all transmitting equipment was operational but signal was not projected (FFT size: 512 points; window: Hamming; overlap: 400 points).

4.6 Calculation of the inverse characteristics

The broadband fidelity of reproduced signals was enhanced by cross-correlation between the transmitting signals and inverse characteristics of the transmitting system. The reciprocal of the amplitude spectrum and the inversion of the phase spectrum of the receiving chirp signal were calculated to compute the inverse spectral characteristics of the transmitting system. The spectral information was transformed by the inverse FFT to temporal information. Transmitting waveforms were cross-correlated to the waveform of the inverse characteristics before they were transmitted.

Fig 4.15 compared frequency properties between the receiving chirp signals before and after the cross-correlation with the inverse characteristics. The cross-correlated chirp signal shows a flat frequency response from 7 kHz to 175 kHz with ±5.5 dB ripple.

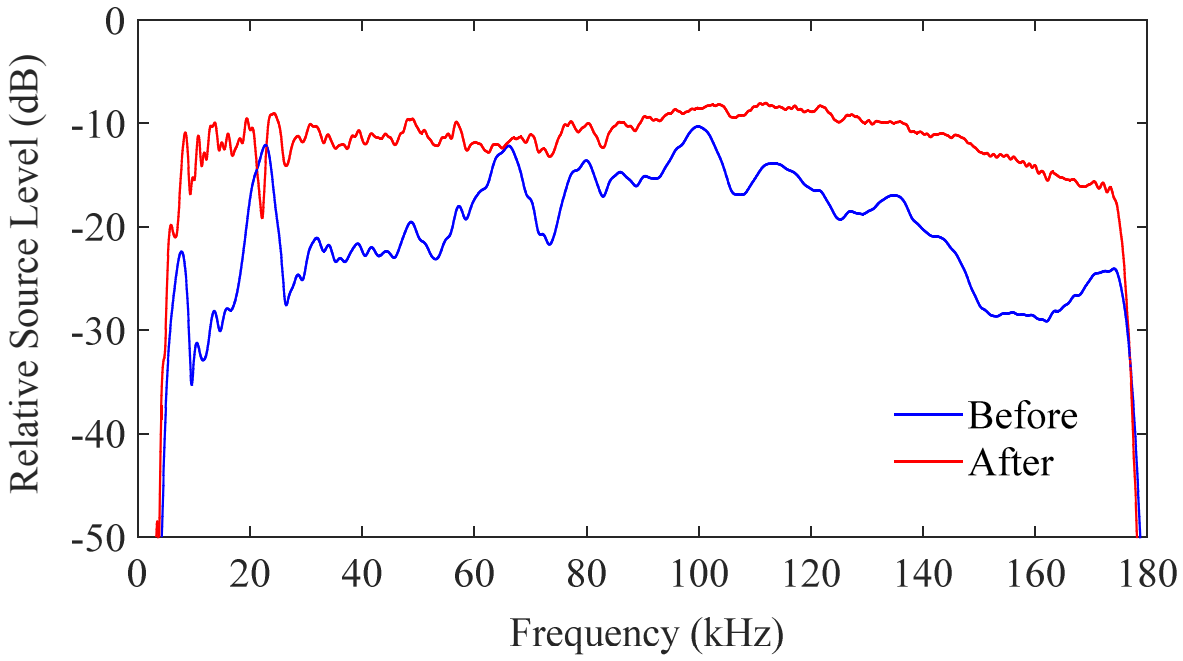


Fig. 4.15 Frequency properties of the receiving chirp signals before and after being cross-correlated to the inverse characteristics of the transmitting system.

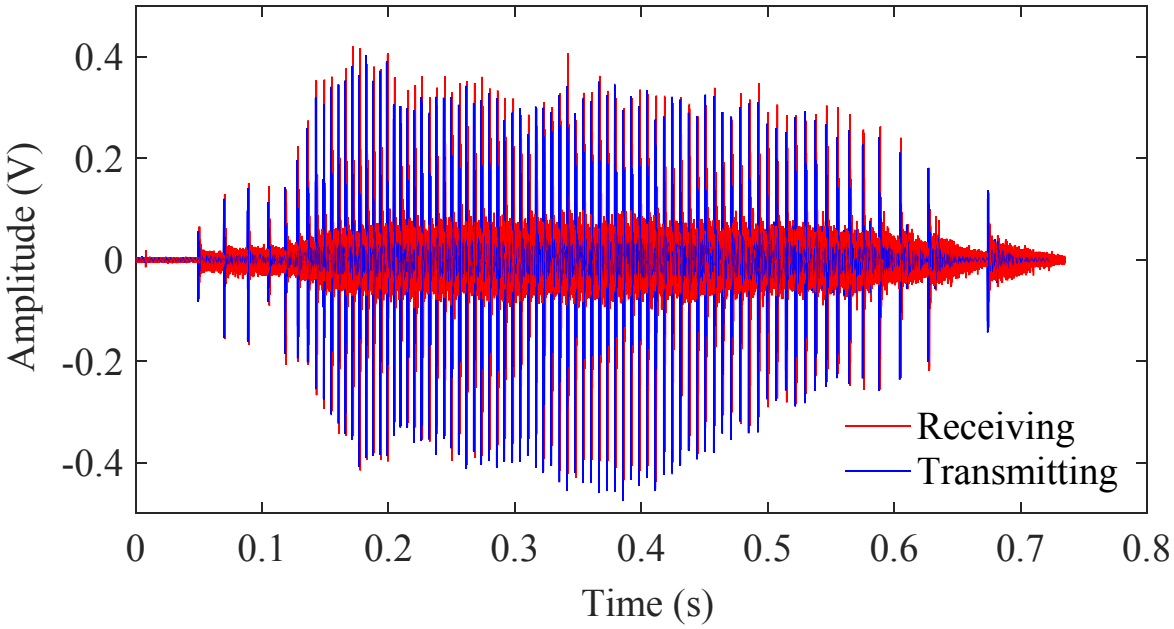
4.7 Test transmission of a PS1 sample

A prerecorded PS1 sample was cross-correlated to the inverse characteristics and then reproduced by the transmitting system to confirm whether the acoustic structure of the receiving PS1 resembled that of the transmitting PS1. The receiving system was the same as Fig. 4.9, excluding an Aquafeeler III preamplifier (AquaSound Inc., Japan) with a flat frequency response to 200 kHz (-3 dB), which bandpass-filtered the signals from 1 kHz to 200 kHz and amplified the signal by 40 dB. This experiment was implemented in an acoustic chamber of 5 m length, 4 m width, and 3 m height at Tokyo University of Marine Science and Technology.

The transmitting PS1 and the receiving PS1 were compared by waveforms, spectrograms and averaged power spectra. The middle pulses in the PS1 calls were used for calculation of the averaged power spectra. The waveform comparison indicated that the pulse train was faithfully reproduced in the temporal domain (Fig. 4.16 a). Configuration of each pulse was similar between the transmitting and receiving PS1 calls (Fig. 4.16 b). The spectrograms and spectra

also show that the frequency range of PS1 is almost realized (Fig. 4.17, 4.18). Reverberation seen in the receiving PS1 was probably due to multiple reflections from the walls, floor, and water surface of the small acoustic chamber.

a) All pulses composing PS1



b) One of the pulses composing PS1

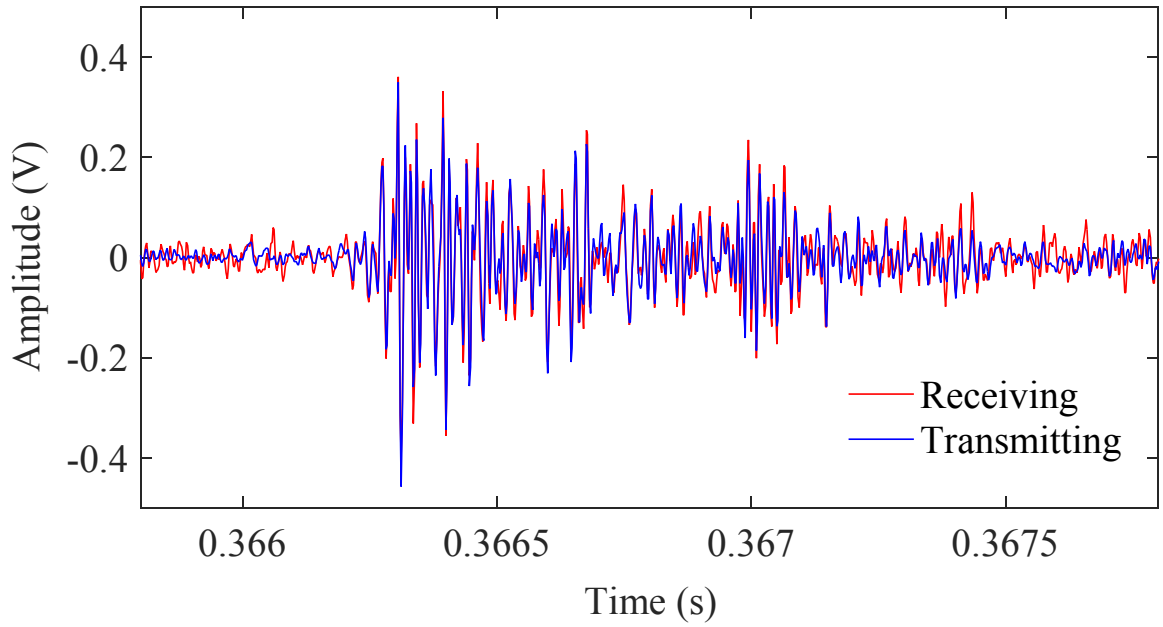
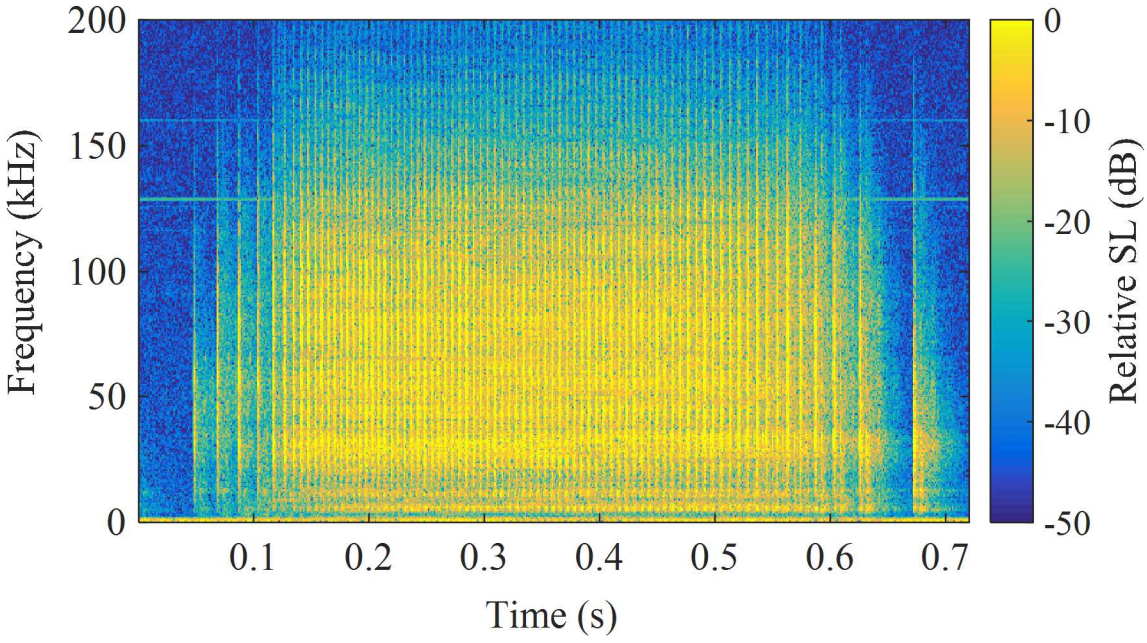


