Training Experimentally Naive Seals for Vocal Learning Experiments

Diandra Duengen[1,2] *and Andrea Ravignani*[1,3,4]

Abstract

Harbor seals (*Phoca vitulina*) are a common zoo species that show a scientifically valuable propensity for vocal learning. Under human care, the seals can be trained to associate vocalizations with cues. This ability is termed vocal usage learning and is characterized by learning to use a vocalization in a specific context. Among mammals, seals are prime candidates to investigate vocal learning. Yet, only a handful of reports exist on harbor seal vocal learning abilities and even less document how these were trained or tested. Here, we investigate how, and if, two experimentally naive harbor seals under human care, can be trained to participate in scientific experiments. We describe the training and testing of two seals in two basic vocal learning experiments. We trained the animals to vocalize upon the presentation of a discriminative stimuli (S_D) through operant conditioning methods and tested their abilities to i) vocalize and refrain from vocalizing on two distinct S_D's, and ii) produce two different vocalizations upon the presentation of two different S_D's. Both seals learned the tasks: the first task was achieved within 118 trials (22 errors to criterion) and 220 trials (40 errors to criterion), the second task within 480 trials (158 errors to criterion) and 380 trials (94 errors to criterion), respectively. Our results confirm that harbor seals are capable of vocal usage learning and further suggest that associating individually distinct vocalizations with different S_D's may be more cognitively demanding than vocalizing and being silent on S_D.

Keywords: Animal training, vocal learning, vocal shaping, pinniped, *Phoca vitulina*

Introduction

A single case study found that a male harbor seal can be trained to vocalize on a discriminative stimulus (S_D) (Schusterman, 2008). In his review of vocal learning in pinnipeds, Schusterman (2008) for the first time described the fruitful training of vocal variability in a harbor seal. He brought one vocalization under stimulus control and selectively shaped it towards a novel 'wa-wa' sound (Goncharova et al., 2024; Schusterman, 2008). Thereby, his case study demonstrates vocal *usage* learning: the ability to use a vocalization, be it learned or innate, in a novel context as a result of experience (Janik & Slater, 2000; Jarvis, 2019). The report provides a valuable foundation to document vocal learning and shaping, potentially inspiring further research. A remarkable case of mammalian vocal

production learning, plasticity in sound production, was a captive male harbor seal, who in the 1980s showed untrained imitation of human speech (Duengen et al., 2023; Ralls et al., 1985). A later attempt to replicate vocal imitation in this species failed (Moore, 1996). Harbor seals are popular in zoos, amongst others for their comparatively easy husbandry and trainability. This, taken together with their propensity for vocal learning, makes them a good study organism. Beyond the contribution to research, the participation in (e.g., acoustic) experiments can enrich individuals, thereby improving animal welfare (Ortiz et al., 2020).

Under human care, vocal usage learning can be tested through operant conditioning experiments (Adret, 1993; Janik & Slater, 2000; Lattenkamp et al., 2018; Seyfarth & Cheney, 2010; Stoeger & Baotic, 2021; Vernes et al., 2021), in which animals are taught to vocalize upon the presentation of an arbitrary S_D (e.g., a hand sign, see Table 1). Vocal learning via operant conditioning was found in a closely related seal species, the grey seal (*Halichoerus grypus*) (Shapiro et al., 2004; Stansbury et al., 2015). Dissecting vocal learning into its complex features - from 'simple' vocalizing upon the presentation of an S_D to vocalizing and withholding vocalizations upon different S_D's, to producing different vocalizations upon other S_D's - requires a chain of trained behaviors (Janik & Slater, 2000; Shapiro et al., 2004; Vernes et al., 2021).

Building on published work in other species (Shapiro et al., 2004; Vergara & Barrett-Lennard, 2017), we describe the training and testing process to probe vocal learning in two harbor seals. Our approach combines classical training methods (e.g., operant conditioning through positive reinforcement, associative learning, shaping by successive approximations) and randomized trial orders (Gellermann, 1933). We prepared two experimentally naive harbor seals for a series of vocal learning experiments (Duengen, under review) by training them to vocalize, withhold from vocalizing, shaping a novel vocalization type and bringing two different vocalizations under stimulus control.

Material & Methods

Setting

The seals were housed in a 230,000-liter freshwater pool in a 300 m^2 enclosure at Zoo Cleves, Germany. Two of the six resident harbor seals were selected to participate in this study as they were vocally active individuals. Seal E was a 6 years old female born at Zoo Osnabrück, Germany. Seal J was a 16 years old male born at Zoo Duisburg, Germany. A third, vocally-active seal dropped out of the experiment due to a severe eye condition. All seals were previously trained to participate in shows, which included behaviors such as object retrieval or flipper slapping. Shows did not take place during the period of the study. Training consisted of general training, conducted by zoo staff, and research training, conducted by the experimenter. All training took place in the main enclosure, where all seals resided, up to three times/week. Seals were tested separately, up to five times per week in the research enclosure, a partitioned area attached to the main enclosure. Training and testing was conducted during visitor hours.

