

3

Vocal Learning in Mammals with Special Emphasis on Pinnipeds

Ronald J. Schusterman

Introduction

At the most general level, questions about the relative rigidity or plasticity of animal communication systems are relevant to a wide variety of subject areas, including behavioral ecology, sociobiology, comparative cognition, and evolutionary and developmental linguistics. As a result, scientists from diverse fields have attempted to identify the form and function of communicative signals produced by a variety of different species, particularly those occurring in the acoustic domain. In the late 1950s, the prevailing view among prominent classical ethologists and behavioral psychologists was that the acoustic signals emitted by animals were ritualized and not modifiable as they are in humans. Behavioral scientists such as B. F. Skinner (1957) typically believed that behavioral changes (as expressed by various nonvocal responses like bar pressing in rats) could occur as a result of environmental experiences but that the vocal responses of nonhuman animals were far less flexible than those of humans. Instead, emotional/motivational constraints, anatomical limitations, and genetic predispositions were believed to control animal vocalizations while environmental or behavioral consequences controlled human spoken language. Despite scattered accounts and observations that some songbirds and parrots appeared capable of mimicking sounds, behavioral scientists of the time generally remained convinced that animal vocalizations were involuntary responses that were exclusively used to express emotions signaling the probability of attack, flight, or courtship. As a result, the potential influence of learning on animal vocal responses was largely discounted and poorly investigated.

This traditional viewpoint was later updated for songbirds, as a result of experimental evidence that revealed a propensity for vocal learning that can be characterized as a sensorimotor skill (see Wilbrecht and Nottebohm, 2003). This skill is acquired when a young bird, during a sensitive period of ontogeny, perceives sound produced by adult conspecifics, remembers what it has heard, and at some later stage of development vocally imitates the adult model. This definition of vocal learning as

the modification of vocal output through reference to extrinsic auditory information has at its core an ethological focus and leans heavily on such constructs as "learning instinct," "critical period," and "sensory gating mechanisms."

This general model of vocal learning in songbirds was derived in large part from a series of field and laboratory observations on the song dialects of white-crowned sparrows carried out by Peter Marler and his colleagues in the San Francisco Bay area of California (e.g., Marler and Tamura, 1962). In the field, these sparrows were observed to exhibit small-scale regional differences in song structure that could not easily be explained by genetic differences between overlapping populations. In order to further investigate whether experiential factors were contributing to these differences in vocal production, Marler's team moved to the laboratory and applied research methodologies that depended on passively tutoring young birds by exposing them to tape recordings of conspecific and extraspecific song stimuli. In doing so, they experimentally demonstrated that while vocal learning does occur in songbirds following passive exposure to auditory information, the process is instinctive, restricted to a sensitive period, and constrained by a species-specific song template. Marler's pioneering studies of vocal learning in songbirds relied on tape-recorded songs and thus eschewed live tutors capable of interacting socially with young birds. A more behaviorally interactive approach developed by Baptista and Petrinovich (1984) later showed that vocal learning may occur beyond the recognized sensitive period and can include anomalous or allospecific vocalizations when avian social tutors replace recorded songs.

The degree to which the vocalizations of some birds may be modified by learning occurring through social feedback continues to be clarified. For example, West and King (1985) found that female cowbirds perform a subtle wing stroke display during a male's serenade in response to acoustic structures in the male's performance. Mature males respond to the wing stroke by altering their songs to include more of those vocal patterns that are more effective at eliciting the female's copulatory posture. In summary, females react more positively to some song variants than others, and such social reinforcement modifies male vocal behavior. Further, other investigators have shown that some young male songbirds with a history of socially interactive experiences can acquire a repertoire of song syllables by means of operant conditioning (Adret, 1993a) or by a process that Marler (1990) has termed "action-based learning." One example of such learning comes from Adret (1993b), who concluded from a study of young zebra finches that operant conditioning, with conspecific songs serving as reinforcers, increases the attention of listeners toward the detail of song phrases, thereby increasing the motivation of the bird to imitate the tape-recorded song. Pepperberg (1986) has also shown that social interactions between a human tutor and an African gray parrot greatly facilitate the imitation of human speech by the parrot.

Studies such as those just described, which show that the call learning by some adult passerines and psittacines are influenced by social experiences, have been expanded to show that operant conditioning of vocalizations can be accomplished with not only social reinforcers but food rewards as well. Innovative operant procedures devised by Manabe and his colleagues (Manabe, Kawahima, and Staddon, 1995; Manabe, Staddon, and Cleaveland, 1997) have used computer-based systems and food reinforcement to train budgerigars to alter frequency characteristics of their calls and to emit these calls in novel contexts.

Collectively, the evidence reviewed here suggests that the production or acoustic structure and contextual usage of natural vocalizations emitted by mature birds of certain species can be modified by the selective distribution of social and food reinforcement. It is likely that the behavioral scientists of just a generation ago would consider such findings to be remarkable. While human utterances have traditionally been viewed and continue to be viewed as subject to powerful learning effects, there remains little evidence for evolutionary continuity in vocal learning processes between human and nonhuman animal species. The introduction of contingency learning—the modification of behavior through selective reinforcement—as a field of special interest within both human and nonhuman animal vocal learning may provide significant clues to the common bases of communication systems. For example, recent research on human speech acquisition indicates that contingency learning and social shaping may serve as general processes that underlie the ontogeny of speech in humans as well as song development in birds (Goldstein, King, and West, 2003). This research has shown that the babbling of eight-month-old infants sounds progressively more speech-like when mothers socially reinforce recognizable syllables emerging during this maturational period than do comparable sounds that are produced by control subjects who are not given such positive feedback. This work has striking parallels to the research previously described on social influences in song learning by male cowbirds.

Thus, the extent to which the vocalizations of nonhuman mammals may be modified through experience is especially interesting from a comparative perspective. While more is currently known about how some mammals learn to respond to different auditory cues, little is understood about if and how readily they learn to adapt their own vocal behavior in response to changing auditory and social environments. In this chapter, I focus on contingency learning as a tool for understanding how the vocalizations that are emitted by mammals may be influenced by reinforcement contingencies. If we are to understand the evolution of spoken language or human speech, then it is necessary to study the vocal behavior of those species that possess adequate vocal complexity. Such an animal group must have “a capacity for vocal learning and a vocal tract with a wide phonetic range” (Fitch, 2000). Moreover, such candidates for studying the evolution of spoken language must vocalize frequently.

This is necessary because vocal behavior is more likely to develop depending on the degree to which it competes with other responses (e.g., limb movement) for access to positive reinforcement (Salzinger, 1973).

The Search for Vocal Learning in Nonhuman Mammals

Support for Classical Perspectives on Flexibility in Vocalizations

In contrast to the research on certain bird species that has identified vocal learning as a significant influence in avian acoustic communication, for the most part, the traditional viewpoint about the relative inflexibility of animal vocal responses has been substantiated to the present day by studies of mammalian vocal production. Results from a good deal of research, especially on nonhuman primates, suggest that the acoustic structure of species-typical mammalian vocalizations develops within a relatively closed genetic program. For example, deafening at birth produces only minor modifications in vocal development in some monkeys (Hammerschmidt, Freudenstein, and Juergens, 2001). Further, nonhuman primates raised either in isolation or with foster mothers of a different species still emit vocalizations that are species typical (Owren, Dieter, Seyfarth, and Cheney, 1993). In a recent review of vocal development in nonhuman primates, Seyfarth and Cheney (1997) point out that while there is abundant evidence of flexibility and modification in the response that individuals perform upon hearing calls, the evidence for plasticity in the contextual usage of emitted vocalizations is more limited and there is virtually no evidence for learned (nonmaturational) modifications in call structure. These observations suggest that it is unlikely that nonhuman primates have voluntary control of their vocal production apparatus, which is more likely under the direct control of affectively linked systems.

It is also significant that experimenters attempting to place primate vocalizations under stimulus control in laboratory situations have typically had limited or no success. While failed attempts have remained generally unpublished, the difficulties inherent in such conditioning procedures are well-known among primatologists (M. Owren, personal communication; R. J. Schusterman, personal observation). These include the observations that only certain call types, combined with specific, appropriate reinforcement types, may be useful in conditioning procedures at all (see Pierce, 1985).

These observations with nonhuman primates are consistent with the view that the vocal signals of mammalian species are automatic and a fixed part of instinctive behavior. This view has been partially explained by theories suggesting that mammalian vocal musculature is constrained by direct emotional controls and is therefore unavailable for modification by means of operant conditioning techniques (Salzinger

and Waller, 1962). If true, this would predict that mammals should lack voluntary control over the usage of their vocal behavior and therefore be unable to learn when and under what conditions and contexts to use a call. According to this concept, a call type that is associated with a particular eliciting stimulus should not easily be mapped onto novel arbitrary stimuli associated with a different or neutral affective context, and therefore, the call type cannot come to acquire novel functions. Based on this rationale, mammals other than humans have generally been considered incapable of learning to use their vocalizations independent of emotional contexts. Such contextual flexibility in vocal responses has recently been termed “vocal usage learning” by Janik and Slater (2000). Moreover, mammals were also believed to be incapable of what Janik and Slater have termed “vocal production learning,” or the ability to arbitrarily alter the physical structure of their sounds or to imitate sounds from social or nonsocial sources.

