INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

2025, 38
Charles I. Abramson
Editor
Peer-reviewed

A Case Study of Associations Between Human Visual-Vocal Commands and Behaviors in a Lactating Steller Sea Lion Pup (*Eumetopias jubatus*)

Masahiro Sasaki[1], Kazuki Tsutsumi[1], Hinano Kinoshita[1], Masahiro Nishijima[1], Chiori Matsumura[1], Ayaka Toyota[1], & Toshimune Kambara[2]

[1] *Kinosaki Marine World, 1090 Seto, Toyooka, Hyogo 669-6192, Japan*

[2] *Department of Psychology, Graduate School of Humanities and Social Sciences, Hiroshima University, 1-1-1 Kagamiyama, Higashihiroshima, Hiroshima, Japan*

This study aimed to examine whether a lactating Steller sea lion pup (*Eumetopias jubatus*) could learn the associations between human commands (verbal/ visual) and behavior. A male Steller sea lion pup named Kanata was artificially cared for from the age of 36 hours. When Kanata was three months old, we began behavior-shaping training using milk as a reinforcer. During the training phase, which lasted approximately three months, Kanata learned nine behaviors. The trainers simultaneously presented Kanata with both hand signals and vocal commands to condition each behavior's association with the human commands. In the test phase, we tested the effects of associative learning between the human commands and each behavior in the three conditions. When trainers simultaneously gave him hand signals and vocal commands (Condition 1), Kanata's performance was significantly higher than for the other conditions (Condition 2: only hand signals and Condition 3: only vocal commands). Additionally, Kanata's performance was significantly higher in Condition 2 than in Condition 3. Although Kanata's performance was stable for all commands in both Conditions 1 and 2, it was significantly low for almost all commands in Condition 3. These results provide the first evidence to suggest that a Steller sea lion pup has an excellent ability of associative learning and that he could shape the associations between the human commands and behaviors in both visual and auditory modalities, but his learning association could be dominant in visual information, compared to auditory information.

Keywords: artificial care, associative learning, hand signals, Steller sea lion pup, training, vocal commands

授乳期のトド仔獣 (*Eumetopias jubatus*) における視覚−音声コマンドを用いた連合学習の事例研究

本研究の目的は、授乳期のトドの仔獣がヒトのコマンド（視覚および音声コマンド）と動作の連合を学習できるかについて調べることであった。オスのトド・カナタは生後 36 時間の時点から人工哺育を施された。私たちはカナタが生後 3 ヵ月の時、ミルクを強化子として行動形成トレーニングを開始した。約 3 ヵ月間のトレーニング期間に、カナタは 9 種類の動作を習得した。トレーナーはハンドシグナル（視覚のコマンド）とボーカルコマンド（音声のコマンド）を同時に提示することで、学習した動作とコマンドを条件付けた。テスト段階では 3 つのコンディションにおいてコマンドと動作の連合学習の効果を調べた。ハンドシグナルとボーカルコマンドを同時に提示したコンディション 1 の成績は、ハンドシグナルのみを提示したコンディション 2 およびボーカルコマンドのみを提示したコンディション 3 よりも有意に高かった。また、コンディション 2 の成績はコンディション 3 の成績よりも有意に高かった。コンディション 1、2 においてカナタの成績はすべてのコマンドで安定していたが、コンディション 3 においてはほぼすべてのコマンドの成績が有意に低かった。これらの結果は、トドの仔獣が視覚と聴覚双方のモダリティを用いてコマンドと動作の連合を学習できる卓越した連合学習能力を持つが、聴覚の情報に比べて視覚の情報が優位に連合学習されることを示唆している。

キーワード：人工哺育、連合学習、ハンドシグナル、トド仔獣、トレーニング、ボーカルコマンド

Please send correspondence to Masahiro Sasaki, Kinosaki Marine World, 1090 Seto, Toyooka, Hyogo 669-6192, Japan. Email: cod.beautiful.gbr@gmail.com

Un estudio de caso sobre las asociaciones entre comandos visuales/vocales humanos y el comportamiento en una cría lactante de león marino de Steller (Eumetopias jubatus)

El objetivo de este estudio fue evaluar la capacidad de aprendizaje asociativo entre señales humanas (visuales y auditivas) y comportamientos en una cría lactante de león marino de Steller (Eumetopias jubatus). Una cría macho, llamado Kanata, recibió cuidados artificiales desde las 36 horas de vida. A los tres meses, se inició un entrenamiento conductual basado en reforzamiento positivo (leche), que se extendió por tres meses. Durante este período, Kanata aprendió nueve comportamientos mediante la asociación simultánea de señales manuales y comandos vocales. En la fase de prueba, se evaluó el rendimiento comportamental del sujeto bajo tres condiciones de asociación con los comandos humanos: comandos visuales y auditivos simultáneos (condición 1), en donde el rendimiento de Kanata fue significativamente mayor que en las otras condiciones; Comandos solo visuales (condición 2) y solo auditivos (condición 3), en donde el desempeño de Kanata fue significativamente mayor en la Condición 2 que en la Condición 3. Adicionalmente, aunque el rendimiento de Kanata fue estable para todos los comandos en las Condiciones 1 y 2, fue significativamente bajo para casi todos los comandos en la Condición 3. Estos resultados proporcionan la primera evidencia que sugiere que una cría de león marino de Steller tiene una excelente capacidad de aprendizaje asociativo y que podría moldear las asociaciones entre los comandos humanos y los comportamientos tanto en las modalidades visuales como auditivas, pero su asociación de aprendizaje podría ser dominante en la información visual, en comparación con la información auditiva.