Procedure and Training

Operant conditioning was facilitated through positive reinforcement, where a correct response was immediately followed by a bridge stimulus (here, a whistle) and fish as a primary reinforcer. Incorrect behavior resulted in a Least Reinforcing Scenario (neutrality for ca. 3 seconds). Positive reinforcement was delivered through pre-cut pieces of mackerel (*Scomber scrombus)* and herring (*Clupea harengus*).

We distinguish between *Training* and *Testing* (see Fig. 1). *Training* sessions were conducted with all seals in the main enclosure, a phase in which multiple behaviors were trained. These sessions served to establish novel behaviors and maintain known behaviors. Establishing novel behaviors was needed for a series of vocal learning experiments (Duengen, under review) and consisted of e.g., stationing and targeting behaviors, as well as ‘go into the water’, and ‘follow the trainer’. Much of this training time included habituating the animals to a novel research enclosure, where testing would later take place (see description below). Training novel behaviors was facilitated through capturing (e.g., ‘go into the water’), or luring (e.g., ‘follow the trainer’), or shaping (‘target’). In *Testing*, the seals were exposed to two dedicated tasks in a separate research enclosure. These tasks consisted of two different experiments, *Experiment 1: Vocalize vs. Silence* and *Experiment 2: Produce Different Vocalizations Upon Different* S_D*s*.

<u>Vocal Training</u>

Seal E and Seal J would occasionally vocalize during general training sessions in April and August 2021. This behavior was captured by opportunistically reinforcing the vocal behavior with fish to increase its occurrence. Seal E offered one vocalization while Seal J offered two different vocalizations. Seal E’s second vocalization was shaped through successive

approximations: The initial vocalization was rewarded whenever it aurally differed from the original vocalization and was produced with an opened muzzle. This led to the formation of a second type of vocalization (E2). When the seals produced vocalizations, we reinforced these with fish and presented a respective hand sign to facilitate the formation of an association between the vocal behavior and the hand sign (S_D, see Tab.1, Fig. 1, 2).

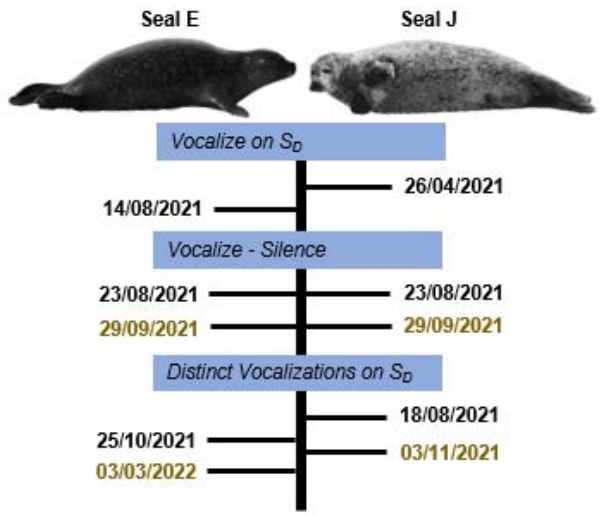

Fig. 1: Illustration of the training (black) and testing (brown) timeline for seal E (left) and seal J (right).

We assumed that the seal would have formed a first association when they responded to approx. 50% of hand signs with the desired behavior. As a next step, we occasionally reinforced the seals for not vocalizing, i.e., for silence. We then trained the seals to associate this silence with a hand sign until they responded correctly to approx. 50% of hand signs (see Fig. 1, Tab. 1). The association of these behaviors and their respective S_Ds were interspersed in the seals' daily research training sessions. During this stage, only the number of training sessions was noted, not the number of each cued behaviors. This opportunistic training approach resulted from the circumstance that at this stage of the study, the seals could not be separated, and a consistent assistant was not available.

Tab. 1: Description of visual SDs used to elicit the desired behaviors of the seals

Behavior	**Seal**	**S_D**
Vocalize	E & J	Hand next to the trainer's right ear
Silence	E & J	Stretched arm in front of the trainer's chest with open hand
Vocalization E1	E	Hand next to the right ear (see *Vocalize*)
Vocalization E2	E	Pointed finger rolling in 90-degree angle in front of the trainer
Vocalization J1	J	Right hand and arm above the trainer's head
Vocalization J2	J	Full hand in front of the trainer's mouth with a slight forward movement

Research Enclosure Training

In order to facilitate one-on-one testing sessions with the seals, a research enclosure was built. This enclosure accessible through a gate to the main enclosure. Separating the seals and working with single individuals in the closed research area was trained through successive approximations to promote a gentle transition from group training to individual training. Initially, the seals were acquainted with the novel enclosure through occasional feeding at the gate. Incrementally, the seals were asked to follow into the area, partially through luring. We took small steps, so that after a short habituation phase and before the seals showed any signs of nervousness, they were send into the pool (adjacent to the research enclosure). Sending into the pool served as a secondary reinforcement. The gate was kept open at all times during this stage. Once the seals' stay in the research enclosure reached a couple of minutes without any signs of distress, closing the gate was initiated. Here, one trainer rewarded the seal's behavior (remaining rested), while another trainer closed and opened the gate in very small increments. After approximately 5 minutes of remaining rested without any signs of distress, we considered the seals comfortable with the research enclosure.

