INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

2020, 33
Heather M. Hill
Editor
Peer-reviewed

Preliminary Study of Discrimination of Human Vocal Commands in Walrus (*Odobenus rosmarus divergens*)

Shiho Endo[1], Naoki Kawaguchi[1], Yusuke Shimizu[1], Asuka Imagawa[1], Tomohiro Suzuki[1], Harumasa Ashikari[1], Yoshihito Wakai[1], and Tsukasa Murayama[2]

[1] Toba Aquarium, Japan
[2] Tokai University, Japan

Walruses appear to use various acoustic signals in different social contexts. The auditory faculty seems to be important for walruses. Can walruses understand another animal's vocal information using their auditory sense? This study tested whether a male walrus could discriminate human speech sounds and perform different actions corresponding to each one under various conditions. The subject, a male walrus (*Odobenus rosmarus divergens*) named Pou, was placed on the ground. The experimenter presented vocal commands to the subject under 3 conditions. (1) The experimenter stood near to the subject and presented each vocal command while wearing cloak and goggles so that the experimenter's eye and body movements would not influence the subject's behavior. (2) A wooden board was placed between the experimenter and the subject so that the subject could not see the experimenter. (3) A wooden board was placed between the experimenter and the subject so that the subject could not see the experimenter, and the experimenter presented each vocal command through an audio speaker. Under each condition, when the subject performed the correct action corresponding to the vocal commands, he was rewarded with a piece of fish. Results demonstrated that the subject responded correctly to almost all of the human vocal stimuli in every condition (10 kinds of stimuli; correct responses were above 80%), including when the experimenter (presenter of the commands) was not visible. This suggests that the subject only responded to vocal cues and not to other inadvertent ones, such as visual or tactile stimuli. This study demonstrated that walruses can hear and identify human speech sounds using their auditory sense and can discriminate auditory cues.

Keywords: human vocal command, walrus

Pinnipeds are thought to utilize their calls to communicate with one another (Lindemann et al., 2006; Schusterman, 1978; Schusterman et al., 2001). Walruses appear to prefer to create groups, and they use various acoustic signals in different social contexts (Kastelein, 2002). For example, they utter calls to intimidate or signal obedience to another individual (Charrier et al., 2011; Kastelein, 2002), and male walruses utter unique calls to draw the attention of females or to keep away rival males (Charrier et al., 2011; Kastelein, 2002). Moreover, mothers and calves use calls to communicate with one another (Charrier et al., 2010). Thus, the auditory faculty is important for walruses. However, limited information exists regarding walruses' auditory abilities (Kastelein et al., 2002; Supin et al., 2001).

Can walruses understand another animal's vocal information using their auditory sense? In some aquariums in Japan, walruses are trained to discriminate human speech sounds. Previous studies have investigated the comprehension of human vocal sounds in animals such as chimpanzees (Hays, 1951), bonobos (Savage-Rumbaugh, 1993), and African gray parrots (Pepperberg, 1990, 2002). Among marine mammals, bottlenose dolphins (Lilly, 1961, 1967), belugas (Murayama et al., 2014; Ridgway et al., 2012), and killer whales (Abramson et al., 2018) have been shown to be able to discriminate human speech sounds. Yet, there is currently no information on walruses' ability to discriminate human speech sounds.

The objective of this study was to determine whether walruses could perform different actions when human vocal commands are presented. This study tested whether a male walrus could correctly follow human

Please send correspondence to Tsukasa Murayama, Tokai University, School of Marine Science and Technology, 3-20-1 Orido, Shimizu-ku, Shizuoka, Japan. Electronic mail may be sent to beluga@scc.u-tokai.ac.jp https://doi.org/10.46867/ijcp.2020.33.00.03

vocal commands under various conditions to examine whether the walrus could discriminate auditory cues in the absence of other cues.

Method

Subject

The subject was a male walrus (*Odobenus rosmarus divergens*) named Pou (948 kg in body weight, 9 years old; Figure 1). He was born in Russia and has been kept at Toba Aquarium in Mie prefecture, Japan. He was trained in a variety of performances with the experimenters every day. Another walrus was kept in the same pool; however, this walrus was isolated during the experiment to avoid any possible influence on the experiment.

Figure 1
The Subject, a Walrus Named Pou.