Fig. 4.16 Transmitting and receiving PS1 waveforms. a) All pulses and b) one of the pulses composing PS1.

a) Transmitting PS1



b) Receiving PS1

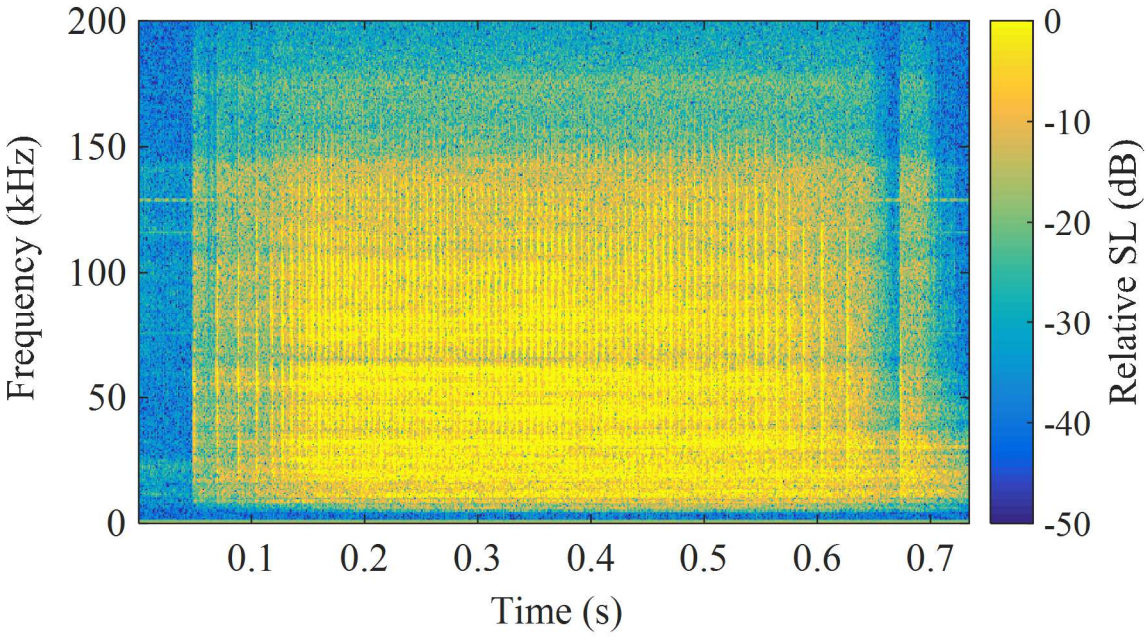


Fig. 4.17 a) Transmitting and b) receiving PS1 spectrograms (FFT size: 1024 points; window: Hamming; overlap: 400 points).

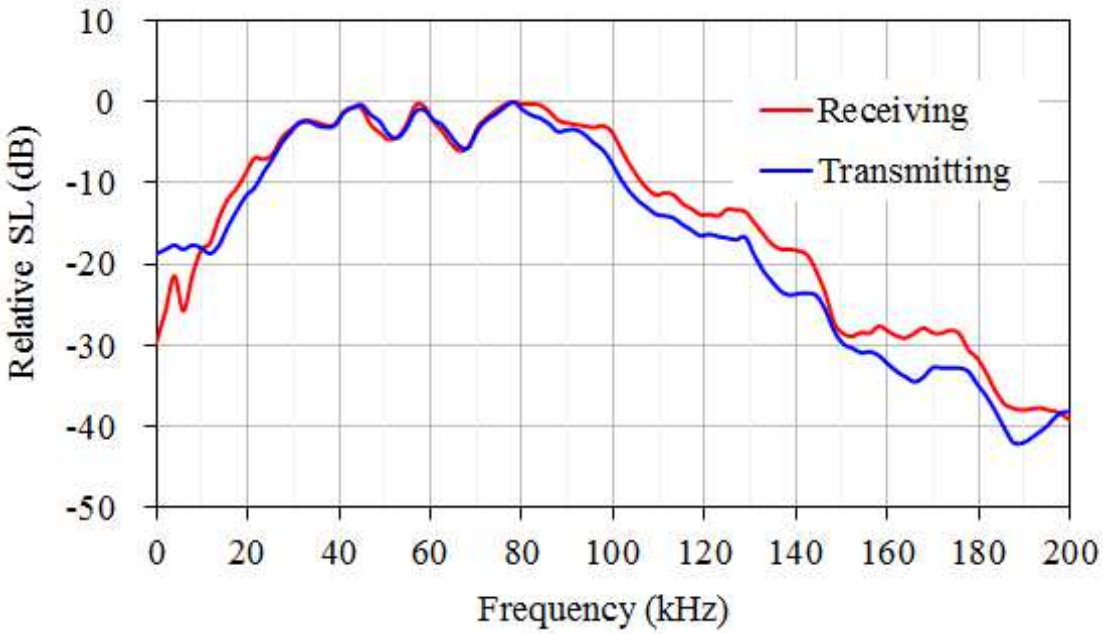


Fig. 4.18 Relative averaged power spectra of the middle pulses in the transmitting and receiving PS1 calls (FFT size: 256 points; window: Hamming; smoothing: 5 points).

4.8 Calibration of the level of transmitting PS1 data for desired source level

In playback experiments, several previously recorded PS1 calls were used as playback stimuli. Because the amplitude varies among the PS1 samples, it should be normalized and then calibrated according to desired SLs as follows:

1) Prerecorded PS1 samples were subjected to the cross-correlation process with inverse characteristics.

2) The maximum amplitude of the transmitting PS1 waveforms was set to 50 mV and they were reproduced by the transmitting system.

3) RMS amplitude of the receiving PS1 waveforms was measured and SLs were calculated. The receiving system was the same as section 4.7.

4) Based on the linear relationship between transmission level and SL (as described in subsection 4.5.2), the level of the transmitting PS1 data was calibrated according to a desired SL by using the computed SL at a transmission level of 50 mV.

4.9 Trial of PS1 playback experiment

4.9.1 Methods

A trial of the PS1 playback experiment was conducted in the Port of Nagoya Public Aquarium on October 29, 2015 to test whether belugas discriminate between PS1 calls from familiar and unfamiliar individuals. The main pool contained NF2, NF3, and NM2, and the holding pool contained NF1 and NM1 (Fig. 4.19). The medical pool was vacant and the broadband underwater speaker was set in the pool to hide its existence as far as possible from the belugas. Two TC 4013 underwater hydrophones (Reson Inc., Denmark) and six AQH-100DTP touch panel hydrophones (AquaSound Inc., Japan) were placed in the three sub-pools as in Fig. 4.19, to record vocal responses from all belugas. The hydrophones were connected to Aquafeeler III preamplifiers (AquaSound Inc., Japan) with a filter from 1 kHz to 200 kHz and a gain of 50 dB, and EZ7510 data recorders (NF Co., Japan) with a sampling rate of 500 kHz and 16 bits. The beluga behavioral response was recorded from three positions—the underwater windows of the main and holding pools and from above the surface of the main pool—using a GZ-V675-R video camera (JVC Co., Japan), an iVIS HF R11 video camera (Canon Inc., Japan), and a GoPro HERO 3+ video camera (GoPro Inc., USA), respectively. All equipment was arranged one day before the playback trial to accustom the belugas to it.

A familiar playback stimulus was selected from PS1 samples from NF2 and an unfamiliar playback stimulus was selected from PS1 samples from SF1 at Shimane Aquarium. NF2 and SF1 were both females and of almost equal age, 15 and 16–18 years old, respectively. The selected PS1 samples had good signal-to-noise ratio with no multivariate outliers based on the nine temporal and spectral parameters (Table 2.2). Each stimulus was played twice, with an interval of 2 s, which was long enough for a response (see subsection 3.4.3; Morisaka et al., 2013). The SL was set to 160 dB (RMS) re 1 μPa at 1 m, which was the typical SL of PS1 calls.

Vocal and behavioral recordings began in the afternoon. The operator of the transmitting system waited for the timing of playback at the medical pool, listening for vocalizations from the underwater hydrophones (Fig. 4.19). Two observers conducted video recordings at the main and holding pools and relayed the beluga behavior to the operator via transceivers. The operator played the prerecorded PS1 calls when there were no interactions, such as aggressive/submissive behavior, between individuals and no high vocal activities between belugas.

To compare the strength of the belugas' responses to familiar and unfamiliar playback stimuli, I investigated the frequency of PS1 production by all belugas and the staying time of NF1 and NM1 at the half of the holding pool closer to the lattice between the holding and medical pools. The individuals were considered to be staying in this area when the heads were visible in the area. The PS1 frequency and staying time were compared between 2 min before (pre-playback) and after (post-playback) each playback.