Palabras clave: cuidado artificial, aprendizaje asociativo, señales manuales, cría de león marino de Steller, entrenamiento, comandos vocales

Researchers have studied the ability of non-human animals to understand human commands in some species. The main findings on command learning relate to associative learning, with cases using visual and acoustic stimuli. Examples of human command learning research using visual stimuli include studies of sign language learning in great apes and dolphins. Gardner and Gardner (1969) reported not only that a female chimpanzee (*Pan troglodytes*) could understand the signs given by humans, but also that the chimpanzee herself used sign language to express the names of objects. Patterson and Cohn (1990) also reported a female lowland gorilla named Koko could learn language using hand signs. In cetaceans, bottlenose dolphins (*Tursiops truncatus*) can learn associations between their behaviors and human sign language (Herman et al., 1984).

However, several findings have been made on the human command learning ability using acoustic information in non-human animals. Dogs and parrots recognize human vocal commands. Chaser, a border collie (*Canis lupus familiaris*), was able to learn the names of toys (Griebel & Oller, 2012; Pilley & Reid, 2011). Alex, a male African grey parrot (*Psittacus Erithacus*), has also been reported to be able to learn many nouns and adjectives (Pepperberg, 1990).

The family of sea lions (Otariidae), belonging to the pinniped family, are highly sociable amphibious animals. Sea lions actively communicate with each other through vocalizations (Schusterman & Van Parijs, 2003). A captive Steller sea lion (*Eumetopias jubatus*) can learn to associate human vocal commands with her behaviors with high accuracy (Sasaki et al., 2022; Sasaki & Kambara, 2024). Additionally, Sasaki et al. (2025) has suggested that she also has motor imitation ability based on human actions. On the other hand, Schusterman and Krieger (1984) reported spectacular results in language learning studies of California sea lions (*Zalophus californianus*), using hand signals to direct objects and behaviors. These studies suggest that sea lion families are capable of associative learning by using both visual and verbal commands.

Although researchers have accumulated knowledge regarding the ability of sea lions to connect human commands to referents, no studies have been conducted to determine the stage of development at which sea lions acquire the ability to associate human commands with referents. The weaning age of sea lions can vary significantly by year and population (Rosas-Hernández et al., 2024). In general, sea lion species in temperate regions are weaned at approximately one year of age (Trillmich, 1990). Due to their long weaning period, it is difficult for human experimenters (trainers) to intervene and train sea lion pups (i.e., babies) in captivity. Such ecological features did not allow us to study the command learning ability of sea lion pups.

We hypothesized that Steller sea lions are capable of associative learning using human commands even at the lactating pup stage, as they communicate immediately after birth through vocalization between mother and pup, and social learning through playing with other pups. It would also be interesting to determine which visual or auditory stimuli are dominant in associative learning in this species which are capable of both visual and auditory communication.

In this study, we focused on a Steller sea lion pup that was artificially nursed from 36 hours of age due to the lack of lactation from his mother, because recent studies have suggested that an adult Steller sea lion can understand vocal commands associated with behaviors (Sasaki et al., 2022; Sasaki & Kambara, 2024).

Steller sea lions and other sea lion species actively communicate with each other through vocalizations (Loughlin, 2009). There are also some reports that wild individuals engage in various social behaviors, including "boundary display," "neck bite," and "open mouth submission," to maintain social relationships and spatial territories (Gentry, 1974). The breeding season is from late May to early July (Pitcher & Calkins, 1981). Steller sea lion mothers and pups actively communicate with each other through their unique vocalizations (Loughlin, 2009). Pups learn various social behaviors by playing with other individuals of the same generation (Gentry, 1974).

We conducted an associative learning study of nine behaviors with human commands using hand signals and vocal commands in a Steller sea lion pup with no prior learning history. We then tested the accuracy rates in three conditions (i.e., when the hand signals and vocal commands were presented simultaneously, when only the hand signals were presented, and when only the vocal commands were presented). Specifically, we investigated (a) whether a lactating pup is capable of associative learning using human commands, and (b) whether there are differences in performance depending on the type of human commands presented (i.e., visual vs. auditory stimuli).

Method

We hypothesized that associative learning with human commands in Steller sea lion pups is possible. To verify this hypothesis, we conducted pre-training, behavior-shaping training, and command conditioning training.

The Subject

The subject was a male Steller sea lion pup named Kanata (domestic registration number in JAZA: 328). He was born at Kinosaki Marine World (Hyogo, Japan) on July 4, 2023. He weighed 22 kg at the time of birth. His mother is Hama, a Steller sea lion that can associate 50 words with each action (Sasaki et al., 2022) and understand two consecutive vocal commands (Sasaki & Kambara, 2024). Because no lactation was observed in Hama, we raised Kanata completely artificially from 36 hours of age.

Kanata was reared in three different animal houses, depending on his stage of growth. First, he lived in the 14 m^2 (5.7 m × 2.5 m) animal house (maternity room) from the start of artificial feeding until the age of 49 days old. Second, Kanata has lived in a 21 m^2 (7.8 m × 2.7 m) animal house from the next period to now. Third, during training and testing, Kanata was temporarily housed in an animal house, as described by Sasaki et al. (2022).

The Experimenters (Trainers)

The experimenters were six trainers (three males: A, B, and C; and three females: D, E, and F). Five of the trainers (Trainers A–E) had reared Kanata since the beginning of his artificial care. All trainers had at least two years of work experience in marine mammal husbandry and training. Trainer A was the first author of this article. He fed and trained Kanata at the highest frequency during the experimental period (227 sessions). Trainers B-E fed and trained at moderate frequency (210–221 sessions). The remaining trainer (F) raised Kanata from the age of 2 months. Thus, Trainer F's time with Kanata was shorter than that of the other trainers (85 sessions).