A final idea with respect to the fixed nature of vocal behavior in mammals can be derived from well-established principles regarding biological constraints on learning. These principles state that there are limits on the extent to which some behaviors can be modified by contingencies that depend on reinforcers that have no natural relationship to the response being emitted. For example, while it is quite easy to train a rat to press a lever to receive food or water reinforcement, it is difficult, if not impossible, to teach the same behavior using shock avoidance. Conversely, the same rat can be easily trained to jump or run using shock avoidance techniques, while these escape behaviors are not readily learned using food reinforcement (Bolles, 1970). One might expect that, in situations where the natural vocalizations of mature animals have no normal functional or causal relationship with food, the modification of vocal behavior using food rewards might be at best difficult or at worst not possible. On the other hand, perhaps vocal conditioning using food rewards can more readily occur in species that emit acoustic signals to detect and locate food (e.g., the use of echolocation signals by dolphins and bats). Of course, any of these possibilities involving vocal conditioning would depend on call production's being under volitional control to some degree.

For all of the above reasons, including the apparent fixedness of mammalian sound production, probable emotional constraints, presumed anatomical inflexibility, and possible biological limitations on vocal learning, it would seem unlikely that comparative studies of mammalian vocalizations would play an important role in the search for precursors of complex communication systems such as human speech. However, there is now accumulating evidence to suggest that signals used in vocal communication by some mammals are indeed accessible to discriminative control and selective shaping by the process of contingency learning using positive reinforcement (see box 3.1).

Box 3.1

Instrumental and operant conditioning in relation to the vocal behavior of mammals

Instrumental conditioning (as first described by Edward Thorndike) and operant conditioning (as articulated by B. F. Skinner) are usually treated as functionally equivalent concepts describing general principles of learning. However, Thorndike, unlike Skinner, emphasized the importance of species-typical or species-specific responses elicited by identifiable stimuli and strengthened by appropriate reinforcement (Thorndike, 1911). Therefore, Thorndike's early ideas about learning actually fit well with our current understanding of biological constraints on learning. Skinner, on the other hand, emphasized that operants were never elicited responses but rather novel or arbitrary responses (Skinner, 1938). Skinner considered these behavioral units to be *emitted* rather than *elicited* responses because these operants have no identifiable eliciting stimulus. Reinforcement increases the future probability of operant responding when its delivery is made contingent on the emitted behavior. Moreover, an arbitrary stimulus can become discriminative for the behavior because it evokes but does not elicit the operant response that it precedes and accompanies.

With respect to vocal learning, it is likely that the general principles shared by both instrumental and operant conditioning paradigms can lead to discriminative control of vocal behavior. Once such control has been established, modification of sound production through selective reinforcement is possible, and thus, vocal learning through the process of contingency learning can occur. The distinction between instrumental and operant conditioning in this regard becomes relevant when we consider both the nature of the response and the reinforcer. The finding that vocal behavior in some species can come to occur reliably in the presence of an arbitrary discriminative cue that is established by using food reinforcement suggests vocal conditioning may more closely follow the instrumental paradigm because the anticipation of food is clearly important. Skinner (1957) himself did not believe that mammalian vocalizations were voluntary, and therefore were not subject to operant control. This is clearly not the case, as the last four decades of research on conditioning of "involuntary" autonomic responses has shown. It is quite possible that Skinnerian conditioning procedures can be used to acquire discriminative control of vocal responses, and it is likely that such procedures, when applied experimentally to the study of sound production in some mammals, will reveal a degree of plasticity and environmental influence heretofore not described. However, there is a strong caveat here because it is well-known that vocalizations like food coos of macaque monkeys as well as baboon grunts and the food calls of chimps are critically dependent on the underlying affective/motivational systems that elicit calling (Owren, personal communication).

Finally, because contingency learning essentially boils down to reinforcement training, it is prudent to carefully evaluate the properties of the reinforcer involved. While the simplest examples involve food (or water) reinforcement, other biological reinforcers, such as access to sexual partners, shelter, or safety can also be substituted into the reinforcement contingency. Further, as classical conditioning experiments have demonstrated, any other cues associated with such biological reinforcement can come to share reinforcement value as well. Thus, arbitrary cues as well as social stimuli can come to enter into situations where behavior is modified through positive reinforcement. As a consequence, it is possible that social reinforcement may be effective at strengthening and influencing sound production in social settings. This can be illustrated by the example of the cowbirds discussed earlier, where a vocal response (song production by a male) can be modified by a discriminative cue (the wing stroke of a female) that has come to be associated with biological reinforcement (copulatory posture).

Evidence in Support of Flexible Vocal Responses

The evidence suggesting inherent inflexibility in the vocal responses of mammals outlined in the previous section is somewhat paradoxical in light of the great malleability in both the structure and function of human utterances, which incidentally are readily accessible to a variety of operant conditioning situations including vocal mimicry. The notion of a total evolutionary disconnect between the vocal communication of humans and other animals was summarized most clearly and succinctly by Lenneberg (1967), who rejected the study of animal vocal communication because he claimed that the processes of human language were “deeply rooted, species-specific, innate properties of man’s biological nature” (p. 394) and because “no living animal represents a direct primitive ancestor of our own kind and, therefore, there is no reason to believe that any one of their traits is a primitive form of any one of our traits” (p. 234–235). Today, however, studies of animal vocal learning and communication are considered relevant to larger scientific investigations of human language evolution. This approach is typified by Hauser, Chomsky, and Fitch (2002) and Fitch, Hauser, and Chomsky (2005), who have both recently pointed out that any faculty involved in language (e.g., vocal learning) is worthy of comparative study because it is part of a language faculty in a broad sense. These faculties may be contrasted, they say, with those mechanisms that are both specific to language and uniquely human, such as recursion, which is involved in the faculty of language in the narrower sense.

Consistent with this approach, Janik and Slater (1997) have recently pointed out that investigations within a comparative context “... enable us to compare and contrast birds and mammals and to consider the possible functional significance of vocal learning” (p. 60). In addition to this review, there have been several other reviews of vocal learning in the past thirty years (Salzinger, 1973; Adret, 1993a; Janik and Slater, 2000). These show that the phonations of such ecologically and phylogenetically diverse species as sea lions, dolphins, beluga whales, monkeys, cats, dogs, and rodents, can—at least in part and under the right circumstances—be acquired by operant conditioning techniques using food reinforcement. Furthermore, while several studies demonstrate that animals can learn to produce vocalizations in response to specific arbitrary cues, complementary studies have found that mammals can also learn to inhibit their vocalizations in the presence of different arbitrary cues (Schusterman and Feinstein, 1965; Schusterman, 1967). These findings, which demonstrate volitional control of sound production as well as inhibition of sound production (both of which occur at very different levels of internal control), show that the vocal response involved is not merely a function of an animal’s state of arousal. They also provide evidence for call usage learning in the species whose vocalizations have been brought under stimulus control, as vocalizations controlled by particular emotional contexts come to be controlled by qualitatively different cues. Some laboratory

experiments similar to those just described also touch on call production learning, where individuals learn to modify certain structural properties of their vocal emissions through experimenter manipulation of reinforcement contingencies. These properties include call rate, pitch, loudness, and duration (e.g., Adret, 1993a). While the acoustic structure of operantly conditioned calls can sometimes be purposefully shifted along these dimensions through selective reinforcement, Janik and Slater (1997) point out that these are all dimensions that are also intimately connected with arousal and affect and that modifications along these dimensions can be accomplished by mechanisms other than direct control of vocal musculature per se (e.g., through respiratory action). This again raises the issue of the influence of emotional constraints on flexibility in vocal responses. There is weaker experimental evidence for modifications of more arbitrary aspects of call structure, such as fundamental frequency and frequency modulation, and it remains to be seen to what extent these call dimensions are broadly available for modification. In addition to experimental data demonstrating that some mammals are able to control and modify their sound production, there is evidence from natural settings showing that some marine mammals, including whales, dolphins, and seals, and perhaps some terrestrial mammals, such as bats, can alter the acoustic structure of their sound emissions on the basis of auditory feedback from other individuals. This observational evidence, which is probably the strongest evidence of vocal learning in nonhuman mammals, is reviewed in detail elsewhere (Tyack and Sayigh, 1997; Janik and Slater, 1997).

While the information available on vocal learning in mammals is relatively sparse and limited to a handful of phylogenetically disparate species, the data that are available support the view that the study of vocal learning in mammals is relevant to the search for the evolutionary and ontogenetic origins of complex communication systems including human speech. Evidence that the vocalizations of several mammals, like humans and some birds, have some degree of emotional (fixed) control in addition to some degree of voluntary (flexible) control is significant. When taken alongside observations suggesting that the calls of these mammals may be modified by learning with respect to both structure and function, these findings suggest that contingency learning may be a general process that supports more complex vocal learning abilities.

More field and laboratory investigations are required to better understand how reinforcement contingencies can modify vocal responses. An example of a relevant field study comes from Seyfarth and Cheney (1986), who have shown that when an infant vervet monkey gives a contextually correct alarm call, it is frequently given positive feedback in the form of a second appropriate alarm call by an adult. These adult calls are consistent with the infant's alarm call and might serve as reinforcers that provide essential feedback on vocal responses. Positive feedback of this nature should facilitate the infant monkey's developing an understanding of an equivalence

relation between different alarm calls and their referents (see Schusterman, Reichmuth Kastak, and Kastak, 2003). However, it is essential that experiments exploring such learning under highly controlled conditions, such as those conducted with human infants (Goldstein et al., 2003) and cowbirds (West and King, 1996), be explored with nonhuman mammals. Such experiments may help to disentangle the ecological, phylogenetic, and ontogenetic variables affecting vocal learning. Additionally, it is clear that more research is needed in the area of vocal conditioning using food reinforcement. Some recent examples of this approach comes from Adret (1993a) and Shapiro, Slater, and Janik (2004), who have outlined progressively more complicated examples of operant conditioning in the context of experimental studies. These include an initial experimental phase in which an animal's specific natural call types come to be controlled by different arbitrary stimuli. In the next phase, the animal has to use its generalization and cognitive skills to vocally respond to classes of stimuli. In the final phase, the subject has to match vocally the stimulus that it has just heard (i.e., barking when it hears a bark and growling when it hears a growl). Experiments such as these may be useful for a wide variety of reasons. For example, they would make possible anatomical investigations of vocal musculature in trained animals which may help to identify structural constraints on vocalizations. Further, because such operant conditioning research can be conducted with precision, such experiments may enable us to set procedural and measurement standards by which we can compare the vocal learning abilities of different mammalian species.