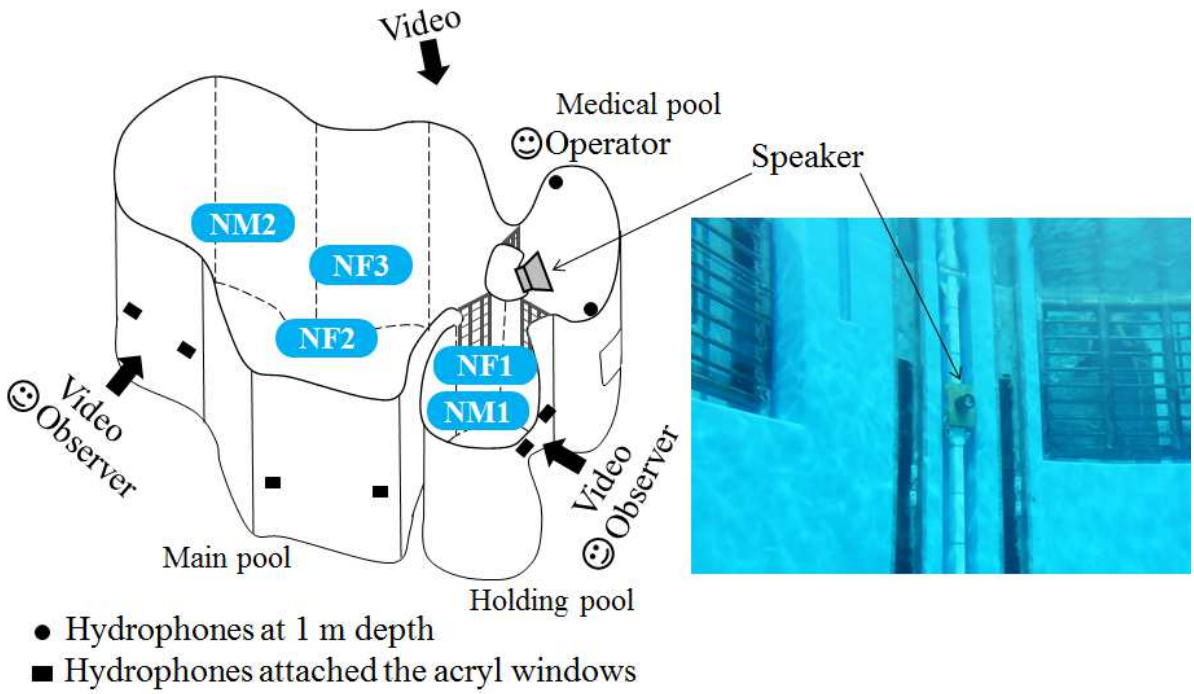


Fig. 4.19 Schematic layout of the playback experiment.

4.9.2 Results

There was one pulse train, other than PS1, in the pre-playback time of the familiar stimulus, and no vocalizations in the post-playback time. One PS1 call was recorded at the pre-playback time of the unfamiliar stimulus, whereas there were no vocalizations in the post-playback time. These results suggested that my playback attempt failed to elicit clear vocal responses.

The staying time of NF1 at the half of the holding pool closer to the speaker was 91 s and 101 s in the pre-playback and post-playback times of the familiar stimulus, respectively, and 86 s and 88 s in the pre-playback and post-playback times of the unfamiliar stimulus, respectively. The staying time of NM1 was 89 s and 110 s in the pre-playback and post-playback times of the familiar stimulus, respectively, and 69 s in the pre-playback time of the unfamiliar stimulus. The staying time of NM1 in the post-playback time of the unfamiliar stimulus could not be measured because of the dead angle of the video. These results suggest that the belugas stayed in the area close to the speaker after the familiar playback for a little longer time than before, while the staying time was comparable before and after the unfamiliar playback. Other belugas, especially the subadult NF3 and calf NM2, in the main pool escaped to the sides of the pool after both playbacks.

4.10 Discussion

Fidelity in transmitting systems is indispensable in eliciting correct responses from target animals. The established transmitting system almost fulfilled the requirements for PS1 playback. This study succeeded in creating a broadband underwater speaker by applying the broadband technique of a previously developed transducer. The transmitting system had flat frequency response (Fig. 4.12), and the broadband reproducibility was enhanced by cross-correlation between the transmitting signal and the inverse characteristics of the transmitting system, up to ±5.5 dB ripple from 7 kHz to 175 kHz (Fig. 4.15). Although the transmitting

system had low sensitivity below 7 kHz, it does not appear crucial as hearing sensitivity in belugas tends to drop sharply at frequencies lower than 8 kHz and higher than 120 kHz (Awbrey et al., 1988; Klishin et al., 2000; Mooney et al., 2008). However, it should be noted that low frequency sounds travel better than high frequency sounds, and belugas might then use the frequency range below 7 kHz especially in long-distance communication. The transmitting system could project sounds at least 130–170 dB re 1 μPa at 1 m and possesses enough power for PS1 playback (Fig. 4.12). The established calibration method calculates the PS1 transmission level in accordance with the desired SL. Mechanical noise and waveform distortion were not found in the transmitting system. When the PS1 test sample was reproduced by the transmitting system, the transmitting and receiving PS1 calls had almost equal acoustic structures (Fig. 4.16, 4.17, 4.18). These results led to the conclusion that the system was useful for PS1 playback experiments.

While peak-to-peak SLs of echolocation clicks often exceed 200 dB (Au et al., 1987; Au, 1993; Rasmussen et al., 2002; Au & Herzing, 2003; Madsen et al., 2004), SLs of communicative vocalizations tend to range from 120 dB to 180 dB (Janik, 2000a; Miller, 2006; Rasmussen et al., 2006; Clausen et al., 2010). Thus, the transmitting system is capable of reproducing various odontocete communicative calls with energy mainly in the frequency range of 7 kHz–175 kHz, as well as the PS1 calls of belugas. For instance, most whistles of the spinner dolphin (*Stenella longirostris*) and spotted dolphin (*Stenella frontalis*) have a fundamental frequency above 10 kHz and harmonics extending past 50 kHz (Lammers et al., 2003). The harmonic cues are possibly used to indicate the orientation of callers since the high-frequency harmonic content varies depending on azimuth. Their burst pulses were ultrasonic with little or no energy below 20 kHz, and might have communicative function in social contexts such as agonistic interactions (Lammers et al., 2003). In bottlenose dolphins, broadband burst pulses spanning approximately from 5 kHz to 150 kHz served as

aggressive/agonistic calls (Blomqvist & Amundin, 2004), and play-fighting dolphins also produced similar burst pulses though followed by whistles (Blomqvist et al., 2005). *Cephalorhynchus* species appear to use narrow-band high-frequency pulse trains for communication (Watkins et al., 1977; Dawson, 1991; Yoshida et al., 2014). Those communicative call functions will be verified experimentally by playback experiments.

Further, the transmitting system is useful in playback studies of the investigation of vocal mimicry and matching skills. Richards et al. (1984) demonstrated the imitation of computer generated tonal sounds by a bottlenose dolphin. Similarly, Murayama et al. (2014) presented the imitation of synthetic tonal sounds by a beluga. Vergara (2011) examined whether belugas could respond to playbacks of two categories of their calls, which were produced after hand signals during public shows, with matching vocalizations. The broadband transmitting system will broaden this study area.

The broadband transmitting system is also helpful in other areas of study, including wildlife management. Marine mammals display avoidance behavior when they were presented with calls of killer whales, a potential predator (Cummings & Thompson, 1971; Deecke et al., 2002), and the playbacks of killer whale calls were used to prevent whales and pinnipeds from interfering in fishing industry (Fish & Vania 1971; Shaughnessy et al., 1981). As some killer whale calls have energy up to 100 kHz (Schevill & Watkins, 1966), broadband playback using the broadband transmitting system may thus elicit the targets' response more effectively.

Codas of socializing sperm whales were broadcasted to guide conspecifics in a narrow confined bay to safe open sea (Goold, 1999), and although the operation was unsuccessful, one whale approached the sound source in a single playback trial. It is suggested that playback of contact calls may be useful for controlling the direction and movement of odontocetes in danger of stranding. The broadband transmitting system enables us to try similar playback on various species to protect them from stranding.

I attempted a playback experiment of PS1 calls using the transmitting system to test whether belugas discriminate PS1 calls of familiar and unfamiliar individuals. The attempt did not elicit clear vocal responses to both of familiar and unfamiliar playback stimuli. The staying time of the two adults in the half of the holding pool closer to the speaker was comparable before and after the playback of the unfamiliar stimulus. However, after the familiar stimulus, the adults showed a slight tendency for a longer staying time in the half of the pool closer to the speaker. This suggests that the response of the belugas to the familiar playback was a little stronger than that to the unfamiliar playback. However, other belugas in the main pool escaped from the speaker after both playbacks. From these ambiguous results, it was apparent that I should redesign the playback protocol. According to my observation and the information from trainers, belugas are curious but sensitive and cowardly to new situations when exposed to new objects or sounds. When the playback experiment was conducted, the belugas might have become more sensitive because they experienced an abnormal situation when I prepared the playback experiment, such as the installation of the speaker, the day before the experiment. Although the speaker was set so that the belugas could not see it, they knew there was something in the medical pool, even if they did not know what it was. Thus, when sounds were emitted by the speaker, the adult belugas in the holding pool could show increased curiosity, whereas the subadult, calf, and adult belugas defending the calf in the main pool might escape from the unknown entity, even if the exposed sound was a PS1 of their associate (or a PS1 of herself for NF2, the adult beluga depending the calf). Therefore, I need to design a playback protocol that can be performed under more natural conditions to assess a response more reflective of natural behavior from belugas.

In summary, a broadband transmitting system was established enabling us to conduct playback experiments of various broadband calls in odontocetes and PS1 calls in belugas. Since my first attempt of PS1 playback experiment failed to elicit clear vocal and behavioral

responses, in future research, I need to redesign the playback protocol.

Chapter 5: Summary and discussion

5.1 PS1 characteristics in belugas

The results from two aquariums are summarized as follows:

Acoustic structure of PS1

PS1 sounds like a ratchet or a door creaking to human ears.

Np: 103.5 ± 42.5 (19–260)

DUR: 0.58 ± 0.3 (0.16–1.48) s

PRR: 197.6 ± 80.9 (32.6–525.5) pulses/s

IPI 1: 4.1 ± 2.1 (1.0–19.6) ms

IPI 2: 5.9 ± 1.6 (4.1–13.1) ms

F_p: 58.8 ± 36.0 (11.7–146.5) kHz

10 BW: 58.4 ± 27.2 (0.0–136.7) kHz

F_l: 31.3 ± 31.1 (2.0–109.4) kHz

F_u: 89.6 ± 37.9 (11.7–175.8) kHz

Some of individual PS1 calls contain tonal or narrowband pulsed components overlapping the main pulse train.

PS1 function

1) PS1 was the most predominant call type in an isolation context.

2) There was no relationship between PS1 bouts and aggressive/submissive behavior.

3) PS1 was used for vocal exchange.

The results of 1), 2), and 3) suggest that PS1 play a role in a contact call.

4) The largest and strongest belugas, which were thought to be top rank, produced PS1 most frequently. This suggests that dominant belugas may produce PS1 to instigate and control the movement of the group members.

5) There was an increase in PS1 production rates in a visual reunion after a long separation,

and those PS1 calls were individually distinctive. This implies that PS1 possibly functions as an individual advertisement.

Individuality in PS1

1) IPI contours showed strong individuality. Visual comparison revealed the individual distinctiveness in IPI contours, and it was supported by the results of statistical analyses on temporal parameters.

2) Spectra had little individuality. Although visual comparison revealed no clear individual distinctiveness in spectra, statistical analyses showed some spectral parameters were individually different.

PS1 variants

PS1 was subcategorized into acoustically different variants in males based on IPI contours. It implies the possibility of PS1 sharing or convergence and other PS1 roles in males.