Rearing Methods

At 36 hours of age, Kanata was housed in the maternity room, separated from his mother, and artificial care was started. For the first month, approximately 10 staff members took turns spending 24 hours a day with Kanata, feeding him a special formula made by adding fish oil to dog milk as a substitute milk fed every two hours. The staff intended to support Kanata's growth as surrogate parents in other ways besides nursing. Therefore, when Kanata was actively moving, the staff played with him by rubbing his body or playing with a ball or a frisbee. Additionally, when Kanata slept, the staff slept with Kanata, who climbed on the staff's body.

After the first month, the staff gradually reduced the number of feedings per day because the amount of substitute milk Kanata could drink at once increased. Additionally, as Kanata came to sleep sustainably for longer periods of time without asking for substitute milk during the night, the staff stopped feeding late at night and only fed him five times during the day. Along with changes in the number of feedings, we reduced the number of trainers who nurtured Kanata to six (i.e., Trainers A–F). We shifted to a lifestyle in which we spent time with Kanata mainly during the daytime.

At 49 days of age, he began living with his mother (Hama) in preparation for living with three other Steller sea lions (one male and two females). At this point, we no longer spent time in the same room with Kanata except for feeding.

Training Procedure

Pre-training Phase

From 25 days old, we began the elementary training of Kanata. At this stage, we trained Kanata in the basics of targeting, stationing (i.e., waiting on the spot), and leading (i.e., moving after the trainer). As the first step, each trainer gave Kanata substitute milk when Kanata touched his nose on the trainer's hand. In this situation, Kanata was placed in front of each trainer. As the second step, each trainer immediately gave Kanata a whistle sound (or a vocalization “yoshi” that means “OK”) as bridging stimulus, when Kanata gently placed his nose on the trainer's hand. After the whistle sound (or the vocalization “yoshi” that means “OK”), the trainer immediately gave him milk. Each trainer repeated this procedure in the second step. Through this procedure, Kanata learned that the whistle (or the vocalization “yoshi” that means “OK”) was sounded when he performed a behavior desired by the trainer and was subsequently rewarded.

Behavior Shaping Phase

Since Kanata was three months old, trainers began training for shaping behaviors. Each training session lasted approximately 5 min. In the session, the trainers conducted five to ten sessions per day. The training was conducted by six trainers (see “The Experimenters” section). Each trainer shaped different types of behavior.

Using pre-training learning, behaviors were shaped through reinforcement with the sound of a whistle (or the vocalization “yoshi” that means “OK”) and substitute milk. We shaped Kanata’s behaviors in two ways: “targeting (including ruling)” and “innovative (including capturing).” In the “targeting” case, trainers reinforced that Kanata would touch his body parts to the trainer's hands or specific landmark objects. Trainers shaped the target behaviors by repeating the small steps to move specific body parts with the trainers’ induction. In the “innovative” case, the trainers shaped new behaviors by reinforcing the behaviors that Kanata performed spontaneously. From September 7, 2023 (when Kanata was three months old) to December 8, 2023 (when Kanata was six months old), the trainers shaped nine types of behavior using the methods described above.

Conditioning Phase: Conditioning Behaviors to the Human Commands

After shaping the target behaviors, we conditioned specific hand signals and vocal commands for each behavior. Each trainer who shaped each behavior conditioned the behavior to the human commands (i.e., hand signals and vocal commands). The training was repeated so that Kanata performed the target behavior immediately after receiving the human commands. When Kanata performed the correct behavior to the human commands, the trainers whistled and gave him milk. During training, the trainers provided both hand signals and vocal commands at the same time.

This procedure was repeated from time to time, as each behavior was shaped. When more than 80% of the correct behaviors were generated within a session for the simultaneous presentation of hand signals and vocal commands, trainers defined Kanata as having completed learning of the behavior and human command pair. Behaviors that had completed conditioning with human commands were shared with other experimenters (trainers) and were allowed to provide these commands to Kanata during the feeding time.

Hand signals, Vocal commands, and Kanata’s actions are shown in Table 1 and video 1 (https://osf.io/by389/?view_only=ac174ddc6ba04823a6c9888e31fa5aed).

Table 1

Learned Behaviors and Associated Commands (Hand Signals and Vocal Commands), Training Period, and Training Type of Each Behavior

Short name	Trainer's hand signal	Trainer's vocal command	Description of Kanata's behavior	Training period (Number of session)	Training type
Wave	Each trainer waves the trainer's hand.	bye-bye	Swinging his foreflipper.	Sep. 7 (1)	Innovative
Spin	Each trainer rotates the trainer's right hand.	omawari	Rotating on an axis perpendicular to the ground.	Sep. 25 to Oct. 4 (8)	Targeting
Leg	Each trainer swings the trainer's left index finger.	chin-chiro-rin	Swinging both hindflippers.	Oct. 13 to Oct. 22 (15)	Targeting
Back	Each trainer spreads both hands in front of the chest.	goron	Lying on his back.	Nov.4 to Nov.6 (5)	Innovative
Turn	Each trainer shows both trainer's palms.	oshiri	Turning around.	Nov. 9 (2)	Innovative
Pose	Each trainer taps the trainer's left thigh once.	pose	Standing still with the left foreflipper on the hip.	Dec. 4 (2)	Innovative
Down	Each trainer lowers the trainer's right hand to the ground.	huse	Lying on his face.	Nov. 10 to Dec. 10 (39)	Targeting
Bark	Each trainer shows the trainer's right fist.	Kanata	Barking once.	First half of Nov. (3)	Innovative
Hand	Each trainer shows the trainer's left hand in front of the cage.	dashite	Taking his right foreflipper out of the cage.	First half of Nov. (28)	Targeting

Note. The training period is the time from the beginning of the training of the shaping behavior to the completion of conditioning behavior with the hand signal or vocal command. Training type means the types of shaping behavior. See the second paragraph in the behavior shaping phase section (in task procedure) for details.