In the following sections, I introduce a group of mammals that may be especially well-suited for controlled studies of flexibility in sound production using food reinforcement—the pinnipeds. Following a brief introduction to these animals, I describe published experiments and new observations on their vocal usage and production learning capabilities. In contrast to several contemporary reports on primate vocal communication (see, e.g., Seyfarth and Cheney, 1997) this material emphasizes plasticity in the production of vocalizations rather than in the comprehension of acoustic signals. The research discussed here represents some of the most comprehensive and successful efforts to operantly condition vocal responses in nonhuman animals. The data will show that some pinnipeds can learn to alter the contexts in which they emit natural vocalizations, as well as modify their call structure along several structural dimensions, to an extent not known to occur in most terrestrial mammals.

The Case for Vocal Learning in Pinnipeds

A Brief Introduction to the Pinnipeds

The pinnipeds are semi-aquatic, carnivorous mammals consisting of 33 living species in three major families: phocids (true seals), otariids (sea lions and fur seals), and

odobenids (walruses). Their ancestors were terrestrial carnivores that originated in temperate waters of the northern hemisphere about 30 million years ago and spread throughout the world's oceans, colonizing isolated shores and feeding in nearby productive waters. Pinnipeds combine two characteristics which occur together in no other mammals. They give birth to and attend their pups on land or ice and they feed exclusively at sea (Bartholomew, 1970). These amphibious mammals are especially interesting mammals to study regarding vocal learning for numerous reasons, some of which are briefly itemized here, and several of which are further explained in the following sections: (1) All pinnipeds spend a significant amount of time at sea as well as ashore. Because their visual range is limited in both media, most species rely heavily on acoustic modes of communication. (2) Like humans, pinnipeds not only produce many of their vocal signals in the larynx, but they have relatively good control over articulatory movements occurring above the larynx. Thus, the sounds produced by the larynx can be influenced by oral and nasal cavities as well as by movement of the tongue, lips, and mouth. Certain species also have highly specialized structures used in sound production. (3) Directly related to the above, some pinnipeds have specialized oral feeding musculature which makes them good candidates for motor conditioning of movements of the lips, mouth, tongue, teeth, and throat muscles. (4) Pinnipeds are diving mammals that have superior breath control relative to terrestrial mammals. Such breath control may be used during sound production both in air and under water but the mechanism for doing this is not well understood. (5) Since all pinniped pups are born on land or ice and utter airborne distress calls with their mouth open, upon entering the water, they must in some manner modify the way they produce their vocal signals. It is interesting to note that many of the sounds emitted by pinnipeds under water do not involve the release of air. (6) Several species that mate aquatically or semi-aquatically emit acoustically complex vocalizations or songs during the breeding season. (7) In at least seven species, there is evidence of intraspecific geographic variation in vocalizations. In some cases, this variation is suggestive of regional dialects similar to those found in humans, birds, and cetaceans (Tyack and Miller, 2002). (8) Vocal mimicry has been clearly demonstrated in a single male harbor seal. This seal, called "Hoover," was reared in New England by human caretakers before being transferred to an aquarium; as he matured, he spontaneously emitted a repertoire perceived as about twelve English words produced through a variety of vowel-like sounds. Anybody who listens to recordings of Hoover cannot fail to be impressed and amused by the seal's New England dialect and flawless belly laugh (Hiss, 1983; Rawls, Fiorelli, and Gish, 1985). Hoover imitated human speech at least as proficiently as any parrot I have ever heard. In an effort to replicate the Hoover finding, another male harbor seal named "Chimo" was trained using operant conditioning to emit human speech-like sounds

but only eventually gave poor approximations of “hello” and “how are you” as well as several vowel sounds (Moore, 1996).

Acoustic Communication in Pinnipeds

Although contemporary pinniped species inhabit unique behavioral and ecological niches, another significant commonality among them is the apparent structural complexity of their phonations and the degree to which they used these vocal signals in communicative contexts (Schusterman and Van Parijs, 2003). Whereas whales and dolphins are known to use many of their sounds for the purpose of detecting prey, pinnipeds do not appear to possess similarly specialized echolocation abilities (Schusterman, Kastak, Levenson, Reichmuth, and Southall, 2000). Indeed, the sounds that are produced by pinnipeds are used almost exclusively for influencing or being influenced by other individuals—in other words, in some form of social communication (see Dawkins, 1995). Their amphibious lifestyle requires that social communication occur both in the atmosphere and the hydrosphere, therefore, most pinnipeds need to communicate acoustically in both media. All pinnipeds exhibit some degree of polygyny in their breeding behavior, and consequently, their vocal behavior is highly influenced by competition for mates among males.

Many pinnipeds, including all of the otariids and some of the terrestrially breeding phocids, congregate seasonally on traditional sites to breed and give birth. These rookeries can be very dense, with up to tens of thousands of individuals inhabiting relatively small stretches of coastline. The vocalizations produced in these settings are often loud and repetitive, and the call types produced include male threat/sexual displays, female threats, mothers calling to their pups, and pups calling for their mothers. Variation in many of these vocalizations tends to be greatest between individuals and more stereotyped within individuals. Playback experiments in the field demonstrate that individuals can be identified by distinctive acoustic cues, and that vocal recognition may occur between females and their dependent pups as well as between males competing for access to females during the reproductive season (see Insley, Phillips, and Charrier, 2003, for a review). In contrast to land breeding pinnipeds, many species of phocid seals interact sexually in the water and are more loosely spaced while ashore. Mature males of these aquatically breeding species tend to have unusual and complex underwater vocal repertoires or songs (for reviews, see Tyack and Miller, 2002; Van Parijs and Schusterman, 2003). Seal pups of most pinniped species rarely swim prior to weaning but, in highly precocial species such as the harbor seal, pups may swim soon after birth and can emit their signature calls in water as well as air (Perry and Renouf, 1988).

In addition to the obvious selective advantage of an acoustic mode of communication imposed by the physical environment, it is likely that crowding and competition

for space have also played significant roles in making the vocal-auditory channel such an important part of pinniped communication systems. Vocal signaling by males functions primarily as a mechanism to disperse other males and may potentially influence females' choice of mates. The calls of male pinnipeds appear to serve as "honest signals," that is, they advertise to others the presence of the male and they are associated with a cost. During the breeding season, dominant and territorial males often produce their vocal signals nearly incessantly, and observations in captivity suggest that the energetic cost associated with this behavior is significant (Schusterman and Gentry, 1970). The high density of animals within breeding colonies appears to favor the production of loud, harsh, redundant calls to make signal detection by others more likely.

Crowding has probably functioned as a selective pressure on parturient females as well. The constant threat of mother-pup separation in a rookery leads to the frequent exchange of calls between females and their pups, and rapid learning of unique call characteristics by females, and, in some species, by pups. Later on mothers and their adolescent offspring probably use these calls to locate one another following extended periods of separation (see Insley et al., 2003). Competition between males and interactions between mothers and their pups comprise the most significant social behaviors in pinniped groups, and it is noteworthy that these social relationships are maintained by vocal signaling that can be easily viewed in the framework of contingency learning: Responses to affiliative calls are associated with positive consequences, and responses to aggressive calls are associated with the avoidance of negative consequences.

The mechanisms involved in pinniped vocalizations are relatively unstudied but very intriguing (see Tyack and Miller, 2002, for a review). The structures that are common to those used by terrestrial mammals include the larynx and the soft tissue of the nasal membranes, lips, mouth, and tongue; however, these structures are modified to various extents in pinnipeds. Some species also appear to have tracheal mechanisms for sound generation that are probably related to adaptations for diving. These include compressible airways comprised of flexible cartilage and membranes in the respiratory tract that produce sounds by vibration when air is passed over and between these and associated structures, all without necessity of inhalation and exhalation. Additionally, some species have highly derived structures that are involved in sound production, for example, the resonating chambers of the elephant seal proboscis, the inflatable nasal hood and septum of the hooded seal, and the unique pharyngeal pouches of the walrus, each of which can be used to generate or modify sounds. Different sound generating mechanisms can be used alone or in combination to produce distinctive call types, and some species are actually capable of emitting two structurally different calls simultaneously.

Modes of sound production may also change as an individual moves from land to water. For example, male California sea lions display vocally by barking along the boundaries of their shoreline breeding territories. At high tide, a male's territory may become partially or wholly submerged; however, the attending male continues to patrol the same boundaries with vocal displays comprised of underwater barks. While these barks sound similar in air and under water, it is clear that the motor patterns involved are different: Males barking in air do so with their mouths open, while males barking underwater do so with their mouths closed, without releasing any air (Schusterman and Balliet, 1969). Obviously, the various mechanisms and motor behaviors involved in sound production by pinnipeds may be quite complex, therefore, it seems likely that individuals must exercise some voluntary control over their sound emissions.

There are good examples from the field demonstrating that California sea lions can voluntarily control the usage of vocal signals and learn the context in which to use a call (see Schusterman et al., 2003). Nonterritorial subadult males, who frequently vocalize on other parts of the rookery, will sometimes "sneak" through a bull's territory, moving low to the ground and inhibiting their vocalizations during their entire period on the territory. If a territorial bull discovers one of these young individuals, he emits a highly directional, loud series of barks aimed directly at him (see Schusterman, 1977, 1978) and the intruder flees without emitting any vocalizations (Schusterman et al., 2003). Indeed, such evidence gleaned from the field is supported by observations made in captivity of male California sea lions that inhibit their barking in the presence of a territorial male but do not inhibit their barking in the absence of a territorial male (Schusterman and Dawson, 1968).