5.2 PS1-like calls described in previous studies of Monodontidae

Rrelatively long broadband pulse trains were described in previous studies of both captive and wild beluga vocalizations. It should be noted that those studies only recorded up to 24 kHz or lower, therefore frequency components were not compared directly in the present study. The PS1-like calls were summarized as follows, according to Mishima et al. (2015):

Sjare and Smith (1986a) reported vocalizations of wild belugas in the Northwest Territories, Canada. The pulse train type H they categorized in group 3 calls is similar to PS1 calls in terms of spectrograms and PRR of 80–290 pulses/s. The duration of the group 3 calls are also similar to PS1 calls, with 0.85 ± 0.44 (0.2–2.7) s. The frequency composition of the group 3 calls is 4.6 ± 1.7 (0.3–12.0) kHz, which is comparable to PS1 calls in the Port of the Nagoya Public Aquarium that have a peak at 6 kHz in the range below 10 kHz. The group 3 calls were produced during rest and socially interactive periods (Sjare & Smith, 1986b).

Bel'kovitch & Sh'ekotov (1993) summarized vocalizations of wild belugas in the White Sea and Amur Estuary, Russia. The intensive pulse trains, "grinding" calls, are similar to PS1 calls with high PRR up to 150 pulses/s of 0.9–1.92 s in duration and dominant frequencies of 1.2–2.2 kHz and 5.6–10.0 kHz. The calls were calling signals produced by a female to a juvenile. Belikov and Bel'kovich (2008) also examined wild belugas in the White Sea, Russia. The pulse train types with low PRR, lPT3 and lPT7, resembled PS1 calls. The IPT3 had 13–630 pulses/s, 0.87 ± 0.43 (0.32–2.28) s in duration, and a dominant frequency of 6.1 ± 1.0 (3.9–8.7) kHz, while the IPT7 had 9–770 pulses/s, 1.36 ± 0.43 (0.6–2.43) s in duration, and a dominant frequency of 5.7 ± 4.8 (0.2–15.0) kHz. These calls were produced during social interactions and quiet swimming (Panova et al., 2012). Alekseeva et al. (2013) reported that pulse trains with low PRR, "groaning" and "grumbling", which are similar to PS1 calls, were produced in the context of sexual behavior.

Karlsen et al. (2002) investigated vocalizations of wild belugas in Svalbard, Norway. The pulse train type II is similar to PS1 calls, with PRR of 104 ± 64 (23–240) pulses/s, 0.55 ± 0.54 (0.07–3.12) s in duration, and frequency range of 0.2–20.0 kHz. The type II calls were produced in the context of milling, travelling, and joining. van Parijs et al. (2003) recorded calls from wild belugas in the same Svalbard area during temporal capture events. In their recordings, the mother of a mother-calf pair produced pulse trains with an average of 27 pulses/s and 1.9 ± 1.3 s in duration. She frequently moved her head toward her calf while producing sounds. The pulse trains from the calf had an average of 18 pulses/s, 0.6 ± 0.5 s in duration, and occasionally had an overlapping tonal component. Another sub-adult female that was temporarily captured also produced pulse trains with an average of 22 pulses/s and 0.3 ± 0.08 s in duration. These calls have a smaller PRR than PS1 calls but are similar in duration.

Chemelnitsky & Ferguson (2012) represented vocalizations of wild belugas in the Churchill

River, Canada. The pulse train type P2 is similar to PS1 calls spectrographically. P2 has a PRR of 207 ± 57 pulses/s, 1.16 ± 0.36 s in duration, and frequency range of 2.8–5.3 kHz. Recchia (1994) reported on captive belugas originating from the river. One of the most discriminant call types, the so-called "buzzsaw" calls, appears to be spectrographically similar to PS1 calls with a minimum duration of 0.2 s, but the information available is limited. Vancouver Aquarium belugas originating from the Churchill River Estuary produced Type A calls in isolation contexts (Vargara et al., 2010). A mother predominantly produced Type A calls the day after the birth of two calves and the death of a calf on a different occasion, as well as whenever she needed to regain or maintain contact with her calf. In addition, there were vocal exchanges of Type A calls between the mother-calf pair. Type A call variants, A1–A5, are broadband rapid pulse trains 1.2–1.9 s in duration and resemble PS1 calls. Average PRRs are 94.6 ± 13.0 pulses/s for A1, 328.9 ± 36.4 pulses/s for A2, 306.4 ± 42.4 pulses/s for A3, 115.0 ± 26.1 pulses/s for A4, and 371.8 ± 40.3 pulses/s for A5. A1 contains a narrowband tonal component consistently at 14.6 + 0.6 kHz overlapping the pulse train. A3 also contains a secondary pulsed synchronous component. These overlapping structures are found in PS1 calls. Vargara et al. (2010) additionally made recordings from temporarily restrained belugas in the Nelson River Estuary and social groups in the St. Lawrence Estuary, Canada. In both areas, the Type A calls were observed.

Moreover, PS1-like calls are found in narwhals (*Monodon monoceros*), which are the belugas' closest living relative, also belonging to the Monodontidae family, living in arctic oceans, and migrating long distances (Heide-Jørgensen et al., 2003). They form large groups composed of small, sexually segregated subgroups (Marcoux et al., 2009) and the subgroups might display a fission-fusion type social structure, although the details remain unclear (Watt et al., 2015). Ford & Fisher (1978) collected narwhal vocalizations in Koluktoo Bay, Canada. The pulsed tones are similar to PS1 calls spectrographically and have highly variable PRRs and of 0.6–

1.3 s in duration. Further, they were repeated a number of times in succession, and each successive pulsed tone was nearly identical in structure. This implied that individual narwhals produced each type of pulsed tone and the succession was vocal exchange. Marcoux et al. (2012) recorded calls from narwhals in the same Koluktoo Bay area. The pulse trains resemble PS1 calls spectrographically. They last 0.12 ± 0.07 (0.02–0.43) s and include 77.4 ± 44.2 (10–279) pulses. When the pulse trains were classified based on the pulse-rate contours, there was no association between pulse train and behavioral categories but there was an association between the pulse train category and the herd. Thus, these pulse trains may serve as individual- or group- specific contact calls.

Shapiro (2006) investigated vocalizations of two male narwhals belonging to different social groups in Admiralty Inlet, Canada. The “combined tonal/pulsed signals” resemble PS1 calls. One of the males had a PRR of 82.3 ± 14.2 (28.1–112.8) pulses/s and 1.6 ± 0.7 (0.6–2.7) s in duration, while the other male had a PRR of 160.8 ± 7.4 (147.5–180.5) pulses/s and 1.2 ± 0.1 (1.0–1.2) s in duration. The calls also contain synchronous tonal components with consistent energy. In addition, the pulse-rate contours of the pulsed components were clearly different between the two males, thus they may use the combined tonal/pulsed signals as individualized contact calls and encode individual information into their pulse-rate contours.

Subsequent recordings of PS1 calls from captive and wild belugas were essential to define PS1. Many samples will interpret the role of the tonal or pulsed components that co-occurred with the main pulse train in PS1 and the role of PS1 variants in males. Further investigation of contact calls in narwhals is also required to understand the evolutionary pathway of contact calls in Monodontidae.

5.3 Evolutionary processes and adaptive significance of contact calls in odontocetes

5.3.1 Pulse- and whistle-type contact calls

The common ancestor of odontocetes used pulses, while whistles emerged after the divergence of Platanistidae, 30 million years ago (Morisaka & Connor, 2007) (Fig. 5.1). However, Pontoporiidae and Phocoenidae families, genus *Cephalorhynchus*, hourglass dolphin (*Lagenorhynchus cruciger*), and Peale's dolphin (*Lagenorhynchus australis*) lost whistles, and eliminated the lower-frequency component of the pulses, resulting in what are called narrow-band high-frequency (NBHF) pulses (Morisaka, 2012). Moriska (2012) established a hypothesis that whistles originally evolved for sexual selection because the first whistling species showed sexual dimorphism, suggesting strong sexual selection. The whistle function was then diverted into group cohesion in Delphinidae, because low-frequency whistle sounds transmit effectively in their environment. The whistle loss and NBHF pulse emergence was selected as a result of the predation pressure from killer whales, which could not hear high-frequency sounds above 100 kHz (Morisaka & Connor, 2007). In non-whistling NBHF species, the cost of producing whistles, or detection by killer whales, appeared to exceed its benefit, or long-distance communication. Northern right whale dolphins (*Lissodelphis borealis*), Pacific white-sided dolphins (*Lagenorhynchus obliquidents*), and dusky dolphins (*Lagenorhynchus obscurus*) may occasionally produce whistles, but mostly produce only broadband pulses (Rankin et al., 2007; Henderson et al., 2011; Vaughn-Hirshorn et al., 2012); therefore, they seem to be in the process of evolution towards becoming non-whistling NBHF species against killer whale predation risk.

The acoustic media of group cohesion was originally pulses as seen in sperm whales (see subsection 1.1.2). Platanistidae, Zipiidae, and Inioidea are solitary or live in small groups, and little is known regarding whether they have contact calls. However, the existence of pulse-type contact calls has also been found in Monodontidae—belugas, and perhaps narwhals (Ford & Fisher, 1978; Shapiro, 2006; Marcoux et al., 2012). Moreover, an isolated harbor porpoise (*Phocoena phocoena*) calf, which belongs to Phocoenidae, produced a certain type of pulse

train toward her mother (Clausen et al., 2010); therefore, the pulse train presumably serves as a contact call. Further, pulse-type contact calls are found in killer whales, the basal species of Delphinidae (see subsection 1.1.2).

After the divergence of killer whales, the acoustic media of contact calls diverted from pulses to whistles. The overlapping tonal component as seen in some of the pulse-type contact calls in belugas, narwhals, and killer whales may be the emergence of whistle-type contact calls. Bottlenose dolphins evolved signature whistles (see subsection 1.1.3), and they are possibly used in Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) (Gridley et al., 2014), common dolphins (Caldwell & Caldwell, 1968), Atlantic spotted dolphins (*Stenella plagiodon*) (Caldwell et al., 1973), Pacific humpback dolphins (*Sousa chinensis*) (van Parijs & Corkeron, 2001), and Guiana dolphins (*Sotalia guianensis*) (de Figueiredo & Simão, 2009).

Non-whistling species in Delphinidae could use pulses as contact calls alternatively. Repeated burst pulse patterns were found in northern right whale dolphins (Rankin et al., 2007), dusky dolphins (Vaughn-Hirshorn et al., 2012), Pacific white-sided dolphins (Henderson et al., 2011; my unpublished data), which are in the course of evolution towards becoming non-whistling NBHF species. These repeated burst pulse patterns may function as contact calls in these species. It was suggested that non-whistling NBHF species of the genus *Cephalorhynchus* also use pulses for communication (Watkins et al., 1977; Dawson, 1991; Yoshida et al., 2014).