Testing Procedure

The trainers tested Kanata to examine his behavioral responses to the human commands (i.e., hand signals and/or vocal commands) given to him in the following three conditions:

Condition 1: Each trainer provided vocal commands and hand signals simultaneously.
Condition 2: Each trainer provided only hand signals.
Condition 3: Each trainer provided only vocal commands.

The test period was from February 6 to May 2, 2024. The tests were conducted on days when each experimenter (Trainers A–F) was working. As each experimenter (trainer) had a different schedule, the period to complete all sessions varied among the experimenters (trainers).

We conducted 10 sessions for each condition, for 30 sessions. Each session consisted of nine human commands (Table 1). Each human command was tried once in each session in a random order. We decided in advance which condition would be tested in each session by randomization (i.e., sometimes the same condition was tested for two consecutive sessions, and sometimes the different conditions were tested alternately). Each of these tests was conducted with each of the six trainers (180 sessions in total).

Each trainer provided Kanata with human commands when Kanata was in the animal house. The trainers bent down to look at Kanata's eyes. The trainer’s action was intended to prevent the trainer’s unconscious cueing (e.g., eye movements and small gestures) during the test.

When Kanata performed the correct behavior in response to the given human commands, the trainer whistled (or vocalized “yoshi” that means “OK”) and gave him substitute milk to drink for approximately two seconds. In the case of incorrect responses, we touched Kanata's mouth to make him still and gave him substitute milk for about two seconds because there was a risk that a chain of errors would occur if each trainer moved on to the next trial without any rewards.

Data Analysis

For each trial, Kanata’s responses to the human commands were coded as a binomial variable with “correct” or “incorrect” response. The experimenters (Trainers A to F) who conducted each session coded Kanata’s responses onsite. All test sessions were video-recorded. After the tests, we checked whether the recorded scores were correct using the videos as evidence. We compared the accuracy rates of human commands among the three conditions (Condition 1: vocal command and hand signal pairs, Condition 2: hand signals only, and Condition 3: vocal commands only) using a chi-square test. Since there were three multiple comparisons of the number of correct answers among conditions, the significance level was adjusted using the Bonferroni method (i.e., 0.05/3 = .0167). Therefore, only cases in which the p-value was less than .016 were considered significant for the performance comparisons among the three conditions. We also compared the correct response rates among the trainers’ genders using a chi-square test.

The accuracy rate (%) was calculated based on 90 trials per trainer (i.e., 10 trials for each of the nine types of human commands) when evaluating Kanata's performance by differences in the trainers, and 60 trials per type of human commands (i.e., 10 trials for each of the six trainers) when evaluating Kanata's performance by types of human commands. A binomial test was used to evaluate the number of correct versus incorrect responses to human commands in each condition. Because the trainers provided a total of nine human command types during the test in each condition, the probability of success (i.e., chance level) was set at 1/9 (11%). “The number of correct responses that would be significantly greater than the number of incorrect responses” would be 13 trials in the case of 60 trials (i.e., 21.7%) or 17 trials in the case of 90 trials for analyses of each human command and behavior pair in each condition (i.e., 18.9%). We set the significance level at 22% in the case of 60 trials and 18 % in the case of 90 trials for the analyses of each pair of the human command and behavior in each condition when evaluating Kanata’s performance using the binomial test.

All tests were two-tailed and were set at alpha = 0.05.

Ethical Notes

This study was conducted in accordance with the institutional ethics of the Kinosaki Marine World, a member of the Japanese Association of Zoos and Aquariums (JAZA). The experimental design of this study adhered to the animal welfare regulations of the JAZA and the guidelines for the treatment of animals in behavioral research and teaching by the Animal Behavior Society and Association for the Study of Animal Behavior (Buchanan et al. 2020). This study was approved by the ethical management branch of the facility (December 1, 2023). Steller sea lions at the facility receive meticulous daily care, including the necessary monthly clinical veterinary treatment, to ensure their health and well-being.