What Is the Experimental Evidence for Vocal Learning in Pinnipeds?

The most recent experimental work to show that individuals of a pinniped species have voluntary motor control over their vocalizations and can learn the context in which to use a call involves grey seals, which are known to naturally emit growls, clicks, and moans in air and underwater (Schusterman, Balliet, and St. John, 1970). In this recent study, Shapiro et al. (2004) used the withholding of anticipated fish rewards (a frustration technique) to initiate constant aerial calling from two young captive grey seals. To obtain discriminative control of each of the seal's vocal responses, the trainer presented a light and then reinforced the first vocal response given by the seal by pressing a clicker (a conditioned reinforcer) that was followed by the presentation of a fish reward. Therefore, the fish reward was contingent upon vocalizing only in the presence of the light, which served as a discriminative stimulus. The investigators report that under these conditions, both seals rapidly learned to control their vocal musculature and emit sounds only when appropriate. Thus, in

the first phase of this study, call usage learning was demonstrated because the young grey seals signaled vocally given one arbitrary discriminative cue and refrained from emitting vocal signals given another.

I have described this initial phase of the experiments by Shapiro et al. (2004) on grey seals in detail because previous research on another pinniped (California sea lions) done forty years earlier contained nearly the same procedural ingredients and resulted in the same outcome (Schusterman and Feinstein, 1965). In a later study, three more sea lions were added to the sample (Schusterman, 1978) and vocalizations from the three were readily elicited by what I then called the "frustration technique" (i.e., by withholding fish rewards while a sea lion worked at an underwater target-pressing task, a response that hitherto had resulted in reinforcement). When a sea lion was shifted to the vocal conditioning task, it was promptly reinforced with fish for vocalizing when the target-pressing response did not "pay off." Next, the vocalization was brought under the control of the size of circular and triangular stimuli. A click burst emitted in the presence of a large or small stimulus was reinforced, and silence in the presence of the opposite sized stimulus was also reinforced. Thus, vocalization or silence in the presence of the appropriate stimuli defined the correct response. For two subjects, discriminative control of sound emissions was complete after one hundred randomized presentations of large and small stimuli. The third subject required nearly 600 trials before reliably learning to control its vocalizations in this context. As expected, for all of the sea lions, most errors made during the early period of learning were "vocal" and only as the experiment continued did "silent" errors occur. In general, these early studies with California sea lions show that call usage learning occurs more rapidly in these pinnipeds than has been shown in monkeys (e.g., Myers, Horel, and Pennypacker, 1965), cats (Molliver, 1963), and dogs (Salzinger and Waller, 1962). Indeed, several investigators have found some monkeys incapable of call usage learning (e.g., Yamaguchi and Myers, 1972). In a relatively recent study on vocal learning in cats, only twenty-two of thirty-two cats were found to be capable of the simplest type of call usage learning; all subjects in this experiment received extensive training before any of them learned a straightforward vocal conditioning task (Farley, Barlow, Netsell, and Chmelka, 1992). These laboratory results comparing terrestrial and marine mammals are consistent with what many marine mammal trainers have reported to me over the years: Namely, vocalizations are as easy to condition in sea lions and dolphins as are gross motor movements of the head, torso, and limbs. In contrast, when it comes to training non-human terrestrial mammals, usually vocalizations are much harder to condition than gross motor movements.

In a follow-up to the sea lion study just described, I was interested in determining whether a sea lion could transfer its knowledge about when to vocalize and when to remain silent to many pairs of visually presented stimulus configurations

that differed simultaneously in both shape and size. A single sea lion was presented with forty-five such problems, each of which repeated for 160 trials. The stimuli from each problem were randomly ordered in successive trials, and then the next problem was presented. For each new problem, differential fish reinforcement was arranged for vocalizing to one stimulus and for remaining silent to the other stimulus. Following 160 trials, a new pairing was presented and differential reinforcement was arranged for vocalizing or remaining silent. The results showed that during the first five problems, correct responses by the subject averaged 59%. However, on the next ten problems, correct responses jumped to 75%, and on the last five problems, correct responses averaged 88% (Schusterman, 1978). Apparently, using conditioned vocalization and silence as indicator responses, the sea lion had solved each new problem more rapidly than the earlier one. In other words, the sea lion had learned the rule: If vocalizing to one stimulus on the earlier trials of a new problem pays off, then on succeeding trials, keep vocalizing in the presence of that stimulus and remain silent in the presence of the other stimulus. It is the behavior reflecting this rule that established the "learning set" or "learning to learn" phenomenon (Harlow, 1949). In later experiments my colleagues and I used the tightly controlled, operantly conditioned vocalizations of our sea lion subjects as indicator responses for determining their aerial and underwater visual acuity (Schusterman and Balliet, 1970) as well as their aerial and underwater auditory sensitivity (Schusterman, Balliet, and Nixon, 1972). The learning set study, the psychophysical studies, and another study on conditioning rate of barking underwater to different sounds (Schusterman, 1978) clearly show that California sea lions use their generalization skills and are capable of forming broad mental concepts to respond vocally to particular classes of stimuli across sensory modalities.

The grey seal study introduced at the start of this section continued after the calls of two subjects were placed under stimulus control using fish reinforcement. In an interesting attempt to investigate vocal production learning using auditory feedback, Shapiro et al. (2004) trained the same seals to emit two different vocalizations (moan and growl) in response to two different auditory stimulus types, which consisted of a small set of tape-recorded moans and growls. While the training of appropriate (or "type" matched) responses proceeded in an encouraging fashion, when the experimenters introduced novel moans and growls to the seals, both subjects biased their response and produced only growls. These findings led the investigators to conclude that earlier, the seals had learned specific associations between specific exemplars of the two classes of calls but were unable, with the training they were given, to generalize novel auditory stimuli belonging to each class by matching their vocal response to the perceived call type. This study, although demonstrating stimulus control over grey seal vocalizations in several different contexts, failed to show that grey seals are capable of one type of production learning, that is, producing sounds that

match the sounds they hear, a capability that in contrast seems to have been demonstrated with "Hoover," the harbor seal (Rawls et al., 1985).

In a discussion of these results, Shapiro et al. (2004) raised what I consider to be an important issue in the topic of vocal learning in pinnipeds. This is the issue of examining the natural vocal development of these animals in order to determine how their vocal repertoire emerges with maturation. To various degrees, such research has already been done in some nonhuman primate species as well as in dolphins (for several reviews, see Snowdon and Hausberger, 1997).

New Experimental Evidence Supporting Vocal Learning in Pinnipeds

Most of my own research on pinniped vocal learning was completed over three decades ago. In those experiments, my colleagues and I worked mostly with individual animals representing a single species, the California sea lion, and much of what was learned was incidental to other research goals. During my career, I was fortunate to learn about vocal behavior in sea lions and some other pinnipeds through my own active experimentation and personal observations. These experiences included opportunities to learn about natural vocal signaling, amphibious sound production, social and environmental factors influencing vocalizations, sound production over annual and developmental scales, and discriminative control of sound emissions using visual and acoustic cues, all under the umbrella of research on perception, cognition, and communication in these animals. Throughout my career, I have remained intrigued and impressed by the changes in vocal behavior that occur as a result of maturational and experiential processes. I often wish that I had made time for more formal studies specifically dealing with vocal learning in the pinnipeds that I worked with; while the captive studies discussed earlier yielded a wealth of useful information, I am certain that focused efforts using similar reinforcement training techniques to systematically modify vocal responses would have been worthwhile.

Recently, I had the opportunity to participate in some new investigations of vocal learning (along with several close colleagues, who are listed in the acknowledgments) in two species of pinniped that are particularly interesting from the standpoint of sound production and vocal learning. The general goal of this effort was to explore vocal learning using food reinforcement with trained animals with known behavioral and vocal histories. These two case studies are ongoing efforts, but the findings thus far are illustrative of many of the concepts described in this chapter and are useful in suggesting what might be gained from applying such conditioning approaches to the study of vocal learning in nonhuman animals.

Case Study 1: Harbor Seal Vocal Conditioning

At first glance, harbor seals would not appear to be particularly interesting subjects for studies of vocal communication. In air, these seals are relatively silent with acous-

tic signaling among adults consisting of short-range growls and grunts to maintain social spacing and access to haul-out spaces. Pups are born in the spring and are highly precocial. The period of maternal dependency is quite short, and pups are weaned and abandoned within a month from parturition, leaving relatively little opportunity for meaningful nonauditory maternal feedback that might influence vocal development. Harbor seal pups produce stereotyped calls that can be emitted in air and underwater, with structural variation sufficient for individual recognition (Perry and Renouf, 1988). Once pups are weaned, however, they cease to produce these signature vocalizations. In contrast to many other pinniped species, harbor seal mothers do not emit stereotyped calls while attending their pups and they do not exchange calls with their pups. The lack of complex sound production by mature harbor seals in air belies the observation of complex underwater vocal signaling among adult males during the breeding season (Hanggi and Schusterman, 1994). These acoustic displays appear to play a significant role in the breeding behavior of this species. The production of underwater acoustic displays by seals is common to all species that mate in the water rather than on land (for a general review, see Van Parijs and Schusterman, 2003).