Auditory Stimuli

The subject was routinely trained with 15 vocal commands (Table 1); therefore, these commands were familiar to the subject. In the experiments, only 10 of these 15 commands (Table 1) were presented to the subject as the sample stimuli in the test session. These 10 vocal commands were presented by the experimenter to the subject, either out loud or through an audio speaker, as described in the *Procedure* section. As noted in the *Introduction*, when the subject heard the vocal stimuli, he performed different actions accordingly.

Table 1
Human Vocal Commands and their Corresponding Actions

Vocal Commands	Actions	Commands Presented During the Test
Ganbaruzo	Raising the fore-fin while uttering voice	*
Hazukashi	Covering the head with right fore-fin	*
O ho ho ho	Shaking the head up and down while uttering voice	*
Onaka	Beating the belly with right fore-fin	*
Ashi	Extending both hind-fins backward	*
Bye bye	Shaking the left fore-fin with raising	*

Vocal Commands	Actions	Commands Presented During the Test
Bu!	Uttering a low voice	*
Ban!	Falling sideways	*
Dekinai	Shaking the head left and right	*
Rei	Bowing	*
Fuse	Being in prone position	
Oyaji	Lying down with flippers folded	
Goron shite	Facing upward	
Fuh shite	Blowing breath strongly	
Maitta	Holding head with right fore-flipper	

Procedure

The experiment took place in the ground space of a pool at the aquarium. The subject was placed on the ground, and the experimenter presented one of the vocal commands listed in Table 1 under the following conditions. When the subject performed the correct action corresponding to the stimulus, he was rewarded with a piece of fish. When he did not respond correctly, the next command was presented after a 3 s interval without any reward. During the experiment, 10-15 trials were performed in each session, and 10 vocal commands were presented in a random order. Three experimenters took part in the experiment, and each experimenter alternated for each session in a random order. Whether the subject responded correctly or not was judged by all three experimenters.

The experiments presented each vocal command under the following three conditions.

Condition 1: The subject sat facing the front. The experimenter stood close to the subject (Figure 2) and presented each vocal command out loud (Figure 3a).

Condition 2: The experimenter stood close to the subject wearing a cloak and goggles so that the experimenter's eye and body movements would not influence the subject's behavior (Figure 3b). The subject and the experimenter stood in the same position during each trial. Then, the experimenter presented each command vocally.

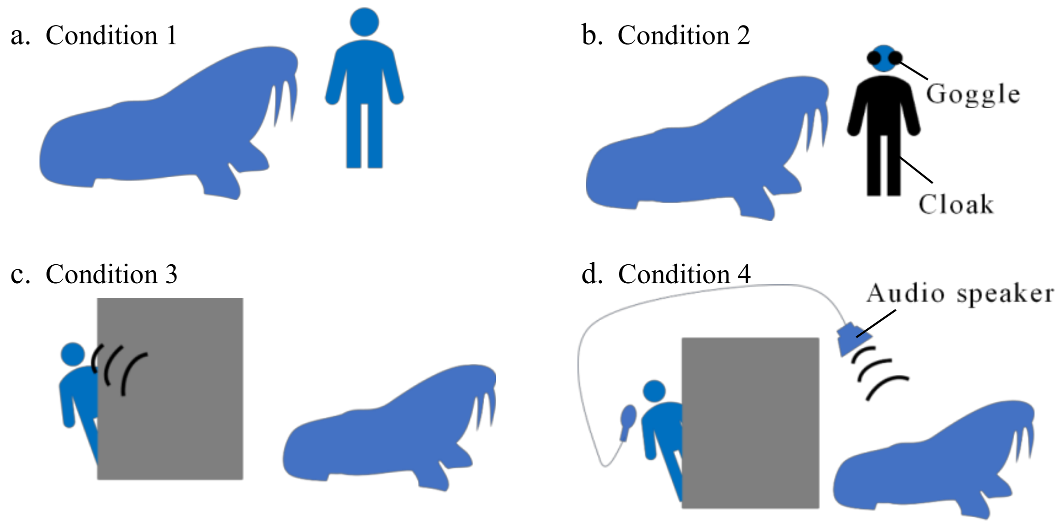
Figure 2
The Subject and the Experimenter who Stood Close to the Subject

Note: The subject sat facing the front in every trial.