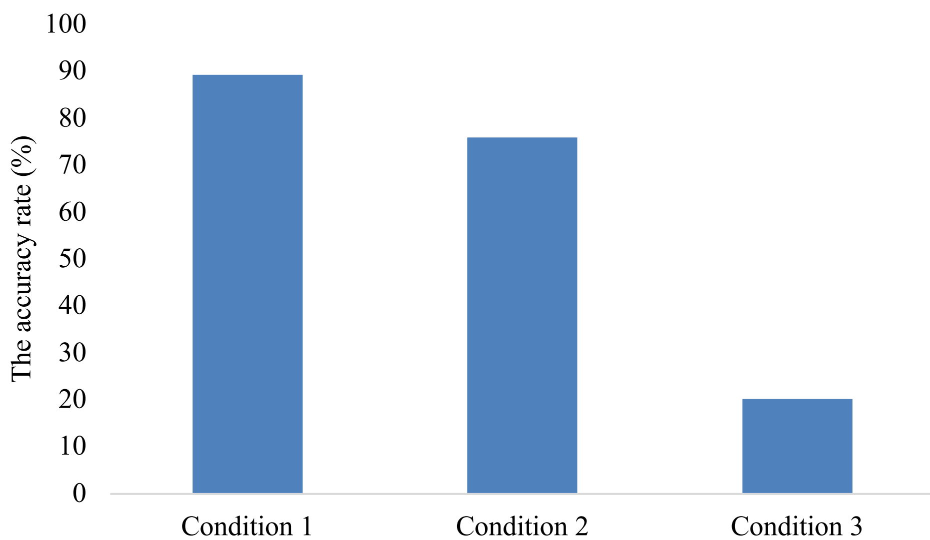
Results

Comparisons Between the Accuracy Rates of the Three Conditions

Figure 1 shows the overall accuracy rates for the nine types of human commands in each condition (i.e., Conditions 1, 2, and 3). The overall accuracy rates for Conditions 1, 2, and 3 were 89.4% (483/540), 75.9% (410/540), and 20.2% (109/540), respectively. Chi-squared tests showed that the accuracy rate of Condition 1 was significantly higher than that of Conditions 2 (X^2 (1, 1080)= 34.5, $p < .001$) and 3 (X^2 (1, 1080)=519.9, $p < .001$). Furthermore, the accuracy rate of Condition 2 was significantly higher than that of Condition 3 (X^2 (1, 1080)= 336.1, $p < .001$).

Figure 1

Mean Accuracy Rates in Three Conditions



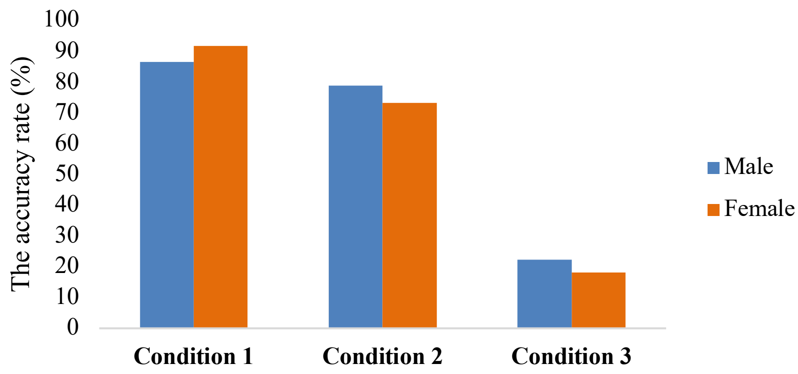
Note. Condition 1: Both vocal commands and hand signals; Condition 2: Hand signals only; Condition 3: Vocal commands only.

Comparisons of the Accuracy Rates by the Gender of the Trainer

Figure 2 shows the overall accuracy rate for the nine human commands in each condition, according to the trainer's gender. There was no significant difference in the accuracy rates of the trainer gender (female and male; Condition 1: X^2 (1, 540)=3.79, $p = .052$; Condition 2: X^2 (1, 540)= 2.29, $p = .130$; Condition 3: X^2 (1, 540)=1.39, $p = .238$, see Figure 2).

Figure 2

Relationship Between Trainer Gender and Accuracy Rate in Each Condition



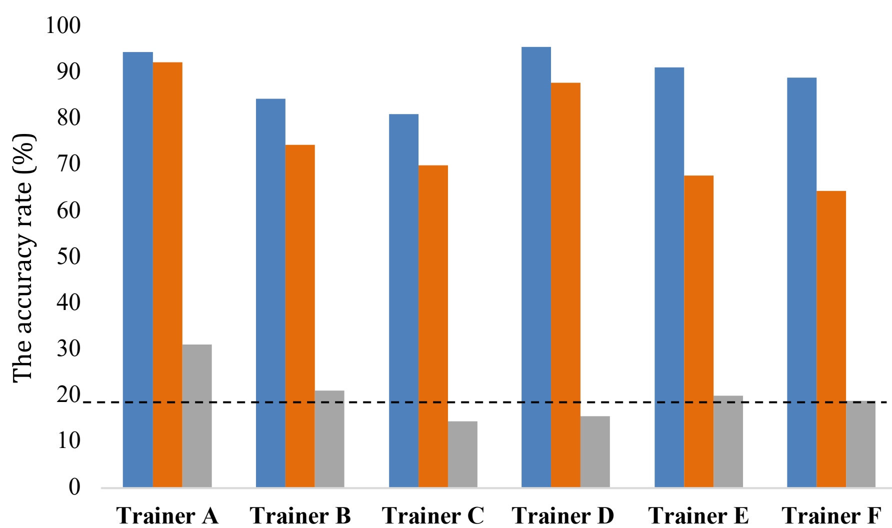
Note. Condition 1: Both vocal commands and hand signals; Condition 2: Hand signals only; Condition 3: Vocal commands only.

The Accuracy Rate of Kanata's Responses by the Trainers in Each Condition

Figure 3 shows the overall accuracy rate for the nine types of human commands in each condition for each trainer. Binomial tests for Kanata's responses to human commands revealed that the number of correct responses was significantly higher than chance for all six trainers in Conditions 1 and 2 ($p < .001$). In Condition 3, binomial tests for Kanata's responses to human commands revealed that the number of correct responses was significantly higher than chance, except for Trainer C, and D (Trainers A, B, E and F ; $p < .05$; Trainer C, and D; $p > .4$).

Figure 3

The Accuracy Rate for Each Trainer in Each Condition



Note. The dotted line indicates the significant revel. Condition 1: Both vocal commands and hand signals; Condition 2: Hand signals only; Condition 3: Vocal commands only.

The Accuracy Rate of Kanata's Responses by the Behavior Types in Each Condition

Table 2 shows the accuracy rates and incorrect response patterns for each of the nine human commands in each condition. Binomial tests for Kanata's responses to human commands revealed that the number of correct responses was significantly higher than the chance level for all nine human commands in Conditions 1 and 2 ($p < .001$). However, in Condition 3, binomial tests for Kanata's responses to human commands revealed that the number of correct responses was significantly higher than the chance level only for Leg, Bark, and Pose ($p < .01$).