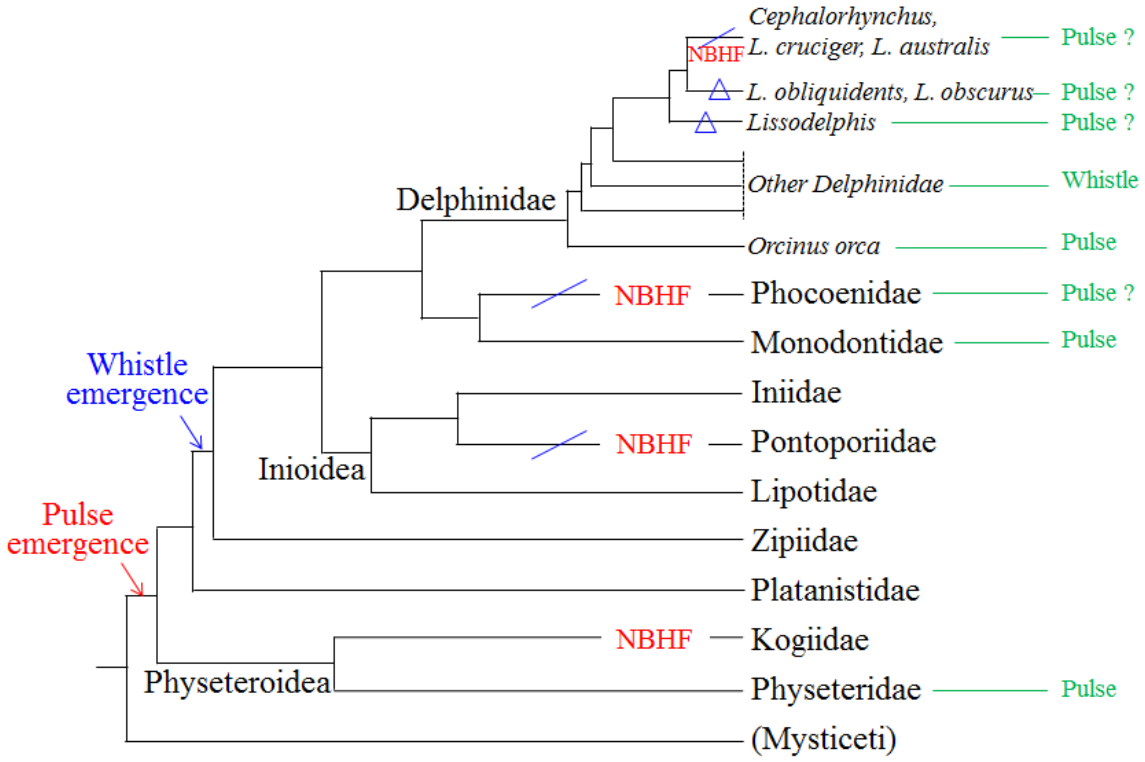


Fig. 5.1 Phylogenetic relationships among odontocetes and evolution of their vocalizations arranged from Morisaka (2012). The diagonal blue line represents whistle loss. The red NBHF indicates narrowband high-frequency pulses. The call type (pulse or whistle) used for contact calls are represented in green.

5.3.2 Species specificity in contact calls

Contact call characteristics are specific to families and/or species and the specificity is the result of the differences in phylogeny, as well as ecology and morphology. Sperm whales use relatively low-frequency, and long-spaced click series for contact calls. Their large body size appears to cause low-frequency characteristics. The IPIs of pulses comprising each click are related to body size, and the individually different IPIs are possibly used for individual recognition (Gordon, 1991; see subsection 1.1.2). The IPIs are a few milliseconds, thus inter-click intervals of the click trains are longer than the IPIs to distinguish between individual and group information.

However, the contact calls of Monodontidae appear to be broadband and short-spaced pulse

trains with or without overlapping tonal or pulsed components. Belugas and narwhals live in pack ice or polynyas in winter, which is a highly noisy and reverberant environment (Brown & Milne, 1967). Thus, they may sharpen the directivity of the broadband contact calls in the wintering area to minimize masking by clutter or ambient noise. Broadband pulses are useful communicative signals to maximize transmission effectiveness according to the surrounding environment by adjusting the directivity. The pulse-type contact calls have short IPIs, possibly to encode individuality in the pulse repetition pattern, because individual distinctiveness in the pulse repetition pattern was reported in both species (see section 5.1 and 5.2). In addition, the overlapping components perhaps play a role in the enhancement of individual recognition in their noisy winter environment.

Harbor porpoises in Phocoenidae appear to use similar spaced pulse trains to Monodontidae, for contact, but use NBHF pulses. According to Morisaka & Connor (2007), they are vulnerable to killer whale predation because of their smaller body size, and the region of their inshore distribution is characterized by a high abundance of killer whales. Moreover, given their body size and distribution, grouping would not be an effective anti-predator strategy, as they are found in small groups. Thus, they selected effective NBHF pulses to prevent killer whales from eavesdropping on their calls.

Killer whale contact calls are extremely short spaced pulsed tones with or without simultaneous, tonal components. Their contact calls seem to be used for group cohesion within not only matrilineal units but also pods; thus, they produce biphonic calls to increase transmission range (Miller et al., 2006; Filatova et al., 2009; see subsection 1.1.2).

Many Delphinidae species, including bottlenose dolphins, and perhaps other species, produce signature whistles. They need long-distance contact calls because of their large group sizes, and use low-frequency whistles as contact calls. Species specificity possibly exists in the acoustic characteristics of signature whistles, especially maximum frequency (Steiner, 1981),

but further investigations are required.

I speculated that northern right whale dolphins, dusky dolphins, and Pacific white-sided dolphins, which are taxonomically close to each other (May-Collado & Agnarsson, 2006), perhaps use sequence-structured, broadband burst pulse patterns as contact calls instead of whistles. Their groups range from several individuals to thousands of individuals, and the group size depends on behavioral context or season (Jefferson et al., 1994; Würsig & Würsig, 1980; Degrati et al., 2008; Morton, 2000; Henderson et al., 2011). The huge group size may need complex, sequence-structured call patterns to suppress interference.

5.3.3 Individuality in contact calls

The degree of individuality in contact calls is linked to social complexity rather than habitat or phylogeny. The societies of sperm whales and killer whales are stable for a long time; therefore, group identity is a greater priority than individual identity. A small degree of individuality, which is recognizable by only the group members, is sufficient (see subsection 1.1.2).

However, bottlenose dolphins live in a fluid and complicated society. Therefore, they needed to evolve not only individual signatures independent of voice cues but also the use of these signatures to address conspecifics, akin to the roles that names play in human social interactions (see subsection 1.1.3). The closest mechanism to their signature has only been found in the contact calls of parrots. Orange-fronted conures (*Aratinga canicularis*) have individually distinctive frequency modulation patterns in contact calls called “chees” (Cortopassi & Bradbury, 2006), and imitate the chees of close associates to address particular individuals in fission-fusion flocks (Balsby & Bradbury, 2009; Balsby et al., 2012). Spectacled parrotlets (*Forpus conspicillatus*) also use contact calls to label conspecifics (Wanker et al., 2005).

Belugas might need to evolve strongly recognizable individuality in contact calls because of

their high mobility and long-term associations in a fluid society. Although slight differences in spectral cues of PS1 might be a by-product of differences in vocal tract morphology and body size, as seen in several species (Boughman & Moss, 2003), obvious individual pattern in IPI contours appeared to specially evolve for individual recognition. In addition, given that a six-month-old calf did not produce PS1 and a 21-month-old calf produced PS1 with fluctuations, it is possible that beluga calves acquire their own IPI contours similarly to bottlenose dolphins. This is also supported by their potential production learning during vocal development (Vergara & Barrett-Lennard., 2008), long association with their mothers (Krasnova et al., 2014), and social interaction in their fluid social structure (Bel'kovitch & Sh'ekotov, 1993; Michaud, 2005; Colbeck et al., 2013). However, it is currently unclear whether IPI contours convey individual signatures independent of voice cues as in the signature whistles of bottlenose dolphins. Moreover, there is a possibility that the overlapping tonal or pulsed component was necessary for individual advertisement to enhance individual recognition. Playback experiments will elucidate them.

Some studies reported the ability of belugas to mimic human speech and synthetic sounds (Ridgway et al., 2012; Murayama et al., 2014), and have demonstrated their object labeling skills using sounds (Murayama et al., 2012). However, apparent copying of individually distinctive IPI contours was not found in either aquarium, although males shared one type of PS1 variant. This suggests that there is little probability that belugas address a particular individual by copying its PS1 as seen in bottlenose dolphins, even if IPI contours served as signatures. Therefore, bottlenose dolphins seem to facilitate a more sophisticated individual recognition system than belugas.

Further investigation regarding beluga society will help us to understand the difference in individual recognition mechanisms of belugas and bottlenose dolphins. In addition, we need to explore individuality in contact calls and the social structure of various species to reveal the

evolutionary process and adaptive meaning of all forms of acoustically embedded individuality.

5.4 Considerations for future PS1 playback experiments in belugas

The established transmitting system in chapter 4 is capable of reproducing broadband PS1 calls faithfully. Therefore, as a next step, playback experiments should be performed to experimentally prove whether belugas discriminate between PS1 calls from unfamiliar and familiar individuals and between familiar associates, and whether IPI contours carry sufficient identity information independent of voice features.

Playback experiments are difficult because there are few playback chances to avoid animals from becoming accustomed to acoustic stimuli. In addition, inappropriate selection of target animals, playback stimuli, composition within a playback sequence, and behavioral context generate the possibility of pseudoreplication (Deecke, 2006). Further, a proper response index should be developed to quantify responses. Therefore, the design of the playback protocol is important to elicit apparent responses from target animals with a small number of trials.

My attempt of a PS1 playback experiment failed to elicit clear vocal and behavioral responses in belugas, possibly due to the abnormal experimental situation. It could be easier to realize playback experiments under natural conditions in wild than in constrained captivity. In bottlenose dolphins, several playback experiments have been conducted both in the wild and in captivity (Lang & Smith, 1965; Dreher, 1966; Caldwell et al., 1972; Sayigh et al., 1999; Janik et al., 2006; Nakahara & Miyazaki, 2011; King & Janik, 2013; Bruck, 2013). However, these protocols may not apply directly to belugas because the two species have different characters: both species are curious but belugas are timid. Therefore, I need to consider this when planning PS1 playback experiments to assess responses more reflective of natural behavior from belugas. Controlled procedures such as habituation-dishabituation methods may also be useful.

References

Alekseeva YI, Panova EM, Bel'kovich VM. 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). Biol Bull. 2013;40:307–17.

Amakasu K, Mishima Y, Sasakura T, Mukai T, Sawada K. Application of multilayer piezoelectric actuators to broadband backscattering measurements of aquatic animals. J Marine Acoust Soc Jpn. 2013;40:126–37.

Antunes R, Schulz T, Gero S, Whitehead H, Gordon J, Rendell L. Individually distinctive acoustic features in sperm whale codas. Anim Behav. 2011;81:723–30.

Arnold BD, Wilkinson GS. Individual specific contact calls of pallid bats (*Antrozous pallidus*) attract conspecifics at roosting sites. Behav Ecol Sociobiol. 2011;65:1581–93.

Au WWL. The Sonar of Dolphins. New York: Springer Verlag; 1993.