Condition 3: A wooden board (1 m × 2 m) was placed 3 m away from the subject and the experimenter presented the commands from behind the board so that the subject could not see the experimenter (Figure 3c). Then, the experimenter presented each vocal command using their voice. (The subject could hear the experimenter's voice directly.).

Condition 4: As in Condition 3, a wooden board was placed between the experimenter and the subject so that the subject could not see the experimenter. Then, the experimenter presented vocal commands through an audio speaker (Figure 3d).

Figure 3
Schematic Diagram of each Experimental Condition



In each condition, 10 vocal commands (Table 1) out of 15 trained commands were presented as a test trial. Each command was presented in a random order; therefore, the number of presentations was not uniform in each session. In each session, 10-15 trials were performed, and the interval between each trial was approximately 3 s. The total number of presentations for each command was 15 in every condition. Each vocal command was presented 10 times in a random order for every condition.

All the research activities adhered to the Ethical Guidelines for the Conduct of Research Animals by Zoo and Aquariums issued by the World Association on Zoos and Aquariums (WAZA), the Code of Ethics issued by the Japanese Association of Zoos and Aquariums (JAZA), and the Japanese Act on Welfare and Management of Animals. All experimental protocols were approved by the Toba Aquarium.

Statistics

Because the subject was trained in 15 commands (Table 1), the chance level was 6.7%. A chi-square test and a binomial test were used to determine whether the frequency of response and the percentages of correct responses were statistically significant.

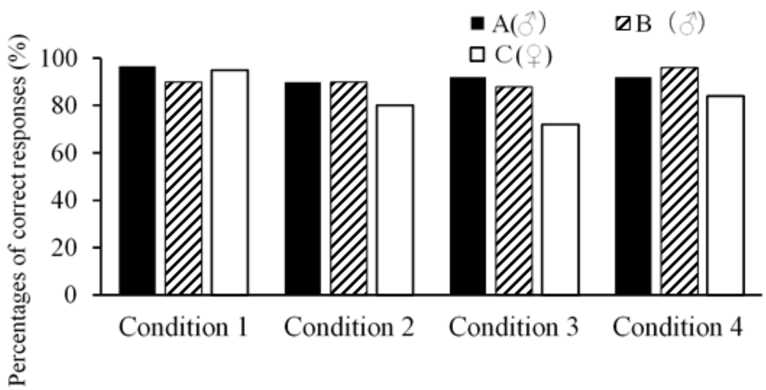
Results

Accuracy Rate for Each Experimenter

In every condition, the stimuli were presented by three experimenters, and they took turns in a random order in each session (i.e., each experimenter changed per trial in a random order). Then, the accuracy rate for each experimenter was calculated. Figure 4 shows the percentages of correct responses by experimenter. The percentage was high for each experimenter, and there was no significant difference among experimenters (p = 0.10 ANOVA), that is, it was demonstrated that there was little difference across experimenters.

Figure 4

Percentages of Correct Responses for each Experimenter (A, B, C)



Accuracy Rate in Each Condition

In all four conditions, the subject performed actions in response to the stimuli.

Condition 1: The frequency of response to each command is shown in Table 2. Although the subject mistook "Onaka" and "Rei" for "Ashi", "Bye bye," and "Fuse" in incorrect responses, the subject performed correct actions for most of the commands. The subject responded correctly to the presented commands for every command ($p < 0.01$, Chi-square test).

Table 2

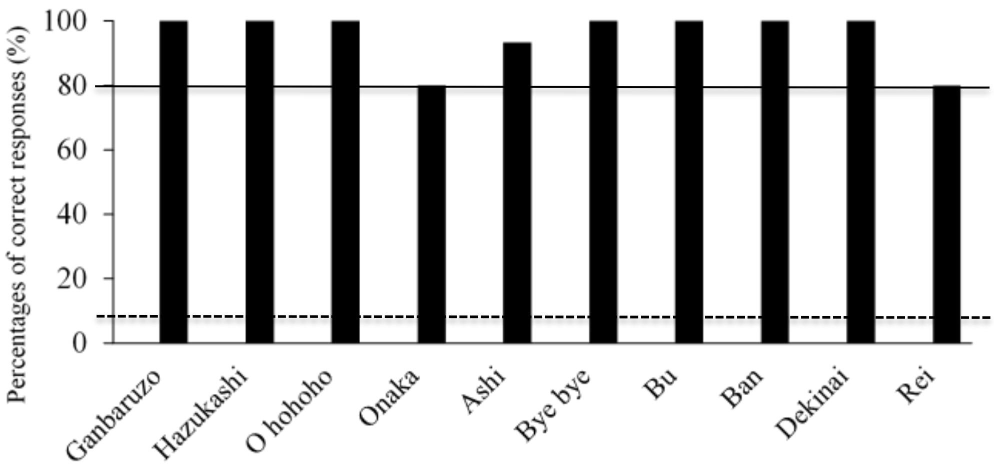
Frequency of Responses of each Command in Condition 1