The subject of the current study was "Sprouts," a 17-year-old male harbor seal (*Phoca vitulina*) housed at Long Marine Laboratory in Santa Cruz, California. Sprouts was born into a captive colony of harbor seals at Sea World, San Diego, where he lived with his mother and other harbor seals for the first nine months of his life. At Long Marine Lab, he lived with other pinnipeds but no conspecifics. Sprouts was relatively silent until the age of seven, when he spontaneously began producing underwater vocal displays (see figure 3.1A) that were similar in general structure to the typical calls of wild harbor seal males. He has continued to produce these stereotyped vocalizations during the spring of each year, during a time period that coincides with the harbor seal breeding season along the central California coast. These calls appear to have changed little over time.

At the age of five, Sprouts was conditioned to produce an airborne growl-like sound by selectively shaping a low-level guttural sound occasionally made while he was having his teeth brushed. This growl was placed under the control of the discriminative signal (S_D) "speak," and it has been part of his training repertoire since that time. Sprouts rarely if ever produced that or any other airborne sound at any time other than when presented with the discriminative cue, and it seemed to have no relationship to the spontaneous underwater breeding vocalizations that appeared at the onset of sexual maturity.

In the spring of 2005, experimental sessions were conducted with Sprouts to explore how his conditioned airborne vocal response might be modified using fish reinforcement. Prior to this study, little selective shaping of his conditioned vocal response had been conducted and his sound production in response to the S_D

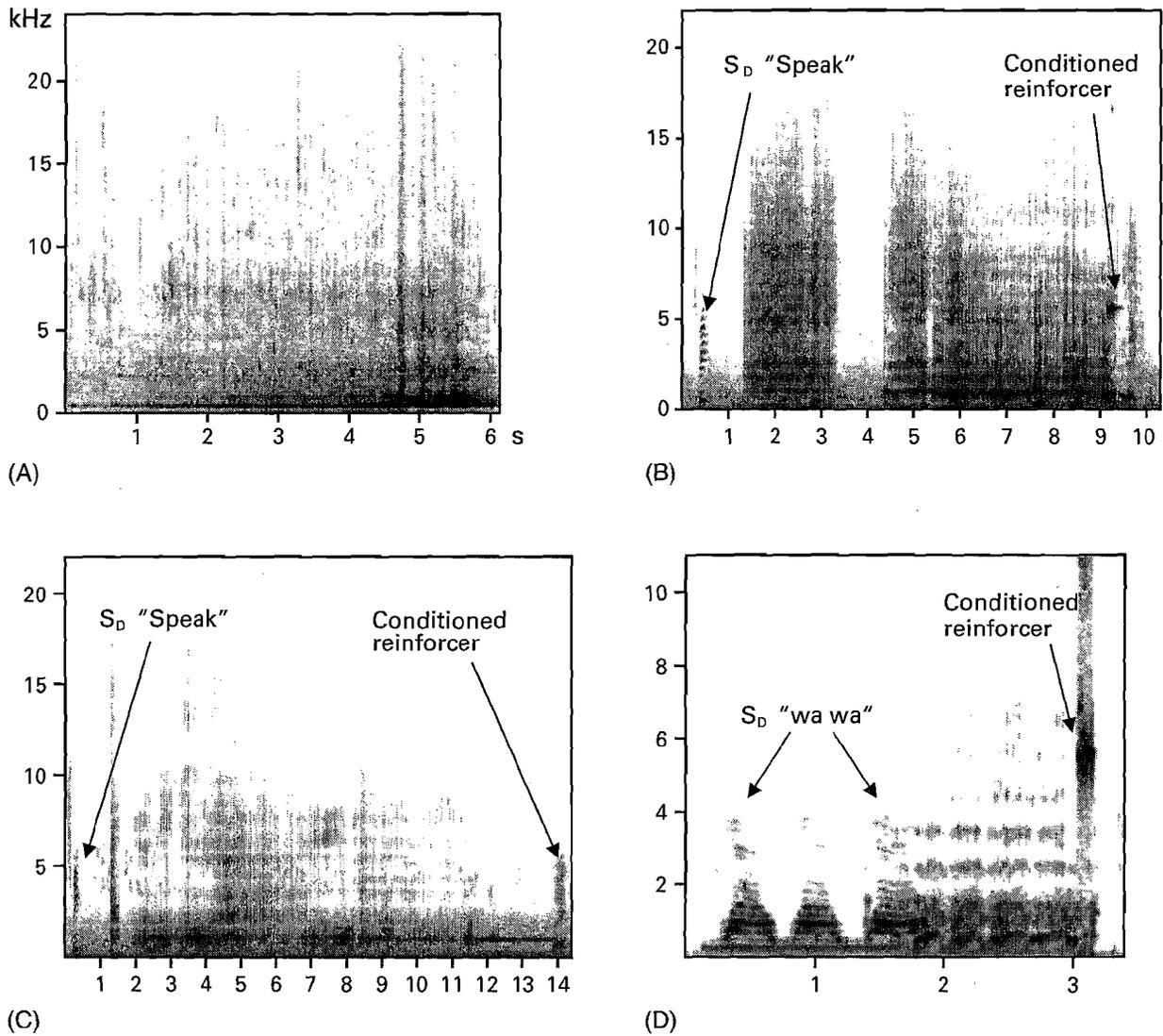


Figure 3.1

Spectrograms of sounds produced by a 17-year-old male harbor seal named Sprouts. Spectrograms show frequency in kilohertz on the vertical axis and time in seconds on the horizontal axis. (A) Underwater roar produced spontaneously during the spring breeding season; the sound is a guttural growl that ramps up in amplitude over 6 seconds to end in three broadband pulses. (B) Example of the grownl-like conditioned vocalization that was produced in air when the seal was given the S_D "speak" prior to selective shaping of the call structure. Note the presence of a discriminative stimulus (a verbal cue) and the conditioned reinforcer (a whistle) given by a human trainer, which are indicated in (C) and (D). (C) Example of an airborne vocalization that was relatively novel compared to previously reinforced vocal responses; this sound was relatively pulsed and tonal and sounded similar to a motorboat. (D) Example of the selectively shaped airborne vocalization that had vowel-like qualities; the S_D was the verbal cue "wa-wa-wa" given by the trainer, and the seal's response was a vocalization that sounded like "wa-wa-wa-wa."

“speak” was relatively stereotyped (see figure 3.1B). During the initial experimental sessions, we recorded these “typical” or baseline vocal responses onto videotape as well as with a directional microphone connected to a digital audiotape recorder. Once we had established his baseline sound production in response to the S_D , we began reinforcing only relatively novel variants of his conditioned vocal response following presentation of the discriminative cue. This contingency led to a very rapid expansion of the sounds produced in response to the S_D and a persistence in vocal behavior until reinforcement was provided. We occasionally encouraged movement of the mouth by gently prompting him with a light touch to move his mouth while vocalizing. As he explored which sound types would result in food reinforcement, the sounds Sprouts produced varied gradually and continuously between growls, snores, sneezes, sputters, moans, and pulses (see, e.g., figure 3.1C). The sounds also became progressively more varied along the dimensions of amplitude, frequency, modulation, and rate. He continued to change his vocalizations on each trial until he was signaled with a whistle that a fish reward would be provided. During his second session of this type, he emitted a sound that had some vowel-like syllabic qualities (a more tonal “wa” that was emitted while the mouth was moving) and this sound was reinforced. At that point, we opted to begin selective shaping of this particular sound, and we continued to reinforce only variants of this sound. Sprouts rapidly learned to produce the “wa” sound in response to the S_D “speak,” and this sound was progressively shaped to become more tonal and more repetitive “wa-wa-wa-wa” (see figure 3.1D).

Once a vocal response such as Sprouts’ “wa-wa” is reliably produced, it can be placed under the control of a particular S_D . This S_D can be arbitrary (a cue that has no structural similarity to the sound being produced) or it can match certain qualities of the desired vocal response (e.g., the S_D could be a tape recording of a particular sound or a vocal cue that sounds similar to the desired response). In this fashion, different acoustic responses can be placed under the control of different discriminative cues. In this example, Sprouts’ “wa-wa” sound was eventually controlled by a similar verbal cue given by the trainer.

Case Study 2: Walrus Vocal Conditioning

Walrus are perhaps the most interesting species among the pinnipeds from the standpoint of sound production, having multiple well-developed modes of sound production, which include manipulation of the larynx, lips, mouth, nose, tongue, and specialized pharyngeal pouches. The soft tissue structures of the mouth are especially muscular and mobile, which is likely a by-product of adaptations for suction feeding. Walrus produce a great diversity of sounds in air and underwater using various combinations of these structures. The sounds produced in the larynx and shaped by the mouth include highly variable barks, coughs, grunts, guttural sounds,

and roars. Highly differentiated whistles are produced by blowing air in and out over the tongue and the lips while a variety of sucking and smacking sounds are produced by movement of the entire mouth. Distinctive taps, knocks, pulses, and gong-like sounds (which are often emitted underwater or at the water's surface) are produced by controlled air movement associated with the pharyngeal sacs (see Tyack and Miller, 2002). At times the gong-like sound is augmented by a flipper striking the throat.

The breeding system of walruses can be classified as polygynous, and adult walruses exhibit a pronounced degree of sexual dimorphism. Males produce diverse and elaborate acoustic displays above and below the water's surface near ice edges where females congregate (Stirling, Calver, and Spencer, 1987). Reproductive behavior has been described but not well studied. Walruses are not easy animals to observe in the wild, especially during the spring breeding season because they gather in groups on the pack (free-floating) sea ice to give birth and mate. Of all pinniped species, they have the longest period of maternal dependency, with mothers nursing their pups for up to three years. Such an extended period of maternal dependency might provide an opportunity for elaboration of a sound-producing ability by the maturing walrus that with development could be considerably more flexible and spontaneous than in other pinniped species.