Au WWL, Herzing DL. Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*). J Acoust Soc Am. 2003;113:598–604.

Au WWL, Penner RH, Turl CW. Propagation of beluga echolocation signals. J Acoust Soc Am. 1987;82:807–13.

Aubin T, Jouventin P, Hildebrand C. Penguins use the two-voice system to recognize each other. Proc R Soc Lond B: Biol Sci. 2000;267:1081–7.

Awbrey FT, Thomas JA, Kastelein RA. Low-frequency underwater hearing sensitivity in belugas *Delphinapterus leucas*. J Acoust Soc Am. 1988;84:2273–5.

Baird RW. The killer whale: foraging specializations and group hunting. In: Mann J, Connor RC, Tyack P, Whitehead H, editors. Cetacean societies: field studies of dolphins and whales. Chicago: University of Chicago; 2000. p. 127–53.

Balsby TJS, Bradbury JW. Vocal matching by orange-fronted conures (*Aratinga canicularis*). Behav Process. 2009;82:133–9.

Balsby TJS, Momberg JV, Dabelsteen T. Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network. PLoS ONE. 2012;7:e49747.

Belikov RA, Bel'kovich VM. High-pitched tonal signals of beluga whales (*Delphinapterus leucas*) in a summer assemblage off Solovetskii Island in the White Sea. Acoust Phys. 2006;52:125–31.

Belikov RA, Bel'kovich VM. Whistles of beluga whales in the reproductive gathering off Solovetskii Island in the White Sea. Acoust Phys. 2007;53:528–34.

Belikov RA, Bel'kovich VM. Communicative pulsed signals of beluga whales in the reproductive gathering off Solovetskii Island in the White Sea. Acoust Phys. 2008;54:115–23.

Bel'kovich VM, Sh'ekotov MN. The belukha whale: natural behaviour and bioacoustics. Woods Hole: Woods Hole Oceanographic Institution; 1993.

Blomqvist C, Amundin M. High-frequency burst-pulse sounds in agonistic/aggressive interactions in bottlenose dolphins (*Tursiops truncatus*). In: Thomas J, Moss C, Vater M, editors. Echolocation in Bats and Dolphins. Chicago: University of Chicago; 2004. p. 425–31.

Blomqvist C, Mello I, Amundin M. An acoustic play-fight signal in bottlenose dolphins (*Tursiops truncatus*) in human care. Aquat Mamm. 2005;31:187–94.

Boughman JW, Moss CF. Social sounds: vocal learning and development of mammal and bird calls. In: Simmons AM, Popper AN, Fay RR, editors. Acoustic Communication. New York: Springer Verlag; 2003. p. 138–224.

Brodie P, Ramirez K, Haulena M. Growth and maturity of belugas (*Delphinapterus leucas*) in Cumberland Sound, Canada, and in captivity: evidence for two growth layer groups (GLGs) per year in teeth. J Cetacean Res Manage. 2013;13:1–18.

Brown JR, Milne AR. Reverberation under Arctic Sea-Ice. J Acoust Soc Am. 1967;42:78–82.

Bruck JN. Decades-long social memory in bottlenose dolphins. Proc R Soc Lond B: Biol Sci. 2013;280:20131726.

Caldwell MC, Caldwell DK. Individualized whistle contours in bottlenose dolphins (*Tursiops truncatus*). Nature. 1965;207:434–5.

Caldwell MC, Caldwell DK. Vocalization of naive captive dolphins in small groups. Science. 1968;159:1121–3.

Caldwell MC, Hall NR, Caldwell DK. Ability of an Atlantic bottlenosed dolphin to discriminate between, and respond differentially to, whistles of eight conspecifics. Proc. the 8th Annual Conference on Biological Sonar of Diving Mammals. 1972;57–65.

Caldwell MC, Caldwell DK, Miller JF. Statistical evidence for individual signature whistles in the spotted dolphin, *Stenella plagiodon*. Cetology. 1973;16:1–21.

Caldwell MC, Caldwell DK, Tyack PL. Review of the signature-whistle hypothesis for the Atlantic bottlenose dolphin. In: Leatherwood S, Reeves RR, editors. The Bottlenose Dolphin. San Diego: Academic; 1990. p. 199–234.

Carter GG, Skowronski MD, Faure PA, Fenton B. Antiphonal calling allows individual discrimination in white-winged vampire bats. Anim Behav. 2008;76:1343–55.

Charrier I, Mathevon N, Jouventin P. How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. J Axp Biol. 2002;205:603–12.

Charrier I, Aubin T, Mathevon N. Mother–calf vocal communication in Atlantic walrus: a first field experimental study. Anim Cogn. 2010;13:471–82.

Chmelnitsky EG, Ferguson SH. Beluga whale, *Delphinapterus leucas*, vocalizations from the Churchill River, Manitoba, Canada. J Acoust Soc Am. 2012;131:4821–35.

Clausen KT, Wahlberg M, Beedholm K, DeRuiter S, Madsen PT. Click communication in harbour porpoises *Phocoena phocoena*. Bioacoustics. 2010;20:1–28.

Colbeck GJ, Duchesne P, Postma LD, Lesage V, Hammill MO, Turgeon J. Groups of related belugas (*Delphinapterus leucas*) travel together during their seasonal migrations in and around Hudson Bay. Proc R Soc Lond B: Biol Sci. 2013;280:20122552.

Connor RC, Krützen M. Male dolphin alliances in Shark Bay: changing perspectives in a 30-year study. Anim Behav. 2015;103:223–35.

Connor RC, Wells RS, Mann J, Read AJ. The bottlenose dolphin: social relationships in a

fission-fusion society. In: Mann J, Connor RC, Tyack P, Whitehead H, editors. Cetacean societies: field studies of dolphins and whales. Chicago: University of Chicago; 2000. p. 91–126.

Cortopassi KA, Bradbury JW. Contact call diversity in wild orange-fronted parakeet pairs, *Aratinga canicularis*. Anim Behav. 2006;71:1141–54.

Cummings WC, Thompson PO. Gray whales, *Eschrichtius robustus*, avoid the underwater sounds of killer whales, *Orcinus orca*. Fish Bull. 1971;69:525–30.

Dawson SM. Clicks and communication: the behavioral and social contexts of Hector's dolphin vocalizations. Ethology. 1991;88:265–76.

Deecke VB. Studying marine mammal cognition in the wild: a review of four decades of playback experiments. Aquat Mamm. 2006;32:461–82.

Deecke VB, Slater PJB, Ford JKB. Selective habituation shapes acoustic predator recognition in harbour seals. Nature. 2002;420:171–3.

de Figueiredo LD, Simão SM. Possible occurrence of signature whistles in a population of *Sotalia guianensis* (Cetacea, Delphinidae) living in Sepetiba Bay, Brazil. J Acoust Soc Am. 2009;126:1563–9.

Degrati M, Dans SL, Pedraza SN, Crespo EA, Garaffo GV. Diurnal behavior of dusky dolphins, *Lagenorhynchus obscurus,* in Golfo Nuevo, Argentina. J Mammal. 2008;89:1241–7.

Dreher JJ. Cetacean communication: small group experiment. In: Norris KS, editor. Whales, Dolphins, and Porpoises. Berkeley: University of California; 1966. p. 529–43.

Esch HC, Sayigh LS, Blum JE, Wells RS. Whistles as potential indicators of stress in bottlenose dolphins (*Tursiops truncatus*). J Mammal. 2009;90:638–50.

Evans RM. Foraging-flock recruitment at a black-billed gull colony: implications for the information center hypothesis. Auk. 1982;99:24–30.

Fagen RM, Young DY. Temporal patterns of behavior: durations, intervals, latencies and sequences. In: Colgan PW, editor. Quantitative Ethology. New York: John Wiley; 1978. p. 79–114.

Fay RR. Hearing in Vertebrates: A Psychophysics Databook. Winnetka, Illinois: Hill-Fay Associates; 1988.

Filatova OA, Fedutin ID, Burdin AM, Hoyt E. Responses of Kamchatkan fish-eating killer whales to playbacks of conspecific calls. Mar Mamm Sci. 2011;27:E26–E42.

Filatova OA, Fedutin ID, Nagaylik MM, Burdin AM, Hoyt E. Usage of monophonic and biphonic calls by free-ranging resident killer whales (*Orcinus orca*) in Kamchatka, Russian Far East. Acta Ethol. 2009;12:37–44.

Fish JF, Vania JS. Killer whale, *Orcinus orca*, sound repel white whales, *Delphinapterus leucas*. Fish Bull. 1971;69:531–5.

Ford JKB. Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. Can J Zool. 1989;67:727–45.

Ford JKB. Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. Can J Zool. 1991;69:1454–83.

Ford JKB, Fisher HD. Underwater acoustic signals of the narwhal (*Monodon monoceros*). Can J Zool. 1978;56:552–60.

Foote KG. Broadband acoustic scattering signatures of fish and zooplankton (BASS). Proc. the 3rd European Marine Science and Technology Conference. 1998;3:1012–25.

Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, Ellison S, Firth D, Friendly M, et al. Package “car.” https://cran.r-project.org/web/packages/car/car.pdf (2015). Accessed Jul 2015.

Fripp D, Owen C, Quintana-Rizzo E, Shapiro A, Buckstaff K, Jankowski K, Wells R, Tyack P. Bottlenose dolphin (*Tursiops truncatus*) calves appear to model their signature whistles on the signature whistles of community members. Anim Cogn. 2005;8:17–26.

Goold JC. Behavioural and acoustic observations of sperm whales in Scapa Flow, Orkney Islands. J Mar Biol Assoc UK. 1999;79:541–50.

Gordon JCD. Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations. J Zool. 1991;224:301–14.

Gridley T. Cockcroft VG, Hawkins ER, Lemon-Blewitt M, Morisaka T, Janik VM. Signature whistles in free-ranging populations of Indo-Pacific bottlenose dolphins, *Tursiops aduncus*. Mar Mamm Sci. 2014;30:512–27.

Hauser DDW, Laidre KL, Suydam RS, Richard PR. Population-specific home ranges and migration timing of Pacific Arctic beluga whales (*Delphinapterus leucas*). Polar Biol. 2014;37:1171–83.

Heide-Jørgensen MP, Dietz R, Laidre KL, Richard P, Orr J, Schmidt HC. The migratory behaviour of narwhals (*Monodon monoceros*). Can J Zool. 2003;81:1298–305.

Henderson EE, Hildebrand JA, Smith M H. Classification of behavior using vocalizations of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*). J Acoust Soc Am. 2011;130:557–67.

Hill HM. The behavioral development of two beluga calves during the first year of life. Int J Comp Psychol. 2009;22:234–53.

Hill HM, Kahn MS, Brilliott LJ, Roberts BM, Gutierrez C. Beluga (*Delphinapterus leucas*) bubble bursts: surprise, protection, or play? Int J Comp Psychol. 2011;24:235–43.

Imaizumi T, Furusawa M, Akamatsu T, Nishimori Y. Measuring the target strength spectra of fish using dolphin-like short broadband sonar signals. J Acoust Soc Am. 2008;124:3440–9.