	Ganba ruzo	Hazu kashi	O ho ho ho	Onaka	Ashi	Bye bye	Bu!	Ban!	Deki nai	Rei	Fuse	Oyaji	Goron shite	Fuh shite	Maitta	No response
Ganba ruzo	15															
Hazu kashi		15														
O ho ho ho			15													
Onaka				12	1	2										
Ashi					14											1
Bye bye						15										
Bu!							15									
Ban!								15								
Deki nai									15							
Rei										12	3					

Then, the percentages of correct responses for each command were calculated (Figure 5) and indicated that they were above the chance level and/or significance level for every command ($p = 0.05$, Binomial test).

Figure 5

Percentages of Correct Responses for each Stimulus in Condition 1



Note: The solid line indicates the significance level ($p = 0.02 < 0.05$, Binomial test), and the dashed line indicates the chance level.

Condition 2: The frequency of responses to each command is shown in Table 3. In response to some commands, the subject sometimes performed an incorrect action. However, with these exceptions, the subject responded correctly, without any confusion, to the presented commands ($p < 0.01$, Chi-square test), even though the experimenter's eye and body movements were hidden.

Table 3

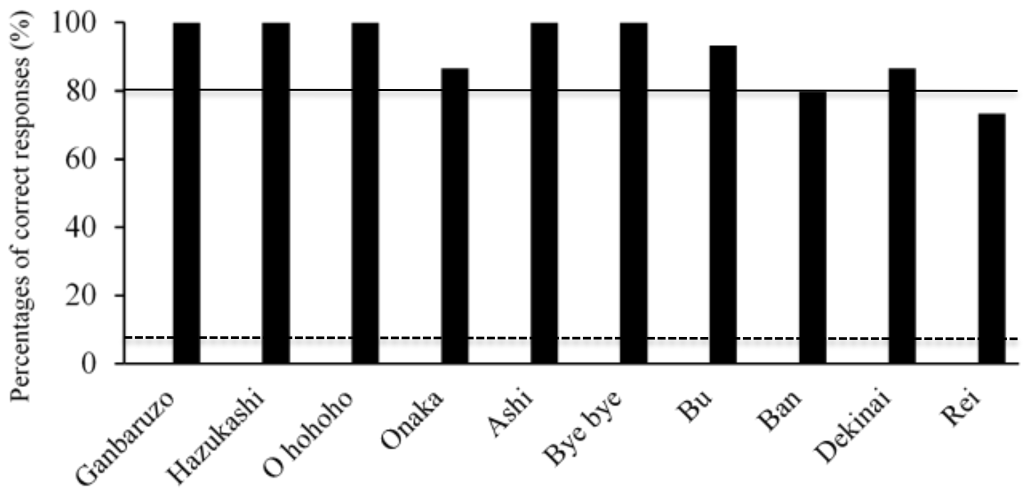
Frequency of Responses for each Command in Condition 2

	Ganba ruzo	Hazu kashi	O ho ho ho	Onaka	Ashi	Bye bye	Bu!	Ban!	Deki nai	Rei	Fuse	Oyaji	Goron shite	Fuh shite	Maitta	No response
Ganba ruzo	15															
Hazu kashi		15														
O ho ho ho			15													
Onaka				13								2				
Ashi					15											
Bye bye						15										
Bu!							14				1					
Ban!								12		2	1					
Deki nai									13	2						
Rei								1		11	3					

The percentages of correct responses are shown in Figure 6 and indicated that the subject performed correct actions to the commands significantly for every command (p = 0.02, Binomial test).