We had the opportunity to observe the vocal development of four captive walruses at Six Flags Marine World in Vallejo, California. The walruses were orphaned in 1994 and recovered from the open pack ice off the coast of Saint Lawrence Island, Alaska. The four animals (one male and three females) were approximately two weeks of age when they were brought into captivity. The pups were hand-raised on formula provided by human caretakers until they were weaned at approximately one year of age. The pups produced short barks as their primary spontaneous vocalizations for the first three to five years of their lives. As the walruses matured beyond that point, their sound-producing repertoires in air and underwater became more varied and began to reflect more of the sound types found in wild populations, including buzzes, growls, gulps, moans, whistles, barks, pulses, and sputters. At the time of these observations, the walruses were eleven years old. The male had begun to seasonally produce a variant of the knocking noises produced by rutting mature males using the pharyngeal pouches; the metallic gong-like sound used in adult acoustic displays had yet to emerge.

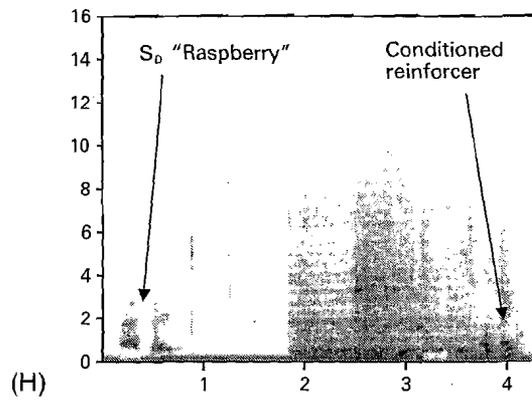
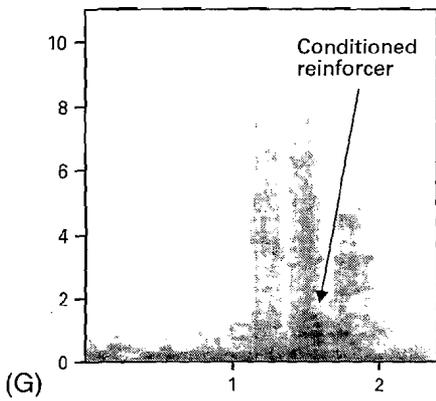
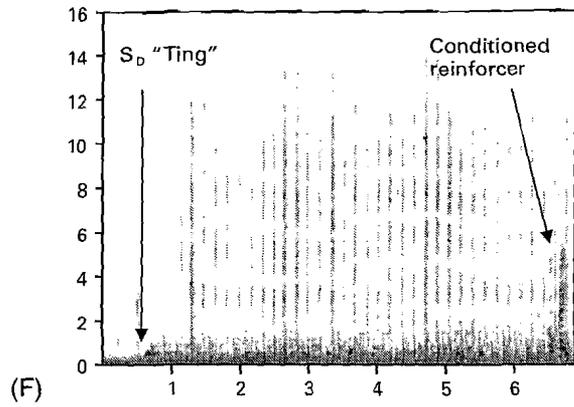
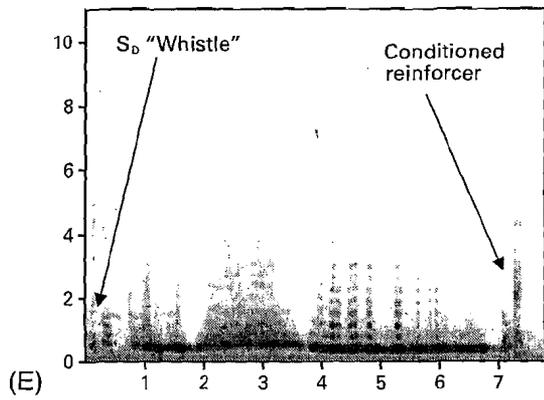
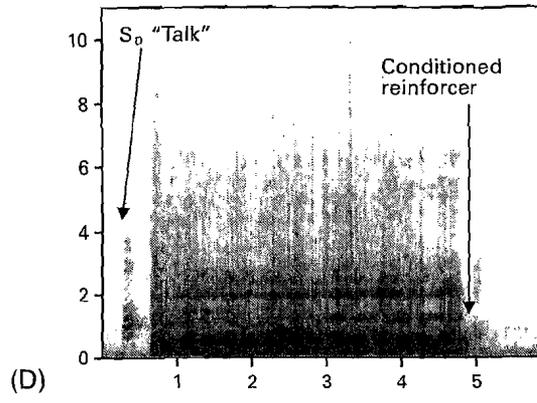
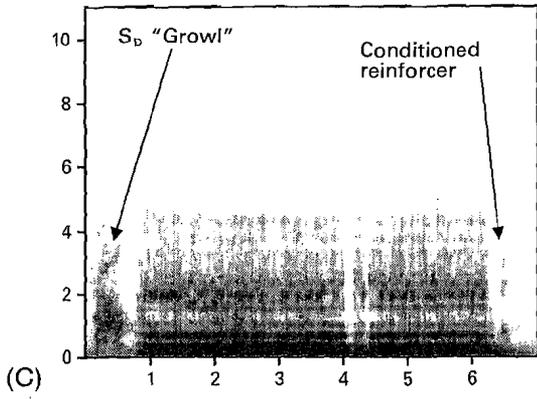
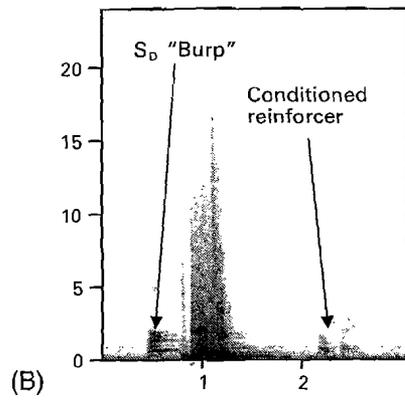
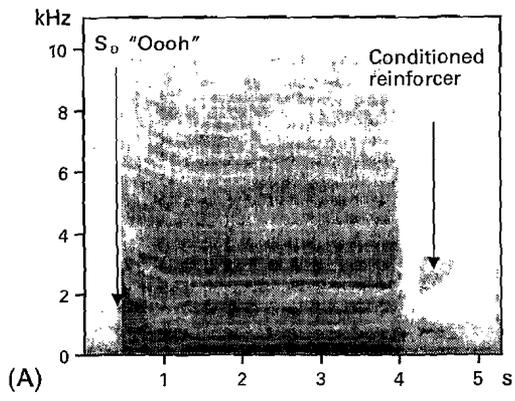
All four of these captive walruses were conditioned to produce sounds in air using food reinforcement. This process occurred in the context of an ongoing husbandry training program. Vocal conditioning began when the walruses were approximately two years old through selective reinforcement of vocalizations that occurred in the context of feeding situations. Delay of food reward increased vocal behavior, and the walruses quickly learned the relationship between sound production and feeding.

The male walrus, "Sivuqaq," was particularly engaged in this training process, but the training of the females followed the same general process.

Once Sivuqaq learned to reliably produce sounds in response to a trainer's prompts, four initial call types were placed under the control of four different discriminative stimuli over a period of about ten months. These call types were coded as "ooh," "growl," "talk," and "burp." All of these calls were shaped from undifferentiated guttural sounds produced in the larynx and emitted through the mouth. The calls were selectively shaped along particular dimensions—for example, "ooh" was shaped for acoustic elements including longer duration, louder intensity, and lack of movement of the mouth and lips during sound emission (see figure 3.2A). The trainer also cued the animal to position itself in an upright posture with its head oriented vertically and his mouth open when giving this vocal response. The call type "burp" was similarly shaped using selective reinforcement to be a very brief, loud, guttural sound (see figure 3.2B). The vocalizations "growl" and "talk" were both conditioned with Sivuqaq in a head-lowered posture. These were both shaped to be lower intensity sounds. "Growl" was shaped to include more moan-like qualities and was emitted with the mouth mainly closed (see figure 3.2C). "Talk" was a similar call that was selectively shaped to include modulation by rapid movement of the mouth and lips with the mouth partially open (see figure 3.2D).

Over the next several years, Sivuqaq was conditioned to produce at least four more distinctive acoustic emissions in response to different discriminative stimuli (see figures 3.2E, 3.2F, 3.2G, and 3.2H). One of these sounds, "whistle," was shaped by selectively shaping mouth position while simultaneously reinforcing exhalation (see figure 3.2E). Another call, coded as "ting," was also shaped as Sivuqaq became more mature. This was a metallic sound produced by the pharyngeal pouches, and it was shaped by prompting Sivuqaq to assume a chest-out posture with his mouth closed and then opportunistically reinforcing small movements and sound emissions originating from near the throat. The selective shaping of this sound resulted in a rapid series of pulsed "tings" correlated with visible movement of the chest and neck (see figure 3.2F). Two other calls, "sniffle" and "raspberry," were shaped by physical manipulation of the nose and mouth during exhalations combined with differential reinforcement of sound emissions that occurred either through nose only ("sniffle"; see figure 3.2G) or through the mouth and lips with the tongue extended ("raspberry"; see figure 3.2H).

Sivuqaq's accuracy at emitting the correct sound (which, of course, involved using the correct anatomical structures) in response to a given discriminative stimulus was quite high. He emitted the correct response at least 80% of the time even with very limited rehearsal. The conditioned vocal responses of the females followed a similar pattern of acquisition and performance. Interestingly, the female walruses were trained to produce at least two distinct sounds types, whistles and tings, not reported



to be emitted by wild female walruses, and not in the natural repertoire of these individuals prior to conditioning.