Testing

All testing occurred in the separate research enclosure. Trials alternated in a predetermined randomized order, which featured a counterbalanced number of trials and ≤ 3 trials of the same type in a row (Gellermann, 1933; Jadoul, Duengen, et al., 2023). Once the seal was stationed, a trial began and the trainer presented the hand sign for that trial (Fig. 2, Tab 1).

Fig. 2: Illustration of the experimental Setup. The experimenter stands in front of the seal and asks for a pre-determined behavior (here: Vocalization J1, see Tab. 1). Responses were noted on the sheet, which contained the randomized order of behaviors (Gellermann, 1933; Jadoul, Duengen, et al., 2023).

In rare cases, when an animal was distracted (e.g., due to disturbance by zoo visitors) or vocalized incessantly, the trainer repeated the trial. The decision to repeat a trial was taken immediately by the trainer. Testing was conducted in the months of September and October 2021, and March and April 2022 (see Fig. 1).

Experiment 1: Vocalize -. Silence

Sessions consisted of 20-50 trials/session, depending on the animal's motivation; Sessions could be terminated when the animal lacked motivation, e.g. was distracted, chewed on fish rather than swallowing, left the testing area, etc. Trials were counterbalanced between presenting the S_D for vocalizing and withholding from vocalizing. The learning criterion was reached when the animal responded ≥ 80% correctly in four consecutive sessions.

Experiment 2: Produce Different Vocalizations Upon Different S_Ds

Each session consisted of 20 trials. Trials were counterbalanced and interchanged between asking seal E for vocal type E1/E2 and seal J for J1/J2. The learning criterion was reached when the animal responded ≥ 80% correctly in four consecutive sessions.

Experiment 2 started when the seals had successfully completed *Vocalize -. Silence* (Fig. 1). Before starting Experiment 2, seal E was trained on a second vocalization. This vocalization, E2, was established by selectively shaping and reinforcing the initial vocalization (E1) over a period of roughly four months, see section Vocal Training

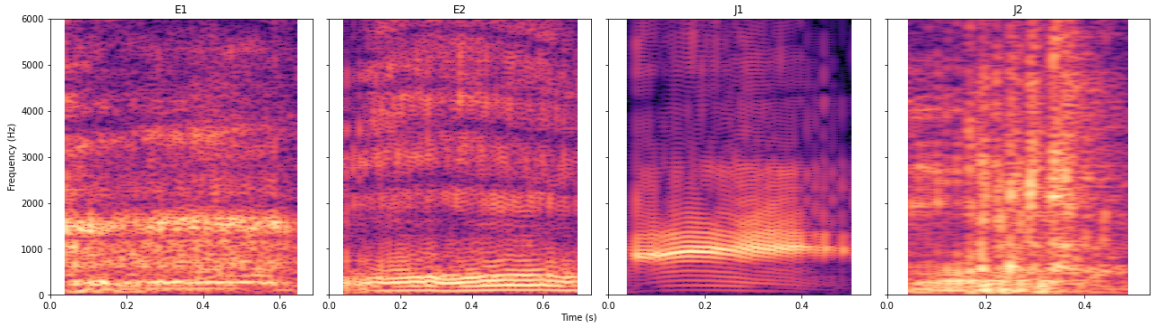

Fig. 3: Spectral representation of the four vocalizations: E1, E2, J1 and J2. Frequency (Hz) on the y-axis, time (s) on the x-axis. Spectrogram was obtained with the Parselmouth package in Python (version 0.4.1, Praat 6.1.38; window length=0.02s, dynamic range=70 dB, max. frequency=7000 Hz (Boersma & Weenink, 2022; Jadoul, De Boer, et al., 2023; Jadoul et al., 2018)).

<u>Documentation</u>

During testing, responses were noted on template response sheets (see Fig. 1). A subset of training and testing sessions were also audio and video recorded (Canon Legria HF25 camcorder). Audio recordings were obtained via a Zoom H6 digital recorder connected to a

Sennheiser ME-67 unidirectional microphone (frequency response 40 – 20.000 Hz +- 2.5 dB) covered by a foam windshield.