Table 2

Kanata's Response to Each Human Command in Each Condition

Commands	Number of correct responses	Accuracy rate (%)	Binomial test	Number of incorrect responses	Classifications of incorrect responses			
					One of the other eight test behavior	Incomplete behavior	No response	Other responses
Condition 1: Both vocal commands and hand signals								
Spin*	47	78.3	$p < .001$	13	7		6	
Hand*	52	86.7	$p < .001$	8	7		1	
Leg*	47	78.3	$p < .001$	13	9	1	1	2
Down*	60	100.0	$p < .001$	0				
Back*	60	100.0	$p < .001$	0				
Turn*	46	76.7	$p < .001$	14	10		4	
Bark*	60	100.0	$p < .001$	0				
Pose*	56	93.3	$p < .001$	4	2	2		
Wave*	54	90.0	$p < .001$	6	4		2	
Condition 2: Hand signals only								
Spin*	42	70.0	$p < .001$	18	4		13	1
Hand*	51	85.0	$p < .001$	9	5		2	2
Leg*	35	58.3	$p < .001$	25	18		5	2
Down*	60	100.0	$p < .001$	0				
Back*	59	98.3	$p < .001$	1				1
Turn*	28	46.7	$p < .001$	32	14		16	2
Bark*	55	91.7	$p < .001$	5			5	
Pose*	38	63.3	$p < .001$	22	12	3	5	2
Wave*	43	71.7	$p < .001$	17	12		1	4
Condition 3: Vocal commands only								
Spin	7	11.7	n.s.	53	27	1	24	1
Hand	0	0.0	n.s.	60	32	2	25	1
Leg*	15	25.0	$p < .001$	45	24		19	2
Down	7	11.7	n.s.	53	30		22	1
Back	0	0.0	n.s.	60	30	1	29	
Turn	2	3.3	n.s.	58	28		28	2
Bark*	55	91.7	$p < .001$	5	2		3	
Pose*	15	25.0	$p < .001$	45	24	1	20	
Wave	8	13.3	n.s.	52	30	1	20	1

Note. * indicates that the accuracy rate was significantly higher than chance.

Discussion

In the present study, we tested whether a lactating Steller sea lion pup could learn the associations between human commands and behaviors. When trainers simultaneously gave him hand signals and vocal commands (Condition 1), Kanata's performance was significantly higher than the other conditions (Condition 2: only hand signals and Condition 3: only vocal commands). Additionally, Kanata's performance was significantly higher in Condition 2 than in Condition 3. Although Kanata's performance was stable for all human commands in both Conditions 1 and 2, it was significantly lower for almost all commands in Condition 3. Moreover, Kanata's performance exceeded the significance level for all trainers in Conditions 1 and 2. However, his performance was below the significance level for two of the six trainers in Condition 3. Kanata's performance did not differ based on the trainer's genders in any condition. These results provide the first evidence that a Steller sea lion pup can learn the associations between human commands and behaviors. In addition, we also provide the first findings that the associations between human commands and behaviors could be formed in both visual and auditory modalities, but the learning association could be dominant in visual information compared to auditory information.

Kanata showed significantly high performance in all nine types of human commands in Condition 1, suggesting that he could learn the associations between the human commands and behaviors well through the training described in a chapter of "Training Procedure." His performance was stable across all six trainers. This finding indicates that Kanata produced his responses by discriminating human commands, not by unconscious cues and peculiar habits unique to each trainer. Schusterman and Krieger (1984) reported that adult California sea lions could learn associations between hand signals and behaviors. Sasaki et al. (2022) and Sasaki and Kambara (2024) reported that an adult Steller sea lion can learn associations between human vocal commands and behaviors. Although these previous studies have suggested that the sea lion family is capable of association learning using human commands, they did not find the ability to learn these associations at a very young age. Kanata's excellent performance in Condition 1 suggests that a Steller sea lion is capable of associative learning using human commands, even at a very early stage of life.

Kanata's performance in Condition 1 was significantly higher than that in the other two conditions, suggesting that Kanata formed the associations between the human commands and behaviors through both visual and auditory modalities. This finding may reflect the ecological characteristics of this species, who use both visual and auditory stimuli to communicate between individuals (Charrier et al., 2022). Wierucka et al. (2018) summarizes that animals often employ multiple sensory modalities for their communication. The multimodal signal demonstrates more effective robustness by maximizing the information in each channel and minimizing the effects of partial loss (Ay et al., 2007). Multimodal communication may enhance message reliability and provide more detailed information to the receiver (Hebets & Papaj, 2005).

Interestingly, Kanata performed very well in Condition 2, where only hand signals were provided. Both the accuracy rates for all trainers and all behaviors were significantly higher than the chance level, similar to the results for Condition 1. Kanata's high performance in Condition 2 suggests that Kanata obtained information mainly based on hand signals, even though hand signals and vocal commands were given simultaneously during the conditioning phase. These results support the findings reported by Shusterman and Krieger (1984) in adult California sea lions and suggest that a Steller sea lion pup can learn associations based on visual information, similar to adult sea lions.

In contrast to the excellent performances in Conditions 1 and 2, which included visual information (hand signals), Kanata's performance was significantly lower in Condition 3, which provided only vocal commands. Wild Steller sea lions communicate using their vocalizations (Campbell et al., 2002; Loughlin, 2009). Moreover, our previous studies in captivity have also reported excellent vocal discrimination in an adult female Steller sea lion, Hama (Sasaki et al., 2022). The results shown by Kanata in this study were in contrast to the findings of the previous studies mentioned above.