Summary and Conclusions of Case Studies

The preliminary observations gleaned from these two ongoing studies reveal the process by which vocal responses can be modified when reinforcement contingencies are related to specific call characteristics. They demonstrate the ease with which vocal responses can come to be produced in response to arbitrary experimental contexts, and they highlight the potential for selective shaping to influence aspects of vocal production including vocalizations produced at the levels of the respiratory tract, the larynx, and the supralaryngeal structures of the oral and nasal cavities that may serve to filter emitted sounds. The results add to the growing body of evidence suggesting that pinnipeds do have some degree of voluntary control over their sound emissions. The example described for the harbor seal is an interesting case. Here is a subject raised for most of his life without conspecifics who still displays apparently species-typical seasonal vocal behavior patterns in response to intrinsic (hormonal) or nonsocial environmental (photoperiod) cues. Despite this apparently innate pattern of sound production, the findings described here indicate that his vocal behavior is still readily accessible to modification by reinforcement contingencies, showing that both fixed and flexible vocal responses can occur concomitantly within an individual under the right set of circumstances. The walrus observations reveal plasticity in sound production in a particularly vocal species with a strong potential for complex communicative interactions during development. This leads to the question of whether reinforcement contingencies may come into play in natural social settings and influence vocal emissions.

The Ontogeny of Vocal Conditioning

If vocal conditioning is relevant in natural settings, then it is reasonable to wonder when this process might begin. Consider the following examples drawn from observations of pinniped behavior during the period of pup dependency. As noted earlier, there is a constant threat of mother-pup separation on a crowded sea lion rookery, resulting in frequent calling between the individuals involved. In most cases, the

Figure 3.2

Spectrograms of eight different types of conditioned vocalizations produced in air by an 11-year-old male Pacific walrus named Sivuqaq. All spectrograms show frequency in kilohertz on the vertical axis and time in seconds (s) on the horizontal axis. Also included are the discriminative stimuli (S_D) and conditioned reinforcers that were verbal cues given by a human trainer. The sounds were identified by their discriminative stimuli as follows: (A) "ooh"; (B) "burp"; (C) "growl"; (D) "talk"; (E) "whistle"; (F) "ting"; (G) "sniffle"; and (H) "raspberry" (note that no discriminative stimulus was recorded for G).

vocal exchange between a mother and her pup is terminated when the pup's calling is positively reinforced with nursing and contact comfort. For example, in northern elephant seals, a close physical association between the mother and her pup is formed soon after parturition and is maintained throughout lactation. The female and pup sleep and rest together, and she responds to its nuzzling and suckling attempts by assuming a nursing position, by rolling on her side and exposing her ventrum toward the pup's face. Given these cues, the pup begins to suckle in a reflexive manner. From the nursing female's viewpoint, these first attempts by the pup to suckle may act as a tactile cue eliciting the milk let-down reflex. This glandular response moves the milk toward her nipple, where it becomes available to the suckling pup. Once successful suckling has been established, a female that has been resting or sleeping while lying on its ventrum frequently waits until the pup emits a high-pitched distress cry before rolling on her side and exposing her teats. I speculate that once vocal utterances precede nursing, the milk let-down reflex of a lactating female may become gradually conditioned as a respondent (see Skinner, 1938) to the cries of the pup. At any rate, the dependent pup is likely to learn that its sound emissions can exert control over the female's orientation postures, giving it ready access to the teat and subsequent nursing. This type of relationship between vocalizing and nursing that I have described in northern elephant seals is ubiquitous and found not only in pinnipeds but in nearly all mammals.

The nursing calls of mammals are similar in many ways to the nursing cries of human infants in that all such calls signal need and their purpose is to elicit attention. Because they are correlated with access to milk, these highly stereotyped and undifferentiated nursing calls may become operantly conditioned. Presently, little is known about the role of learning in the vocal ontogeny of most mammalian species (Boughman and Moss, 2003). However, it is possible that the early relationship established between calling and nursing is relevant to changes in vocal behavior that may occur with maturation.

For example, in some songbirds where there is ontogenetic continuity between food begging and subsong, the origin of song learning in birds can be traced back to the earliest sounds they produce, namely, food-begging calls (Nottebohm, 1972). If this is true for some birds, then perhaps analogous ontogenetic processes in mammals may help us to trace how nursing calls in neonatal mammals play a role in the more complex development of vocalizations by mature individuals.

When we account for such early experiences, it becomes clear that the findings discussed earlier showing discriminative control of vocal behavior established with food reinforcement may not be inconsistent with theories on biological constraints on learning; rather, these findings suggest that one of the original functions of vocal behavior is related to nursing. Thus, infant calls may modify and be modified by nursing and social interactions. This synergistic interaction may be the earliest known social mechanism for discriminative vocal behavior.

General Discussion

Animal vocal communication as a form of social learning has been divided into three kinds of phenomena: receiver comprehension of calls, call usage, and call production (Seyfarth and Cheney, 1997; Janik and Slater, 2000). The first deals with behavioral changes that occur as a result of learning about the function of certain calls that are perceived, the second and third deal with changes that occur as a result of learning that affects the context and structure of calls that are produced. There has been general agreement based on evidence for a variety of mammals that learning and memory play a very important role in responses to the calls of others, that is, in the comprehension of vocalizations. In contrast, vocal usage and vocal production show much less flexibility in nonhuman mammals as compared to birds and humans.

The traditional ideas used to explain this lack of vocal plasticity have been that the production and usage of vocalizations by mammals is innate, that is, relatively fixed with little potential for any modification by learning. The only changes in vocal responses believed to occur were dependent on maturational factors and not experiential factors. These ideas were promoted by investigators who primarily researched vocal learning in terrestrial species of mammals, particularly nonhuman primates, and who did not take a broad comparative approach that included vocal learning in marine mammals. Having begun my career working with primates, I probably would have persisted in this thinking had I not had the opportunity to shift my efforts to the study of marine mammals. It was the research of John Lilly that first brought to my attention the ease with which the operant conditioning paradigm could be used to study the significance of vocal usage and production learning in marine mammals such as dolphins (Lilly and Miller, 1962). Today, the important role of learning and experience in sound production and usage by cetaceans such as dolphins has been well documented (McGowan and Reiss, 1997; Tyack and Sayigh, 1997), and these findings have been viewed as significant exceptions to traditional ideas about mammalian vocal communication. Until now, however, aside from the unusual case of Hoover, the harbor seal, much less has been made of experimental evidence for call production and usage learning of another marine mammal group—the pinnipeds (see Schusterman, 1978; Rawls et al., 1985). In this chapter there has been an attempt to make up for this deficiency by demonstrating that several different pinniped species show remarkable flexibility in the ways that they can use and modify their vocal emissions. In general, the material reviewed here has shown that changes in vocal responses may occur as a result of contingency learning using operant conditioning paradigms with food reward as a positive reinforcer.

Pierce (1985) argues that food reward may not be an appropriate reinforcer for modifying the ways that calls are used and produced, particularly in nonhuman primates. I counter this argument with the hypothesis that perhaps all mammalian infants acquire some experience with call usage in their nursing and contact-comfort

interaction with their mothers. This could explain why all mammals that have been studied using operant paradigms show some degree of vocal usage learning, even though most nonhuman terrestrial mammals do not appear to learn how to use their calls as efficiently as some diving mammals like whales, dolphins, seals, sea lions, and walruses.

Why do marine mammals appear more likely to learn to use and produce vocalizations compared to nonhuman terrestrial mammals? Hypotheses related to this question are fairly straightforward but clearly require further comparative research. This needs to be done not only on the pinnipeds but also on those terrestrial mammals like elephants and bats that depend on vocal communication to a much greater extent than some other terrestrial mammals. Without knowing the detailed structural mechanisms that produce pinniped vocalizations, we can still point to at least two important adaptations that are probably strongly involved in voluntary control and modification of vocal emissions both in air and underwater. These are adaptations related to (1) breath control and therefore laryngeal control and (2) control over articulatory movements of tongue, lips, mouth, and teeth occurring in the oral cavity above the level of the larynx. We look forward in the near future to collaborative studies in which we tease out variables related to the mechanics of laryngeal control, supralaryngeal filtering, and the brain structures mediating voluntary vocal behavior. These studies should clarify the degree to which pinnipeds, rather than nonhuman primates, have a capacity for vocal production learning, which more closely resembles that of humans. Such convergence in the precise control of vocal emissions, although probably related—at least in part—to diving and feeding activities in pinnipeds, should provide clues to the type of morphological, neural, and behavioral mechanisms designed for human spoken language.

Acknowledgments

First, I would like to thank Colleen Reichmuth from Long Marine Laboratory in Santa Cruz, California, and Debbie Quihuis from Six Flag Marine World in Vallejo, California, for being involved with most of the training of the harbor seal and the four walruses who served as subjects of this investigation. The research reported here could not have been accomplished without the participation of these people as well as the animals. Next, I want to thank Marla Holt for working with me on the spectrographic analysis of the sounds and for helping, along with Colleen Reichmuth, with the preparation of this chapter. Colleen's reorganization of some of my earlier material was particularly important for the structure and meaning of the final chapter. In addition to these members of the team, I would also like to thank Kristy Lindemann for her assistance at both Long Marine Laboratory and Marine World. Kathy Streeter from the New England Aquarium provided me with information

about harbor seal vocal conditioning that was extremely helpful. Tecumseh Fitch and Vincent Janik sent me several preprints and reprints as well as messages of encouragement that helped to guide me in all stages of writing this chapter. All of the reviewers who commented on the original version provided invaluable feedback. I especially want to thank Michael Owren for his extremely detailed and thorough critique. Last, but not least, I want to thank Uli Griebel and Kim Oller for inviting me to participate in the wonderful workshop on "The Evolution of Communicative Creativity" held at the Konrad Lorenz Institute in Altenburg, Austria, hosted by Gerd Müller. The comradeship of the participants was splendid, and the organizers provided a very entertaining program including tours of a castle and a monastery on the Danube and dining on fine food and wine in and around Vienna, Austria. This research, writing, and chapter preparation were partially supported by Office of Naval Research Grant N00014-04-1-0284.