Data Analysis

We annotated audio recordings in Praat v6.1.38 (Boersma & Weenink, 2022). Relevant acoustic parameters of vocalizations were then batch-extracted with the Parselmouth package (a Python library for Praat, v0.4.1; Jadoul, De Boer, et al., 2023; Jadoul et al., 2018) in Python. We chose acoustic parameters that were previously analyzed in harbor seal vocalizations, i.e., duration, spectral center of gravity, dominant frequency, percentage voiced, and median harmonicity (Raimondi, under review, see Tab. 2; Van Parijs et al., 2000).

Tab. 2: Definition of the acoustic parameters and description of the extraction procedure

Name of Acoustic Parameter	**Description of Acoustic Parameter**
Duration (s)	Length of the vocalization in time.
Spectral center of gravity (Hz)	The weighted average frequency of the vocalization's spectrum. The spectrum was calculated with Praat's "Sound: To Spectrum", default parameters. Then, the center of gravity was calculated using "Spectrum: Get center of gravity", default parameters.
Percentage voiced (%)	The pitch was calculated with Praat's "Sound: To Pitch" with default parameters, except pitch floor 50 and pitch ceiling 400). The number of voiced frames was divided by the number of all frames.
Median harmonicity (dB)	Harmonicity was calculated with Praat's "Sound: To Harmonicity (cc)", default parameters except minimum pitch 50). After this the median of the harmonicity values were calculated.
Dominant frequency (Hz)	The spectrum was calculated with Praat's "Sound: To Spectrum" (default parameters). Then, we selected the frequency bin with the highest power.

Results

Experiment 1: Vocalize - Silence

Seal E completed Experiment 1 within 5 sessions i.e., 118 trials (23.6 trials/session ± 4.5) with 22 errors to criterion, the amount of false trials until reaching the learning criterion. Seal J completed Experiment 1 within 6 sessions i.e. 220 trials (36.6 trials/session ± 9.4) with 40

errors to criterion (see Fig. 4, 5 and 6). Before testing, Seal E had 9 preceding training sessions, and seal J had 15 preceding training sessions.

Experiment 2: Produce Different Vocalizations Upon Different S_Ds

Seal E completed Experiment 2 within 480 trials (24 sessions with 20 trials/session) with 158 errors to criterion. Seal J completed Experiment 2 within 380 trials (19 sessions with 20 trials/session) with 94 errors to criterion (see Fig. 4, 5 and 6). Four sessions were omitted from analysis as these were terminated early by seal J after a few trials. Before testing, Seal E had 47 preceding training sessions; this number includes the extra training she required to establish a novel vocalization type. Seal J had 23 preceding training sessions.

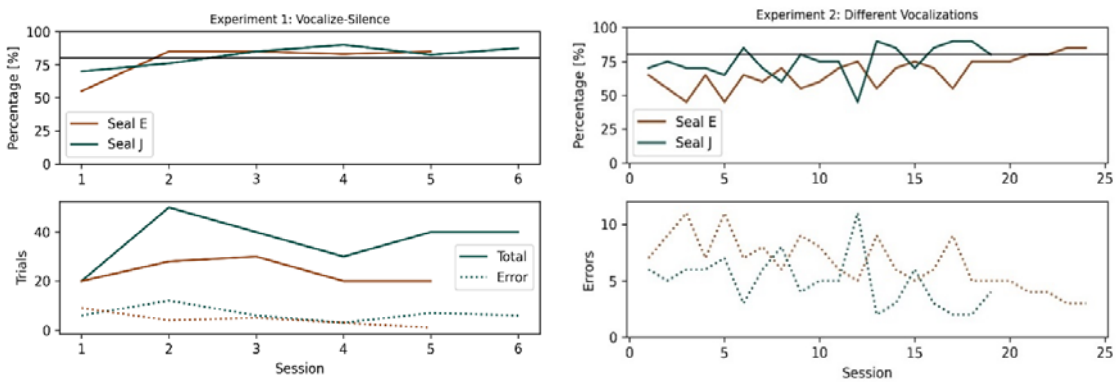

Fig. 4: The upper graphs depict the learning curves of seals E (brown) and J (dark green). The solid line at 80 % on the y-axis marks the learning criterion's threshold. The lower graph shows the seals' errors to criterion (dotted line). Note that the overall trial number per session (solid line) varied for Experiment 1, but was consistent with 20 trials/session in Experiment 2.