Although Kanata’s performance was stable for all trainers in Conditions 1 and 2, his performance for two trainers (Trainers C, and D) was below the significance level only in Condition 3. There was no common gender or frequency of training among these three trainers, making it difficult to make a reference from this study to the cause of the differences in performance. No such trainer-dependent performance differences have been observed in previous studies on other pinnipeds (captive walrus (*Odobenus rosmarus divergens*); Endo et al., 2020., Steller sea lion; Sasaki et al., 2022). Vocal commands may be a stimulus for the Steller sea lion pup that is susceptible to individual differences. Although we do not have strong evidence to support this hypothesis, it may be related to the fact that vocal communication during puphood in Steller sea lions occurs between a limited number of closely related individuals (i.e., mothers and children).

Kanata's performance in Condition 3 lacked stability, even when compared across behavior types. In Condition 3, only three of the nine commands (i.e., “Kanata,” meaning “bark” behavior, “Chin-chillo-rin,” meaning “leg” behavior, and “Pose,” meaning “pose” behavior) exceeded the significance level for the accuracy rate. The command with the highest accuracy rate was “Kanata,” which means “bark” behavior, at 91%, nearly three times higher than the correct rate for “Chin-chillo-rin,” which means “leg” behavior, and “Pose,” which means “pose” behavior. Moreover, the most common incorrect response in Condition 3 was “bark” (169/431), except for no response (190/431) (see Table 2). This response tendency for vocal commands to be prominent in the “bark” is common to all commands. This response is thought to reflect the ecology of this species. Campbell et al. (2002) reported that Steller sea lion mothers and pups communicate through their vocalizations and recognize each other by their calls. Therefore, when mothers call to their pups, the pups respond immediately. The trainers intimately treated Kanata shortly after his birth. Therefore, it is quite possible that a similar response, like the mother-infant calling, could be observed between trainers and Kanata in Condition 3. Based on this background, the response characteristics may have resulted in a “bark” response to many vocal commands given by the trainers.

Considering the overall results of all conditions, the results of this study suggest that Kanata could learn the associations between human commands and behaviors through both visual and auditory information but was better at understanding visual information than auditory information. The findings of this study would be also supported by previous human studies. Previous findings have shown that humans recognize the associations between linguistic and visual information more correctly or quickly than the associations between linguistic and auditory information (e.g., Cosper et al., 2022; Kambara et al., 2013) or between linguistic information in a first language (L1) and linguistic information in a second language (L2; e.g., Liu et al., 2021; Obinna et al., 2025) after associative learning. These previous studies have focused on the differences between modalities of referents (meanings), whereas we compared the differences among modalities of commands in this study. Language learning, which is cross-modal but visually dominant, was reported by Kambara et al. (2013) in their study of associative learning between unfamiliar words in L1 and perceptual features as meanings. Interestingly, the features of associative learning reported in humans have been suggested in Steller sea lion pups across species. Although Steller sea lions can communicate using both visual and auditory stimuli, there may be qualitative differences between communication using visual and acoustic information.

Prior research on wild Steller sea lions suggests that visual communication involves a greater variety of contexts than vocal communication does. Vocal communication in Steller sea lions has been reported in the context of a threat, courtship, and recognition between mother and pup (Campbell et al. 2002; Loughlin, 2009). On the other hand, visual communication may be used in the complex context of social behavior. Gentry (1974) has identified 98 different types of motor elements in his observations of play bouts between individual pups in the field. He also reported that through interindividual play, pups flexibly changed their behavior and exhibited a variety of behaviors that could not be defined individually. In light of the above findings, visual dominance in command learning may be related to the development of visual communication, which has more complex and diverse contexts than vocal communication does.

This study provides the first evidence that a Steller sea lion can learn human commands cross-modally from the pup stage. Although Kanata was able to learn the associations between some vocal commands and behaviors, he was better at learning the associations between hand signals and behaviors. At this time, we cannot discuss whether the visually dominant associative learning observed in this study is a characteristic of Steller sea lions, a characteristic of pups, or an individual difference. This is because no previous study has compared these characteristics. Kanata was the best individual to demonstrate visual dominance results in associative learning. He had no prior training experience with specific human commands as captive adults do. Kanata's situation is very rare, but if individuals that are artificially nursed like Kanata appear in the future, comparative studies may be able to address the above questions. It would also be interesting to investigate whether Kanata can learn many types of human vocal commands in the future, such as the female Steller sea lion Hama reported by Sasaki et al. (2022). We will continue to explore the potential and characteristics of associative learning abilities using human commands in captive Steller sea lions.

Acknowledgments

We would like to thank all staff involved in the artificial nursing of Kanata. We would like to thank Mr. Tatsuo Sawada of the Oita Marine Palace Aquarium “Umitamago” and Mr. Takaya Maruno of the Katsurahama Aquarium for their valuable advice on the artificial nursing of Kanata. We would also like to thank Ms. Yumiko Nagaoka of the Sumida Aquarium for donating milk and salmon oil and for sharing valuable breeding data. We would also like to thank Dr. Akitsugu Konno of Azabu University for his valuable advice in this study. The last author was supported by KAKENHI Grant-in-Aid for Scientific Research (B). Finally, we would like to thank the Steller sea lions at the Kinosaki Marine World for providing us with unlimited energy.