Janik VM. Pitfalls in the categorization of behaviour: a comparison of dolphin whistle classification methods. Anim Behav. 1999;57:133–43.

Janik VM. Source levels and the estimated active space of bottlenose dolphin (*Tursiops truncatus*) whistles in the Moray Firth, Scotland. J Comp Physiol A. 2000a;186:673–80.

Janik VM. Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). Science. 2000b;289:1355–7.

Janik VM, Slater PJB. Vocal learning in mammals. Adv Stud Behav. 1997;26:59–99.

Janik VM, Slater PJB. Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. Anim Behav. 1998;56:829–38.

Janik VM, Slater PJB. The different roles of social learning in vocal communication. Anim Behav. 2000;60:1–11.

Janik VM, Sayigh LS. Communication in bottlenose dolphins: 50 years of signature whistle research. J Comp Physiol A. 2013;199:479–89.

Janik VM, Dehnhardt G, Todt D. Signature whistle variations in a bottlenosed dolphin, *Tursiops truncatus*. Behav Ecol Sociobiol. 1994;35:243–8.

Janik VM, Sayigh LS, Wells RS. Signature whistle shape conveys identity information to bottlenose dolphins. Proc Natl Acad Sci U S A. 2006;103:8293–7.

Jarek S. Package “mvnormtest.” https://cran.r-project.org/web/packages/mvnormtest/mvnormtest.pdf (2015). Accessed July 2015.

Jefferson TA, Newcomer MW, Leatherwood S, van Waerebeek K. Right whale dolphins - *Lissodelphis borealis* (Peale, 1848) and *Lissodelphis peronii* (Lacépède, 1804). In: Ridgway SH, Harrison SR, editors. Handbook of Marine Mammals Vol. 5: The first book of dolphins. London: Academic; 1994. p. 335–62.

Johnston RE. Chemical communication in rodents: from pheromones to individual recognition. J Mammal. 2003;84:1141–62.

Karlsen JD, Bisther A, Lydersen C, Haug T, Kovacs KM. Summer vocalizations of adult male white whales (*Delphinapterus leucas*) in Svalbard, Norway. Polar Biol. 2002;25:808–17.

Killebrew DA, Mercado E, Herman LM, Pack AA. Sound production of a neonate bottlenose dolphin. Aquat Mamm. 2001;27:34–44.

King SL, Janik VM. Bottlenose dolphins can use learned vocal labels to address each other. Proc Natl Acad Sci U S A. 2013;110:13216–21.

King SL, Sayigh LS, Wells RS, Fellner W, Janik VM. Vocal copying of individually distinctive signature whistles in bottlenose dolphins. Proc R Soc Lond B: Biol Sci. 2013;280:20130053.

Klishin VO, Popov VV, Supin AY. Hearing capabilities of a beluga whale, *Delphinapterus leucas*. Aquat Mamm. 2000;26:212–28.

Kondo N, Watanabe S, Izawa E. A temporal rule in vocal exchange among Large-billed crows *Corvus macrorhynchos* in Japan. Ornithol Sci. 2010;9:83–91.

Krasnova VV, Chernetsky AD, Zheludkova AI, Bel'kovich VM. Parental behavior of the beluga whale (*Delphinapterus leucas*) in natural environment. Biol Bull. 2014;41:349–56.

Kureta Y. Vocal exchanges between familiar cotton-top tamarins during social isolation. Jap J Anim Psychol. 2000;50:95–102.

Lammers MO, Au WWL, Herzing DL. The broadband social acoustic signaling behavior of spinner and spotted dolphins. J Acoust Soc Am. 2003;114:1629–39.

Lammers MO, Au WWL, Aubauer R, Nachtigall PE. A comparative analysis of the pulsed emissions of free-ranging Hawaiian spinner dolphins (*Stenella longirostris*). In: Thomas J, Moss C, Vater M, editors. Echolocation in Bats and Dolphins. Chicago: University of Chicago; 2004. p. 414–9.

Lang TG, Smith HAP. Communication between dolphins in separate tanks by way of an electronic acoustic link. Science. 1965;150:1839–44.

Lemasson A, Ouattara K, Bouchet H, Zuberbühler K. Speed of call delivery is related to context and caller identity in Campbell's monkey males. Naturwissenschaften. 2010; 97:1023–7.

Madsen PT, Kerr I, Payne R. Echolocation clicks of two free-ranging, oceanic delphinids with different food preferences: false killer whales *Pseudorca crassidens* and Risso's dolphins *Grampus griseus*. J Exp Biol. 2004;207:1811–23.

Marcoux M, Auger-Méthé M, Humphries MM. Encounter frequencies and grouping patterns of narwhals in Koluktoo Bay, Baffin Island. Polar Biol. 2009;32:1705–16.

Marcoux M, Auger-Méthé M, Humphries MM. Variability and context specificity of narwhal (*Monodon monoceros*) whistles and pulsed calls. Mar Mamm Sci. 2012;28:649–65.

Marler P, Hobbet L. Individuality in a long-range vocalization of wild chimpanzees. Z Tierpsychol. 1975;38:97–109.

Masataka N, Biben M. Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. Behaviour. 1987;101:311–9.

Mathevon N, Charrier I, Jouventin P. Potential for individual recognition in acoustic signals: a comparative study of two gulls with different nesting patterns. C R Biol. 2003;326:329–37.

May-Collado L, Agnarsson I. Cytochrome *b* and Bayesian inference of whale phylogeny. Mol Phylogenet Evol. 2006;38:344–54.

McCowan B, Reiss D. Whistle contour development in captive-born infant bottlenose dolphins (*Tursiops truncatus*): role of learning. J Comp Psychol. 1995a;109:242–60.

McCowan B, Reiss D. Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): a re-evaluation of the signature whistle hypothesis. Ethology. 1995b;100:194–209.

McGarigal K, Cushman S, Stafford S. Multivariate Statistics for Wildlife and Ecology Research. New York: Springer; 2000.

Medwin H. Speed of sound in water: a simple equation for realistic parameters. J Acoust Soc Am. 1975;58:1318–19.

Meschersky IG, Shpak OV, Litovka, DI, Glazov DM, Borisova EA, Rozhnov VV. A genetic analysis of the beluga whale *Delphinapterus leucas* (Cetacea: Monodontidae) from summer aggregations in the Russian Far East. Russ J Mar Biol. 2013;39:125–35.

Michaud R. Sociality and ecology of the odontocetes. In: Ruckstuhl KE, Neuhaus P, editors. Sexual Segregation in Vertebrates: Ecology of the Two Sexes. Cambridge: Cambridge University; 2005. p. 303–26.

Miller CT, Wang X. Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. J Comp Physiol A. 2006;192:27–38.

Miller PJO. Diversity in sound pressure levels and estimated active space of resident killer whale vocalizations. J Comp Physiol A. 2006;192:449–59.

Miller PJO, Bain DE. Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca*. Anim Behav. 2000;60:617–28.

Miller PJO, Shapiro AD, Tyack PL, Solow AR. Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. Anim Behav. 2004;67:1099–107.

Mishima Y, Miyamoto Y, Sasakura T. Development of the Dolphin Speaker. (In Japanese). J Marine Acoust Soc Jpn. 2013;40:27–36.

Mishima Y, Morisaka T, Itoh M, Matsuo I, Sakaguchi A, Miyamoto Y. Individuality embedded in the isolation calls of captive beluga whales (*Delphinapterus leucas*). Zoological Lett. 2015;1:1.

Mooney TA, Nachtigall PE, Castellote M, Taylor KA, Pacini AF, Esteban J-A. Hearing pathways and directional sensitivity of the beluga whale, *Delphinapterus leucas*. J Exp Mar Biol Ecol. 2008;362:108–16.

Moore PW, Dankiewicz LA, Houser DS. Beamwidth control and angular target detection in an echolocating bottlenose dolphin (*Tursiops truncatus*). J Acoust Soc Am. 2008;124:3324–32.

Morgan DW. The vocal and behavioral reactions of the beluga, *Delphinapterus leucas*, to playback of its sounds. In: Winn HE, Olla BL, editors. Behavior of Marine Animals: Current Perspectives in Research. New York: Plenum; 1979. p. 311–43.

Morisaka T. Evolution of communication sounds in odontocetes: a review. Int J Comp Psychol. 2012;25:1–20.

Morisaka T, Connor RC. Predation by killer whales (*Orcinus orca*) and the evolution of whistle loss and narrow-band high frequency clicks in odontocetes. J Evol Biol. 2007;20:1439–58.

Morisaka T, Shinohara M, Taki M. Underwater sounds produced by neonatal bottlenose dolphins (*Tursiops truncatus*): I. Acoustic characteristics. Aquat Mamm. 2005a;31:248–57.

Morisaka T, Shinohara M, Taki M. Underwater sounds produced by neonatal bottlenose dolphins (*Tursiops truncatus*): II. Potential function. Aquat Mamm. 2005b;31:258–65.

Morisaka T, Yoshida Y, Akune Y, Mishima H, Nishimoto S. Exchange of “signature” calls in captive belugas (*Delphinapterus leucas*). J Ethol. 2013;31:141–9.

Morisaka T, Karczmarski L, Akamatsu T, Sakai M, Dawson S, Thornton M. Echolocation signals of Heaviside's dolphins (*Cephalorhynchus heavisidii*). J Acoust Soc Am. 2011;129:449–57.

Mortensen T, Larsen MT, Eriksen PK, Francis DTI, Atkins PR. Octave-bandwidth transducers for ultra-bandwidth echo-sounding system. Proc Inst Acoust. 1999;21:27–41.

Morton A. Occurrence, photo-identification and prey of Pacific white-sided dolphins (*Lagenorhyncus obliquidens*) in the Broughton Archipelago, Canada 1984–1998. Mar Mamm Sci. 2000;16:80–93.

Moscicki MK, Hoeschele M, Bloomfield LL, Modanu M, Charrier I, Sturdy CB. Note types and coding in Parid vocalizations: The *chick-a-dee* call of the boreal chickadee (*Poecile hudsonicus*). J Acoust Soc Am. 2011;129:3327–40.

Murayama T, Iijima S, Katsumata H, Arai K. Vocal imitation of human speech, synthetic sounds and beluga sounds, by a beluga (*Delphinapterus leucas*). Int J Comp Psychol. 2014;27:369–84.

Murayama T, Fujii Y, Hashimoto T, Shimoda A, Iijima S, Hayasaka K, Shiroma N, Koshikawa M. Preliminary study of object labeling using sound production in a beluga. Int J Comp Psychol. 2012;25:195–207.

Murray SO, Mercado E, Roitblat HL. Characterizing the graded structure of false killer whale (*Pseudorca crassidens*) vocalizations. J Acoust Soc Am. 1998;104:1679–88.

Nakahara F, Miyazaki N. Vocal exchanges of signature whistles in bottlenose dolphins (*Tursiops truncatus*). J Ethol. 2011;29:309–20.