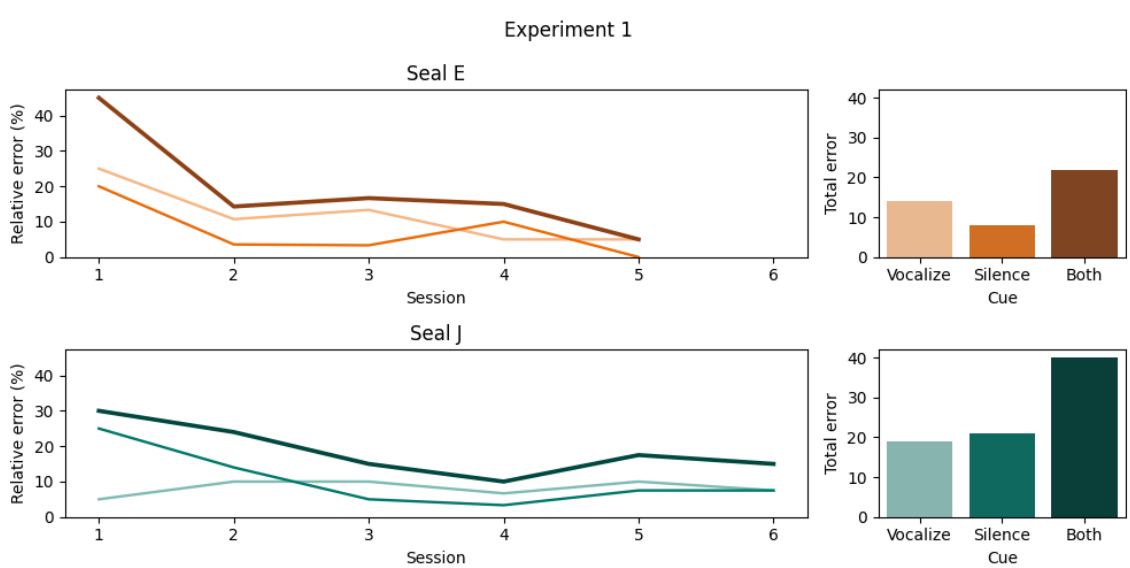

Fig. 5: The lineplot graphs (left, one per seal, seal E: shades of brown, seal J: shades of blue) depict the error to

criterion over each session for each S_D (Vocalize vs. Silence) separately. The barplots (right, seal E: top, seal J: bottom) show the overall scores for each S_D over all sessions.

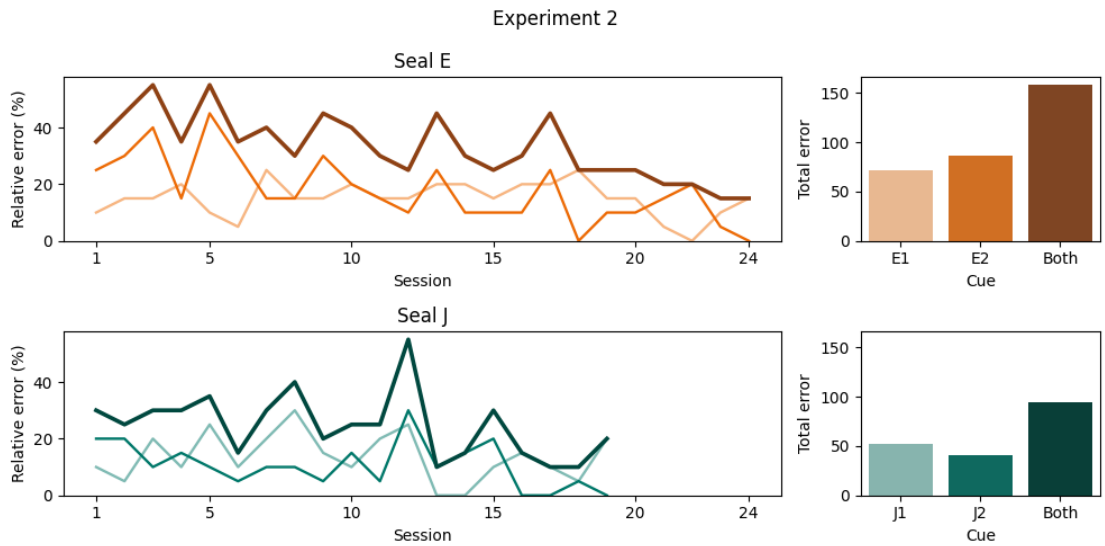

Fig. 6: The lineplot graphs (left, one per seal, seal E: shades of brown, seal J: shades of blue) depict the error to criterion over each session for each S_D (Vocalize vs. Silence) separately. The barplots (right, seal E: top, seal J: bottom) show the overall scores for each S_D over all sessions.

Discernibility of Vocalization Types

To confirm that the seals produced different vocal types (E1 vs. E2 and J1 vs. J2) and that the types were not mere human subjective categories, we inspected the acoustic parameter composition between call types and within seals (Fig. 7).