Figure 6

Percentages of Correct Responses for each Stimulus in Condition 2



Note: The solid line indicates the significance level (p = 0.05, Binomial test) and the dashed line indicates the chance level.

Condition 3: The frequency of responses to each command is shown in Table 4. The subject responded incorrectly to some commands. However, he responded correctly for other presented commands (p < 0.01, Chi-square test), even though the experimenter was not visible.

Table 4

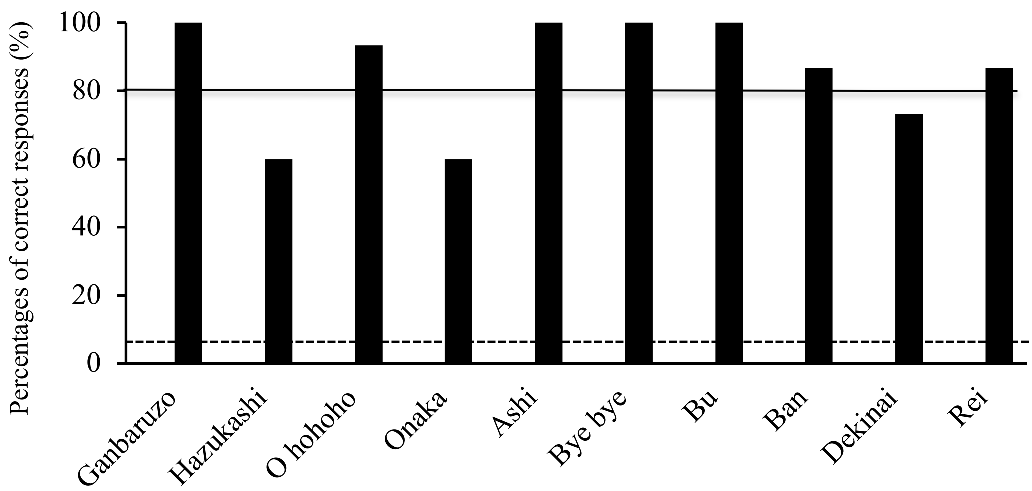
Frequency of Responses in each Command in Condition 3

	Ganba ruzo	Hazu kashi	O ho ho ho	Onaka	Ashi	Bye bye	Bu!	Ban!	Deki nai	Rei	Fuse	Oyaji	Goron shite	Fuh shite	Maitta	No response
Ganba ruzo	15															
Hazu kashi		9			6											
O ho ho ho			14													1
Onaka				9	1	4										1
Ashi					15											
Bye bye						15										
Bu!							15									
Ban!					1			13		1						
Deki nai					4				11							
Rei					1			1		13						

The percentages of correct responses for each command were high and exceeded the chance level (6.7%) and the significance level (p = 0.02 Binomial test; Figure 7).

Figure 7

Percentages of correct responses for each stimulus in Condition 3



Note: The solid line indicates the significance level (p = 0.05, Binomial test), and the dashed line indicates the chance level.

Condition 4: The frequency of responses to each command is shown in Table 5. Even though the experimenter was not visible, the subject responded correctly for most of the presented commands (p < 0.01, Chi-square test).