Nordhausen K, Oja H, Tyler DE. Tools for exploring multivariate data: the package ICS. J Stat Softw. 2008;28:1–31.

Nousek AE, Slater PJB, Wang C, Miller PJO. The influence of social affiliation on individual vocal signatures of northern resident killer whales (*Orcinus orca*). Biol Lett. 2006;2:481–4.

Panova EM, Belikov RA, Agafonov AV, Bel'kovich VM. The relationship between the behavioral activity and the underwater vocalization of the beluga whale (*Delphinapterus leucas*). Oceanology. 2012;52:79–87.

Parr LA, Winslow JT, Hopkins WD, de Waal FBM. Recognizing facial cues: individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). J Comp Psychol. 2000;114:47–60.

Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S. Animal behaviour: elephants are capable of vocal learning. Nature. 2005;434:455–6.

Quick NJ, Janik VM. Bottlenose dolphins exchange signature whistles when meeting at sea. Proc R Soc Lond B: Biol Sci. 2012;279:2539–45.

Radford AN. Vocal coordination of group movement by green woodhoopoes (*Phoeniculus purpureus*). Ethology. 2004;110:11–20.

Rankin S, Oswald J, Barlow J, Lammers M. Patterned burst-pulse vocalizations of the northern right whale dolphin, *Lissodelphis borealis*. J Acoust Soc Am. 2007;121:1213–8.

Rasmussen MH, Miller LA, and Au WWL. Source levels of clicks from free-ranging white-beaked dolphins (*Lagenorhynchus albirostris* Gray 1846) recorded in Icelandic waters. J Acoust Soc Am. 2002;111:1122–5.

Rasmussen MH, Lammers M, Beedholm K, Miller LA. Source levels and harmonic content of whistles in white-beaked dolphins (*Lagenorhynchus albirostris*). J Acoust Soc Am. 2006;120:510–7.

Recchia CA. Social behaviour of captive belugas, *Delphinapterus leucas*. PhD Thesis. Woods Hole Oceanographic Institution/Massachusetts Institute of Technology; 1994.

Rendell L, Whitehead H. Vocal clans in sperm whales (*Physeter macrocephalus*). Proc R Soc Lond B: Biol Sci. 2003;270:225–31.

Rendell L, Whitehead H. Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurement. Anim Behav. 2004;67:865–74.

Rendell L, Whitehead H. Coda playbacks to sperm whales in Chilean waters. Mar Mamm Sci. 2005;21:307–16.

Richard PR, Heide-Jørgensen MP, Orr JR, Dietz R, Smith TG. Sumer and autumn movements

and habitat use by belugas in the Canadian High Arctic and adjacent areas. Arctic. 2001;54:207–22.

Richards DG, Wolz JP, Herman LM. Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus*. J Comp Psychol. 1984;98:10–28.

Ridgway S, Carder D, Jeffries M, Todd M. Spontaneous human speech mimicry by a cetacean. Curr Biol. 2012;22:860–1.

Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D. Package 'MASS'. https://cran.r-project.org/web/packages/MASS/MASS.pdf (2015). Accessed Jul 2015.

Robeck TR, Monfort SL, Calle PP, Dunn JL, Jensen E, Boehm JR, Young S, Clark ST. Reproduction, growth and development in captive beluga (*Delphinapterus leucas*). Zoo Biol. 2005;24:29–49.

Robisson P, Aubin T, Bremond JC. Individuality in the voice of the emperor penguin *Aptenodytes forsteri*: adaptation to a noisy environment. Ethology. 1993;94:279–90.

Salmi R, Hammerschmidt K, Doran-Sheehy DM. Individual distinctiveness in call types of wild western female gorillas. PLoS ONE. 2014;9:e101940.

Sayigh LS, Tyack PL, Wells RS, Scott MD. Signature whistles of free-raging bottlenose dolphins *Tursiops truncatus*: stability and mother–offspring comparisons. Behav Ecol Sociobiol. 1990;26:247–60.

Sayigh LS, Tyack PL, Wells RS, Solow AR, Scott MD, Irvine AB. Individual recognition in wild bottlenose dolphins: a field test using playback experiments. Anim Behav. 1999;57:41–50.

Schevill WE, Watkins WA. Sound structure and directionality in *Orcinus* (killer whale). Zoologica. 1966;51:70–6.

Schulz TM, Whitehead H, Gero S, Rendell L. Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. Anim Behav. 2008;76:1977–88.

Schulz TM, Whitehead H, Gero S, Rendell L. Individual vocal production in a sperm whale (*Physeter macrocephalus*) social unit. Mar Mamm Sci. 2011;27:149–66.

Shapiro AD. Preliminary evidence for signature vocalizations among free-ranging narwhals (*Monodon monoceros*). J Acoust Soc Am. 2006;120:1695–705.

Shaughnessy PD, Semmelink A, Cooper J, Frost PGH. Attempts to develop acoustic methods of keeping Cape fur seals *Arctocephalus pusillus* from fishing nets. Biol Conserv. 1981;21:141–58.

Sibly RM, Nott HMR, Fletcher DJ. Splitting behaviour into bouts. Anim Behav. 1990;39:63–9.

Sjare BL, Smith TG. The vocal repertoire of white whales, *Delphinapterus leucas*, summering in Cunningham Inlet, Northwest Territories. Can J Zool. 1986a;64:407–15.

Sjare BL, Smith TG. The relationship between behavioral activity and underwater vocalizations of the white whale, *Delphinapterus leucas*. Can J Zool. 1986b;64:2824–31.

Slater PJB, Lester NP. Minimising errors in splitting behaviour into bouts. Behaviour. 1982;79:153–61.

Smolker R, Pepper JW. Whistle convergence among allied male bottlenose dolphins (Delphinidae, *Tursiops* sp.). Ethology. 1999;105:595–617.

Soltis J, Leong K, Savage A. African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. Anim Behav. 2005;70:589–99.

StantonTK. Broadband acoustic sensing of the ocean. J Marine Acoust Soc Jpn. 2009;36:95–107.

Starkhammar J, Moore PW, Talmadge L, Houser DS. Frequency-dependent variation in the two dimensional beam pattern of an echolocating dolphin. Biol Lett. 2011;7:836–9.

Steiner WW. Species-specific differences in pure tonal whistle vocalizations of five western North Atlantic dolphin species. Behav Ecol Sociobiol. 1981;9:241–6.

Sugiura H. Temporal and acoustic correlates in vocal exchange of coo calls in Japanese

macaques. Behaviour. 1993;124:207–25.

Svetochev VN, Svetocheva ON. Summer migration activity of the beluga whale *Delphinapterus leucas* in Dvina Bay, the White Sea. Dokl Biol Sci. 2013;448:17–21.

Trimble M, Charrier I. Individuality in South American sea lion (*Otaria flavescens*) mother–pup vocalizations: implications of ecological constraints and geographical variations? Mammalian Biol. 2011;76:208–16.

Tyack P. Whistle repertoires of two bottlenosed dolphins, *Tursiops truncatus*: mimicry of signature whistles? Behav Ecol Sociobiol. 1986;18:251–7.

van Parijs SM, Corkeron PJ. Evidence for signature whistle production by a Pacific humpback dolphin, *Sousa chinensis*. Mar Mammal Sci. 2001;17:944–9.

van Parijs SM, Lydersen C, Kovacs KM. Sounds produced by individual white whales, *Delphinapterus leucas*, from Svalbard during capture (L). J Acoust Soc Am. 2003;113:57–60.

Vaughn-Hirshorn RL, Hodge KB, Würsig B, Sappenfield RH, Lammers MO, Dudzinski KM. Characterizing dusky dolphin sounds from Argentina and New Zealand. J Acoust Soc Am. 2012;132:498–506.

Vergara V. Acoustic communication and vocal learning in belugas (*Delphinapterus leucas*). PhD Thesis. University of British Columbia; 2011.

Vergara V, Barrett-Lennard LG. Vocal development in a beluga calf (*Delphinapterus leucas*). Aquat Mamm. 2008;34:123–43.

Vergara V, Michaud R, Barrett-Lennard L. 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. 2010;23:278–309.

Volodina EV, Volodin IA, Isaeva IV, Unck C. Biphonation may function to enhance individual recognition in the dhole, *Cuon alpinus*. Ethology 2006;112:815–25.

Wanker R, Sugama Y, Prinage S. Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. Anim Behav. 2005;70:111–8.

Watkins WA. The harmonic interval: fact or artifact in spectral analysis of pulse trains. In: Tavolga WN, editor. Marine Bio-acoustics. Vol. 2. New York: Pergamon; 1967. p. 15–43.

Watkins WA, Schevill WE. Sperm whale codas. J Acoust Soc Am. 1977;62:1485–90.

Watkins WA, Schevill WE, Best PB. Underwater sounds of *Cephalorhynchus heavisidii* (Mammalia: Cetacea). J Mammal. 1977;58:316–20.

Watt CA, Petersen SD, Ferguson SH. Genetics and fatty acids assist in deciphering narwhal (*Monodon monoceros*) social groupings. Polar Biol. 2015;38:1971–81.

Watwood SL, Tyack PL, Wells RS. Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. Behav Ecol Sociobiol. 2004;55:531–43.

Watwood SL, Owen ECG, Tyack PL, Wells RS. Signature whistle use by temporarily restrained and free-swimming bottlenose dolphins, *Tursiops truncatus*. Anim Behav. 2005;69:1373–86.

Weihs C, Ligges U, Luebke K, Raabe N. klaR analyzing German business cycles. In: Baier D, Decker R, Schmidt-Thieme L, editors. Data Analysis and Decision Support. Berlin: Springer; 2005. p. 335–43.

Weilgart L, Whitehead H. Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. Behav Ecol Sociobiol. 1997;40:277–85.

Whitehead H, Weilgart L. Patterns of visually observable behaviour and vocalizations in groups of female sperm whales. Behavior 1991;118:275–96.

Whitehead H, Weilgart L. The sperm whale: social females and roving males. In: Mann J, Connor RC, Tyack P, Whitehead H, editors. Cetacean societies: field studies of dolphins and whales. Chicago: University of Chicago; 2000. p. 154–172.

Würsig B, Würsig M. Behavior and ecology of the dusky dolphin, *Lagenorhynchus obscurus*, in the South Atlantic. Fish Bull. 1980;77:871–90.

Yoshida YM, Morisaka T, Sakai M, Iwasaki M, Wakabayashi I, Seko A, Kasamatsu M, Akamatsu T, Kohshima, S. Sound variation and function in captive Commerson's dolphins (*Cephalorhynchus commersonii*). Behav Process. 2014;108:11–9.

Yosida S, Kobayasi K, Ikebuchi M, Ozaki R, Okanoya K. Antiphonal vocalization of a subterranean rodent, the naked mole-rat (*Heterocephalus glaber*). Ethology. 2007;113:703–10.

Yurk H, Barrett-Lennard L, Ford JKB, Matkin CO. Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. Anim Behav. 2002;63:1103–19.