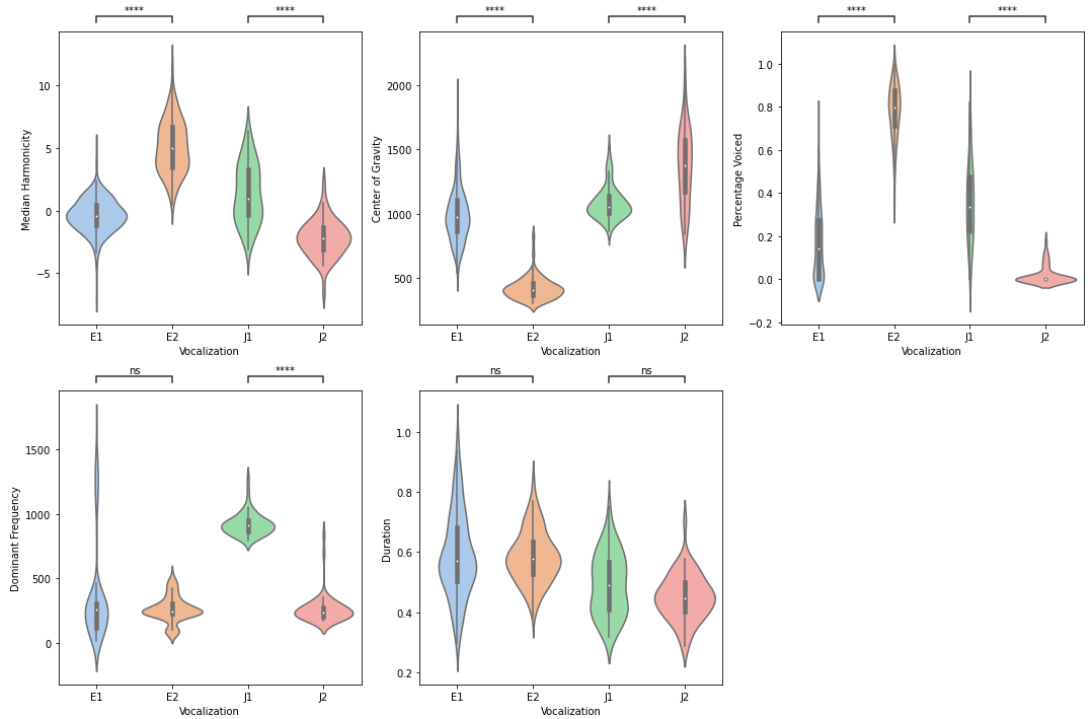

Fig. 7: Visualized spectral parameter distribution of each seal's vocalization types. Significance levels are indicated by asterisks, where ****: $p <= 1.00^{e-04}$, and ns: $5.00e-02 < p <= 1.00^{e+00}$.

A Mann-Whitney-U test revealed that seal E's two vocalization types differed significantly in three of the five extracted spectral parameters, namely center of gravity (statistic=52608.0, p<0.0001), percentage voiced (statistic=171.5, p<0.0001), and median harmonicity (statistic=457.0, p<0.0001). Seal J's two vocalization types differed significantly in four of the five extracted spectral parameters, namely center of gravity (statistic=939.0, p<0.0001), percentage voiced (statistic=4349.5, p<0.0001), median harmonicity (statistic=3825.0, p<0.0001), and dominant frequency (statistic=4470.0, p<0.0001). These significant differences were not affected when correcting for multiple comparisons (five per seal, corresponding to the number of parameters tested).

Discussion

This study provides a protocol to train experimentally naive seals to a vocal learning experiment. In particular, we show how two harbor seals can be trained to associate different S_Ds with different vocal behaviors, which indicates their ability of vocal usage learning.

The task in experiment 1, *Vocalize - Silence*, was learned fairly fast, i.e., within 118 and 220 trials (seal E and J respectively). The task in experiment 2, *Produce Different Vocalizations Upon Different S_Ds*, was learned within 480 trials and 380 trials (seal E and J, respectively). A qualitative comparison between the average trial number of each experiment (*Vocalize and Silence*, x□ = 169 trials; *Produce Different Vocalizations Upon Different S_Ds*, x□ = 430 trials) reveals that the seals needed >2.5 times more trials to learn to *Produce Different Vocalizations Upon Different S_Ds*. Perhaps, experiment 2 was more cognitively demanding than experiment 1; this difference would support the hypothesis that vocal usage learning features components of varying complexity (Janik & Slater, 2000; Shapiro et al., 2004; Vergara & Barrett-Lennard, 2017), which reflects in the time required to learn tasks.

During experiment 1, the seals were asked to vocalize or remain silent on S_D. Both seals succeeded at this task (Fig. 4). Investigating the error to criterion (ETC) value separated per S_D revealed that seal J responded almost equally correct (or incorrect) to either of the SDs. Seal E, however, showed a preference for vocalizing over silence (Fig. 4). In experiment 2 the seals had to produce different vocalizations upon presentation of different S_Ds. Here, especially seal E showed a much higher ETC value for the E2 vocalization, than for the E1 vocalization (Fig. 6). A reason for this may be the active shaping of the E1 vocalization. This shaping directly preceded experiment 2, causing a temporarily higher exposition of vocalization E1 over E2, which could have led to a type preference. However, seal J showed

similarly higher ETC values for one vocal type (J1) over the other (J2). It is not rare that animals show side or S_D preferences (Erdsack et al., 2022; Ortiz et al., 2019; Schluessel et al., 2014).