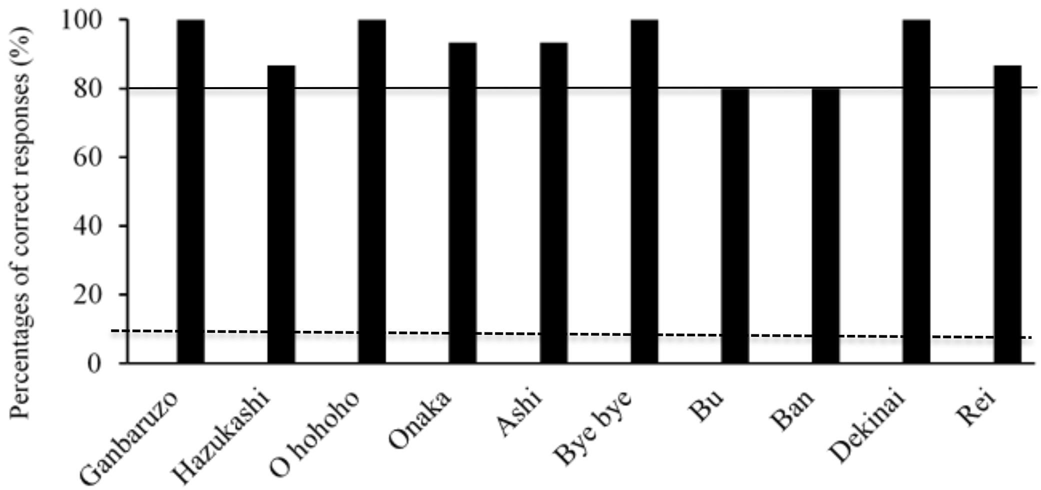
Table 5

Frequency of Responses for each Command in Condition 4

	Ganba ruzo	Hazu kashi	O ho ho ho	Onaka	Ashi	Bye bye	Bu!	Ban!	Deki nai	Rei	Fuse	Oyaji	Goron shite	Fuh shite	Maitta	No response
Ganba ruzo	15															
Hazu kashi		13			2											
O ho ho ho			15													
Onaka				14		1										
Ashi					14					1						1
Bye bye						15										
Bu!							12							1		2
Ban!					1			12		2						
Deki nai									15							
Rei									1	13						1

The percentages of correct responses for all the commands exceeded the chance levels (6.7%) and were significantly high (p = 0.02, Binomial test; Figure 8).

Figure 8
Percentages of Correct Responses for each Stimulus in Condition 4.



Note: The solid line indicates the significance level (p = 0.05, Binomial test), and the dashed line indicates the chance level.

Discussion

The subject responded correctly to almost all of the human vocal commands in every condition, including when the speaker was not visible. However, in Condition 3, the percentages of correct responses were low for some commands. It is thought that the subject did not hear the commands because the board muffled the experimenter's voice. In Condition 4, the subject responded correctly to most of the commands. This means that he was indeed responding to the human speech sounds and not nonvocal cues. Moreover, the same results were obtained for different experimenters. These results demonstrate that the subject could hear and discriminate human speech sounds.

Walruses appear to discriminate other walruses' calls and communicate with one another through their vocalization (Berta & Sumich, 1999). In the present study, it is demonstrated that the walrus could discriminate the vocalization of another species (i.e., human speech sounds). Because the walrus performed actions correctly to the human vocal commands, the walrus could not only discriminate but could also correctly process human speech sounds. That is, it was clarified that the walrus could respond to human vocal commands in an operant conditioning context.

In previous studies, several species have been shown to discriminate and understand human speech sounds. A chimpanzee named Viki was trained to listen to human vocal sounds and imitate them (Hayes, 1951), and a bottlenose dolphin was taught to pronounce human vocal sounds in the form of the alphabet (Lilly, 1961). A male gray parrot named Alex could understand human vocal sounds and was asked to answer several questions (Pepperberg, 1990, 2002). Moreover, Murayama et al. (2014) described that a male beluga named Nack could discern human speech sounds and imitate them correctly. These studies indicate that some animals can discriminate human speech sounds.

Dolphins could perform correct actions corresponding to presented auditory commands in a condition discrimination task (reviewed in Herman, 1986, 1988; Murayama et al., 2017), that is, dolphins are thought to be able to discriminate auditory stimuli and unite auditory stimuli with corresponding behaviors. In the present study, the walrus responded correctly to most of the vocal commands, suggesting that the walrus was able to convert auditory cues to behavioral responses, as reported with the dolphin (reviewed in Herman, 1986, 1988).

Although the auditory ability of walruses is unclear in many ways, this study demonstrated that walruses could hear and identify human speech sounds using their auditory sense. As mentioned above, artificial sounds synthesized through a computer, gestural languages, lexigrams, and sign languages were employed in dolphins, sea lions, and some primates in the animal language study. However, some vocal sounds were employed in this study. The findings of this study revealed that walrus could distinguish between vocal sounds of other species (i.e., human voice, and understand human vocal commands, as demonstrated in gray parrots; reviewed by Pepperberg, 2002). This may be the key to understanding not only the auditory abilities but also the cognitive abilities of walruses.