References

Ay, N., Flack, J., & Krakauer, D. C. (2007). Robustness and complexity co-constructed in multimodal signalling networks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362,* 441–447. https://doi.org/10.1098/rstb.2006.1971

Campbell, G. S., Gisiner, R. C., Helweg, D. A., & Milette, L L. (2002). Acoustic identification of female Steller sea lion (*Eumetopias jubatus*). *Journal of Acoustic Society of America, 111*, 2920–2927. https://doi.org/10.1121/1.1474443

Charrier, I., Pitcher, B., & Harcourt, R. (2022). Mother–pup recognition mechanisms in Australia sea lion (*Neophoca cinerea*) using uni- and multi-modal approaches. *Animal Cognition, 25*, 1019–1028. https://doi.org/10.1007/s10071-022-01641-5

Cosper, S. H., Männel, C., & Mueller, J. J. (2022). Mechanisms of associative word learning: Benefits from the visual modality and synchrony of labeled objects. *Cortex, 152*, 36–52. https://doi.org/10.1016/j.cortex.2022.03.020

Endo, S., Kawaguchi, N., Shimizu, Y., Imagawa, A., Suzuki, T., Ashikari, H., Wakai,Y.,& Murayama, T. (2020). Preliminary study of discrimination of human vocal commands in walrus *(Odobenus rosmarus divergens*). *International Journal of Comparative Psychology, 33*, 1–11. https://doi.org/10.46867/ijcp.2020.33.00.03

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee: A standardized system of gestures provides a means of two-way communication with a chimpanzee. *Science*, *165*(3894), 664–672. https://doi.org/10.1126/science.165.3894.664

Gentry, R. L. (1974). The development of social behavior through play in the Steller sea lion. *American Zoologist*, *14*(1), 391–403. https://doi.org/10.1093/icb/14.1.391

Griebel, U., & Oller, D. K. (2012). Vocabulary learning in a Yorkshire terrier: Slow mapping of spoken words. *Plos One, 7*, e30182. https://doi.org/10.1371/journal.pone.0030182

Hebets, E., & Papaj, D. (2005). Complex signal function: Developing a framework of testable hypotheses*. Behavioral Ecology and Sociobiology, 57*(3), 197–214. https://doi.org/10.1007/S00265-004-0865-7

Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition,16*(2), 129–219. https://doi.org/10.1016/0010-0277(84)90003-9

Kambara, T., Tsukiura, T., Shigemune, Y., Kanno, A., Nouchi, R., Yomogida, Y., & Kawashima,R. (2013). Learning-dependent changes of associations between unfamiliar words and perceptual features: A 15-day longitudinal study. *Language Sciences, 35*, 80–86. https://doi.org/10.1016/j.langsci.2012.05.001

Liu, X., Horinouchi, H., Yang, Y., Ando, M., Obinna, UJ., Namba, S., & Kambara, T. (2021). Pictorial referents facilitate recognition and retrieval speeds of associations between novel words in a second language (L2) and referents. *Frontiers in Communication,* 6, https://doi.org/10.3389/fcomm.2021.605009

Loughlin, T.R. (2009) Stellar sea lion: Eumetopias jubatus. *Encyclopedia of Marine Mammals* (2ndEdition),1107-1110. https://doi.org/10.1016/B978-0-12-373553-9.00253-4

Pepperberg, I. M. (1990). Cognition in an African gray parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. *Journal of Comparative Physiology, 104*, 41–52. https://doi.org/10.1037/0735-7036.104.1.41

Obinna, U., Kabir, R. S., Yoshio, M., Yanamoto, D., & Kambara, T. (2025). Effects of pictorial referents on the associative learning of spoken word forms in a Nigerian language: An experimental investigation with L1 Japanese users. *Language, Cognition and Neuroscience*, 1–13. https://doi.org/10.1080/23273798.2025.2509548

Pitcher, K. W., & Calkins, D. G. (1981). Reproductive biology of Steller sea lions in the Gulf of Alaska. *Journal of Mammalogy, 62*, 599–605. https://doi.org/10.2307/1380406

Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioral Processes, 86*, 184–195. https://doi.org/10.1016/j.beproc.2010.11.007

Rosas-Hernández, M. P., Hernández-Camacho, C. J., Aurioles-Gamboa, D., & Trites, A. W. (2024). Age at weaning of California sea lions depends on colony latitude. *Journal of Experimental Marine Biology and Ecology*, *581*, 152059. https://doi.org/10.1016/j.jembe.2024.152059

Sasaki, M., Chiba, M., Ito, E., Tsutsumi, K., Ito, K., & Sunobe, T. (2022). Human vocal commands verify audiodiscrimination ability in the Steller sea lion *Eumetopias jubatus*. *International Journal of Comparative Psychology*, *35*(1), 1–11. https://doi.org/10.46867/ijcp.2022.35.5615

Sasaki, M., & Kambara, T. (2024). A Case Study of Spontaneous Category Formation and Behavioral Expression in a Language-Trained Steller Sea Lion *Eumetopias jubatus*. *International Journal of Comparative Psychology, 37*, 1–14. http://doi.org/10.46867/ijcp.6592

Sasaki, M., Kinoshita, H. & Konno, A. (2025). Do as I do imitation in a steller sea lion Eumetopias jubatus. *Animal Cognition, 28*, 49. https://doi.org/10.1007/s10071-025-01971-0

Schusterman, R. J., & Krieger, K. (1984). California sea lions are capable of semantic comprehension. *The Psychological Record, 34*, 3–23. http://doi.org/10.1007/BF03394849

Schusterman, R. J., & Van Parijs, S.M. (2003). Pinniped vocal communication: An introduction. *Aquatic Mammals*, 29, 177–180. https://www.aquaticmammalsjournal.org/index.php

Trillmich, F. (1990). The behavioral ecology of maternal effort in fur seals and sea lions. *Behaviour*, *114*(1-4), 3–20.

Wierucka, K., Pitcher, B., Harcourt, R., & Charrier, I. (2018). Multimodal mother–offspring recognition: The relative importance of sensory cues in a colonial mammal. *Animal Behaviour, 146,* 135–142. https://psycnet.apa.org/doi/10.1016/j.anbehav.2018.10.019

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

Submitted: *December 20th, 2025*
Resubmitted: *May 2nd, 2025*
Accepted: *May 5th, 2025*