In classical cognition tasks, considering the Clever Hans effect is important to avoid falsified results. The Clever Hans effect occurs, when an experimenter involuntary cues an animal. It leads to the animal learning a task based on other cues than intended (Sebeok & Rosenthal, 1981). Vocal usage learning, however, is independent of this phenomenon, as it is per definition learning to vocalize in association with a certain context (Janik & Slater, 2000; Janik & Slater, 1997). Whether or not the tested animals picked up on the intended S_Ds or unintended cues is irrelevant. Vocal usage learning would be displayed in any case, solely on other cues than anticipated (Shapiro et al., 2004).

To study the learning process will require more fine-grained documentation of the training sessions from the moment of capturing a novel behavior, including trial numbers. This was unfortunately not possible in this study during Training. Further, a consistent number of trials per session enables the comparison of learning processes. Following such guidelines will further facilitate the direct comparison of vocal learning abilities in different individuals or species (Lattenkamp et al., 2021). Future studies should address this and track stimulus exposition during training to better understand underlying (vocal) learning mechanisms.

Overall, our study shows that harbor seals, just like the closely related grey seal (Shapiro et al., 2004; Stansbury et al., 2015), can be trained to participate in vocal learning experiments. Through successive approximations they can be trained to remain rested in an external enclosure and there tested for the behavior of interest. Experiment 1 shows fast learning of differentiating vocalizing and remaining silent upon demonstration of different visual S_Ds. Experiment 2 shows that the seals are further capable of distinguishing two different vocalization types in response to different S_Ds. Future studies should investigate whether harbor seals are also capable of vocal comprehension learning, i.e., understanding vocalizations in specific contexts.

Conclusion

This study aimed at preparing two experimentally naive harbor seals for vocal learning experiments. Our report complements a previous one on vocal shaping and usage learning in a male harbor seal (Schusterman, 2008). Here, we show that harbor seals are good candidates for research as they can be readily trained for participating scientific experiments. Further, this report confirms that male harbor seals are capable of vocal usage learning, and adds that female harbor seals are also capable of it. Our results further suggest that vocal learning may imply different levels of complexity, which to the best of our knowledge, has not

been shown before in this species. Future studies should incorporate a tracking of the learning progress for a quantitative documentation of vocal usage learning, i.e., during training all presented S_Ds should be noted. It would be further interesting to build on Shapiro et al. (2004) and Stansbury et al. (2015) experimental design on higher levels of usage learning, to see whether harbor seals are capable of such 'more complex' vocal usage learning.

Author Contributions

Conceptualization: DD, AR; Data Curation: DD; Formal Analysis: DD; Funding Acquisition: AR; Investigation: DD; Methodology: DD, AR; Project Administration: DD; Resources: DD, AR; Software: DD; Supervision: AR; Validation: DD; Visualization: DD; Writing – Original Draft Preparation: DD; Writing – Review & Editing: DD, AR

Funding

Center for Music in the Brain was funded by the Danish National Research Foundation (DNRF117). The Comparative Bioacoustics Group was funded by Max Planck Group Leader funding to AR.

Ethical Statement

Ethical review and approval were waived for this study, due to no permit requirement of the general training procedures that were conducted (Az. 81-04.78, Landesamt für Natur, Umwelt und Verbraucherschutz NRW).

Informed Consent Statement

Not applicable

Data Availability Statement

The data can be obtained from the corresponding author upon request.

Acknowledgments

We thank Dr. Sara Torres Ortiz and Prof. Dr. Vera Schluessel for providing helpful comments on the manuscript. We thank Yannick Jadoul for helpful thoughts along the way, the discussions on stimulus randomization, and the brainstorming on plotting. Center for Music in the Brain was funded by the Danish National Research Foundation (DNRF117). The Comparative Bioacoustics Group was funded by Max Planck Group Leader funding to A.R.

Conflicts of Interest

The authors declare no conflict of interest.

References

Adret, P. (1993). Vocal learning induced with operant techniques: an overview. *Netherlands Journal of Zoology*, *43*, 125-125.

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* In *Glot International* (Version 6.2.12) retrieved 17 April 2022 from http://www.praat.org/

Duengen, D., Fitch, W. T., & Ravignani, A. (2023). Hoover the talking seal. *Current Biology*, *33*(2), R50-R52. https://doi.org/10.1016/j.cub.2022.12.023

Duengen, D., Jadoul, Y., Ravignani, A. (under review). Vocal usage learning and vocal comprehension learning in harbor seals.

Erdsack, N., Dehnhardt, G., & Hanke, F. D. (2022). Serial visual reversal learning in harbor seals (Phoca vitulina). *Animal cognition*, 1-11.