References

Abramson, J. Z., Hernandez-Lloreda, M. V., Garcıa, L., Colmenares, F., Aboitiz, F., & Call, J. (2018). Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*). *Proceedings of the Royal Society B*, *285,* 20172171. http://dx.doi.org/10.1098/rspb.2017.2171

Berta, A., & Sumich, J. L. (1999). *Marine mammals: Evolutionary biology.* Academic Press.

Charrier, I., Aubin, T., & Mathevon, N. (2010). Mother-Calf vocal communication in Atlantic walrus: A first field experimental study. *Animal Cognition*, *13*, 471–482.

Charrier, I., Burlet, A., & Aubin, T. (2011). Social vocal communication in captive Pacific walruses *Odobenus rosmarus divergens*. *Mammalian Biology, 76,* 622-627. http://doi.org/101016/j.mambio.2010.10.006.

Hayes, C. (1951). *The ape in our house*. Harper.

Herman, L. M. (1986). Cognition and language competencies of bottlenosed dolphins. In R. A. Schusterman, J. A. Thomas & F. G. Wood (Eds.), *Dolphin cognition and behavior: A comparative approach* (pp. 221–252). Lawrence Erlbaum.

Herman, L. M. (1988). Cognitive characteristics of dolphins. In L. M. Herman (Ed.), *Cetacean behavior: Mechanisms & functions* (pp. 363–429). John Wiley & Son.

Kastelein, R. A. (2002). Walrus. In W. Perrin, B. Wursig, & J. G. M. Thewissen (Eds.), *Encyclopedia of marine mammals* (pp. 1294–1300). Academic Press.

Kastelein, R.A., Mosterd, P., van Santen, B., & Hagedoorn, M. (2002). Underwater audiogram of a Pacific walrus (*Odobenus rosmarus divergens*) measured with narrow-band frequency-modulated signals. *Journal of Acoustical Society of America*, *112,* 2173–2182.

Lilly, J. C. (1961). *Man and dolphin: Adventures of new scientific frontier*. Doubleday.

Lilly, J. C. (1967). *The mind of the dolphin*. Doubleday.

Lindemann, K. L., Kastak, C. R., & Schusterman, R. J. (2006). The role of learning in the production and comprehension of auditory signals by pinnipeds. *Aquatic Mammals, 32*, 483–490.

Murayama, T., Iijima, S., Katsumata, H., & Arai, K. (2014). Vocal imitation of human speech, synthetic sounds and beluga sounds, by a beluga (*Delphinapterus leucas*). *International Journal of Comparative Psychology*, *27,* 369–384. https://escholarship.org/uc/item/51v1z12b

Murayama, T., Suzuki, R., Kondo, Y., Koshikawa, M., Katsumata, H., & Arai, K. (2017). Spontaneous establishing of cross-modal stimulus equivalence in a beluga whale. *Scientific Report, 7*, 9914. https://doi.org/10.1038/s41598-017-09925-4

Pepperberg, I. M. (1990). Cognition in an African gray parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. *Journal of Comparative Psychology*, *104,* 41–52.

Pepperberg, I. M. (2002). *The Alex studies: Cognitive and communicative abilities of gray parrots*. Harvard University Press.

Ridgway, S., Carder, D., Jeffries, M., & Todd, M. (2012). Spontaneous human speech mimicry by a cetacean. *Current Biology*, *22*, 860–861. https://doi.org/10.1016/j.cub.2012.08.044

Savage-Rumbaugh, E. S. (1993). *Kanzi: A most improbable ape*. NHK Publishing.

Schusterman, R. J. (1978). Vocal communication in pinnipeds. In H. Markowitz & V. J. Stevens (Eds.), *Behavior of captive wild animals* (pp. 247–285). Nelson-Hall.

Schusterman, R. J., Southall, B. L., Kastak, D., & Kastak, C. R. (2001). Pinniped vocal communication: Form and function. *Proceedings of the 17th International Congress on Acoustics Proceedings*, Rome, Italy.

Supin, A. Ya,, Popov, V. V., & Mass, A. M. (2001). *The sensory physiology of aquatic mammals.* Kluwer Academic Publishers.

Financial conflict of interest: No stated conflicts.
Conflict of interest: No stated conflicts.

Submitted: *June 11th, 2018*
Resubmitted: *June 29th, 2019*
Accepted: *February 13th, 2020*