Gellermann, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *The journal of genetic psychology*, *42*, 206-208.

Goncharova, M., Jadoul, Y., Reichmuth, C., Fitch, W. T., & Ravignani, A. (2024). Vocal tract dynamics shape the formant structure of conditioned vocalizations in a harbor seal. *Annals of the New York Academy of Sciences*.

Jadoul, Y., De Boer, B., & Ravignani, A. (2023). Parselmouth for bioacoustics: automated acoustic analysis in Python. *Bioacoustics*, 1-17.

Jadoul, Y., Duengen, D., & Ravignani, A. (2023). PyGellermann: a Python tool to generate pseudorandom series for human and non-human animal behavioural experiments. *BMC research notes*, *16*(1), 135.

Jadoul, Y., Thompson, B., & De Boer, B. (2018). Introducing parselmouth: A python interface to praat. *Journal of Phonetics*, *71*, 1-15. https://doi.org/10.1016/j.wocn.2018.07.001

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1-11. https://doi.org/10.1006/anbe.2000.1410

Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. In P. J. B. Slater, J. S. Rosenblatt, C. T. Snowdon, & M. Milinski (Eds.), *Advances in the Study of Behavior, Vol 26* (Vol. 26, pp. 59-99). Elsevier Academic Press Inc. https://doi.org/10.1016/s0065-3454(08)60377-0

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, *366*(6461), 50-54.

Lattenkamp, E. Z., Hörpel, S. G., Mengede, J., & Firzlaff, U. (2021). A researcher's guide to the comparative assessment of vocal production learning. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200237.

Lattenkamp, E. Z., Vernes, S. C., & Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *J Exp Biol*, *221*(Pt 14). https://doi.org/10.1242/jeb.180729

Moore, B. R. (1996). The evolution of imitative learning. In C. Heyes & B. Galef Jr (Eds.), *Social learning in animals: The roots of culture*. Academic Press.

Ortiz, S. T., Maxwell, A., Hernandez, A., & Hansen, K. A. (2020). *Does participation in acoustic experiments improve welfare in captive animals? A case study of three grey seals (Halichoerus grypus)*. Cold Spring Harbor Laboratory. https://dx.doi.org/10.1101/2020.08.15.252460

Ortiz, S. T., Maxwell, A., Krasheninnikova, A., Wahlberg, M., & Larsen, O. N. (2019). Problem solving capabilities of peach-fronted conures (Eupsittula aurea) studied with the string-pulling test. *Behaviour*, *156*(5-8), 815-846.

Raimondi, T., Haas, C.E., et al. (under review). Less variable auditory input increases vocal individuality in harbour seal pups (*Phoca vitulina*).

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, *63*(5), 1050-1056. https://doi.org/10.1139/z85-157

Schluessel, V., Kraniotakes, H., & Bleckmann, H. (2014). Visual discrimination of rotated 3D objects in Malawi cichlids (Pseudotropheus sp.): a first indication for form constancy in fishes. *Animal cognition*, *17*, 359-371.

Schusterman, R. J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. *The evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication*, 41-70.

Sebeok, T. A., & Rosenthal, R. E. (1981). The Clever Hans phenomenon: Communication with horses, whales, apes, and people. *Annals of the New York Academy of Sciences*.

Seyfarth, R. M., & Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain and Language*, *115*(1), 92-100.

Shapiro, A. D., Slater, P. J., & Janik, V. M. (2004). Call usage learning in gray seals (*Halichoerus grypus*). *Journal of Comparative Psychology*, *118*(4), 447-454. https://doi.org/10.1037/0735-7036.118.4.447

Stansbury, A. L., de Freitas, M., Wu, G.-M., & Janik, V. M. (2015). Can a gray seal (*Halichoerus grypus*) generalize call classes? *Journal of Comparative Psychology*, *129*(4), 412. https://doi.org/10.1037/a0039756

Stoeger, A. S., & Baotic, A. (2021). Operant control and call usage learning in African elephants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1836), 20200254. https://doi.org/10.1098/rstb.2020.0254

Van Parijs, S. M., Hastie, G. D., & Thompson, P. M. (2000). Individual and geographical variation in display behaviour of male harbour seals in Scotland. *Animal Behaviour*, *59*, 559-568. https://doi.org/10.1006/anbe.1999.1307

Vergara, V., & Barrett-Lennard, L. (2017). Call Usage Learning by a Beluga (Delphinapterus leucas) in a Categorical Matching Task. *International Journal of Comparative Psychology*, *30*. https://doi.org/10.46867/ijcp.2017.30.01.08

Vernes, S. C., Kriengwatana, B. P., Beeck, V. C., Fischer, J., Tyack, P. L., Ten Cate, C., & Janik, V. M. (2021). The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200236. https://doi.org/10.1098/rstb.2020.0236