References

- Adret P. 1993a. Vocal learning induced with operant techniques: An overview. *Netherlands Journal of Zoology* 43: 125–42.
- Adret P. 1993b. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour* 46: 149–59.
- Baptista LF, Petrinovich L. 1984. Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Animal Behaviour* 32: 172–81.
- Bartholomew GA. 1970. A model for the evolution of pinniped polygyny. *Evolution* 24: 546–59.
- Bolles RC. 1970. Species-specific defense reactions and avoidance learning. *Psychological Review* 77: 32–48.
- Boughman JW, Moss CF. 2003. Social sounds: Vocal learning and development of mammal and bird calls. In *Acoustic Communication*, ed. J Simmons, AN Popper, RR Fay, pp. 138–224. New York: Springer-Verlag.
- Dawkins MS. 1995. *Unraveling Animal Behaviour*, 2nd edition. Harlow: Longman.
- Farley GR, Barlow SM, Netsell R, Chmelka JV. 1992. Vocalizations in the cat: Behavioral methodology and spectrographic analysis. *Experimental Brain Research* 89: 333–40.
- Fitch WT. 2000. The evolution of speech: A comparative review. *Trends in Cognitive Sciences* 4: 258–67.
- Fitch WT, Hauser MD, Chomsky N. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97: 179–210.
- Goldstein MH, King AP, West MJ. 2003. Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences* 100: 8030–5.
- Hammerschmidt K, Freudenstein T, Juergens U. 2001. Vocal development in squirrel monkeys. *Behaviour* 138: 1179–1204.
- Hanggi EB, Schusterman RJ. 1994. Underwater acoustic displays and individual variation in male harbour seals, *Phoca vitulina*. *Animal Behaviour* 48: 1275–83.
- Harlow HF. 1949. The formation of learning sets. *Psychological Review* 56: 51–65.
- Hauser MD, Chomsky N, Fitch WT. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569–79.
- Hiss A. 1983. Hoover. *New Yorker* 3 Jan: 25–7.

- Innsley SJ, Phillips AV, Charrier I. 2003. A review of social recognition in pinnipeds. *Aquatic Mammals* 29: 181–201.
- Janik VM, Slater PJB. 1997. Vocal learning in mammals. In *Advances in the Study of Behavior*, ed. PJB Slater, JS Rosenblatt, CT Snowdon, M Milinski, pp. 59–99. San Diego, CA: Academic Press.
- Janik VM, Slater PJB. 2000. The different roles of social learning in vocal communication. *Animal Behaviour* 60: 1–11.
- Lenneberg EH. 1967. *Biological Foundations of Language*. Oxford: Wiley.
- Lilly JC, Miller AM. 1962. Operant conditioning of the bottlenose dolphin with electrical stimulation of the brain. *Journal of Comparative and Physiological Psychology* 55: 73–9.
- Manabe K, Kawahima T, Staddon JER. 1995. Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior* 63: 111–26.
- Manabe K, Staddon JER, Cleveland JM. 1997. Control of vocal repertoire by reward in budgerigars (*Melopsittacus undulates*). *Journal of Comparative Psychology* 111: 50–62.
- Marler P. 1990. Song learning: The interface between behavior and neuroethology. *Philosophical Transactions of the Royal Society of London B* 329: 109–14.
- Marler P, Tamura A. 1962. Song “dialects” in three populations of white-crowned sparrows. *Science* 146: 1483–6.
- McGowan D, Reiss D. 1997. Vocal learning in captive bottlenose dolphins: A comparison with humans and nonhuman animals. In *Social Influences on Vocal Development*, ed. CT Snowdon, M Hausberger, pp. 178–207. Cambridge, England: Cambridge University Press.
- Molliver ME. 1963. Operant control of vocal behavior in the cat. *Journal of the Experimental Analysis of Behavior* 6: 197–202.
- Moore BR. 1996. The evolution of imitative learning. In *Social Learning in Animals: The Roots of Culture*, ed. CM Heyes, BG Galef, pp. 245–65. San Diego, CA: Academic Press.
- Myers SA, Horel JA, Pennypacker HS. 1965. Operant control of vocal behavior in the monkey. *Psychonomic Science* 3: 389–90.
- Nottebohm F. 1972. The origins of vocal learning. *American Naturalist* 106: 116–40.
- Owren MJ, Dieter JA, Seyfarth RM, Cheney DL. 1993. Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychology* 26: 389–406.
- Pepperberg I. 1986. Acquisition of anomalous communicating systems: Implications for studies on interspecies communication. In *Dolphin Cognition: A Comparative Approach*, ed. RJ Schusterman, JA Thomas, FG Wood, pp. 289–302. London: Lawrence Erlbaum Associates.
- Perry EA, Renouf D. 1988. Further studies of the role of harbour seal (*Phoca vitulina*) pup vocalizations in preventing separation of mother–pup pairs. *Canadian Journal of Zoology* 66: 934–8.
- Pierce JD. 1985. A review of attempts to condition operantly alloprimate vocalizations. *Primates* 26: 202–13.
- Rawls K, Fiorelli P, Gish S. 1985. Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology* 63: 1050–6.
- Salzinger K. 1973. Animal communication. In *Comparative Psychology: A Modern Survey*, ed. DA Dewsbury, DA Rethlingshafer, pp. 161–93. New York: McGraw-Hill.
- Salzinger K, Waller BW. 1962. The operant control of vocalization in the dog. *Journal of the Experimental Analysis of Behavior* 5: 383–9.
- Schusterman RJ. 1967. Perception and determinants of underwater vocalization in the California sea lion. In *Les Systemes Sonars Amimaux, Biologie et Bionique*, ed. RG Busnel, pp. 535–76. Jouy-en-Josas, France: Laboratoire de Physiologie Acoustique.
- Schusterman RJ. 1977. Temporal patterning in sea lion barking (*Zalophus californianus*). *Behav Biol* 20: 404–8.
- Schusterman RJ. 1978. Vocal communication in pinnipeds. In *Studies of Captive Wild Animals*, ed. H Markowitz, V Stevens, pp. 247–309. Chicago: Nelson Hall.

- Schusterman RJ, Balliet RF. 1969. Underwater barking by male sea lions (*Zalophus californianus*). *Nature* 222: 1179–81.
- Schusterman RJ, Balliet RF. 1970. Conditioned vocalization as a technique for determining visual acuity thresholds in the seal lion. *Science* 169: 498–501.
- Schusterman RJ, Dawson RG. 1968. Barking, dominance, and territoriality in male sea lions. *Science* 160: 434–6.
- Schusterman RJ, Feinstein SH. 1965. Shaping and discriminative control of underwater click vocalizations in a California sea lion. *Science* 150: 1743–4.
- Schusterman RJ, Gentry R. 1970. Development of a fatted male phenomenon in California sea lions. *Developmental Psychobiology* 4: 333–8.
- Schusterman RJ, Balliet RF, Nixon J. 1972. Underwater audiogram of the California sea lion by the conditioned vocalization technique. *Journal of the Experimental Analysis of Behavior* 17: 339–50.
- Schusterman RJ, Balliet RF, St. John S. 1970. Vocal displays underwater by the gray seal, harbor seal and Steller sea lion. *Psychonomic Science* 18: 303–5.
- Schusterman RJ, Kastak D, Levenson DH, Reichmuth CJ, Southall BL. 2000. Why pinnipeds don't echolocate. *Journal of the Acoustical Society of America* 107: 2256–64.
- Schusterman RJ, Reichmuth Kastak C, Kastak D. 2003. Equivalence classification as an approach to social knowledge: From sea lions to simians. In *Animal Social Complexity: Intelligence, Culture, and Individualized Societies*, ed. FBM DeWaal, PL Tyack, pp. 179–206. Cambridge, MA: Harvard University Press.
- Schusterman RJ, Van Parijs SM. 2003. Pinniped vocal communication: An introduction. *Aquatic Mammals* 29: 177–80.
- Seyfarth RM, Cheney DL. 1986. Vocal development in vervet monkeys. *Animal Behaviour* 34: 1640–58.
- Seyfarth RM, Cheney DL. 1997. Some general features of vocal development in nonhuman primates. In *Social Influences on Vocal Development*, ed. CT Snowdon, M Hausberger, pp. 249–73. Cambridge, England: Cambridge University Press.
- Shapiro AD, Slater PJB, Janik VM. 2004. Call usage learning in gray seals (*Halichoerus grypus*). *Journal of Comparative Psychology* 118: 447–54.
- Skinner BF. 1938. *The Behavior of Organisms*. New York: Appleton-Century-Crofts.
- Skinner BF. 1957. *Verbal Behavior*. East Norwalk, CT: Appleton-Century-Crofts.
- Snowdon CT, Hausberger M. 1997. *Social Influences on Vocal Development*. Cambridge, England: Cambridge University Press.
- Stirling I, Calver W, Spencer C. 1987. Evidence of stereotyped underwater vocalizations of male Atlantic walrus (*Odobenus rosmarus rosmarus*). *Canadian Journal of Zoology* 65: 2311–21.
- Thorndike EL. 1911. *Animal Intelligence*. New York: Macmillan.
- Tyack PL, Miller EH. 2002. Vocal anatomy, acoustic communication and echolocation. In *Marine Mammal Biology: An Evolutionary Approach*, ed. AR Hoesel, pp. 142–84. Oxford: Blackwell Science.
- Tyack PL, Sayigh LS. 1997. Vocal learning in cetaceans. In *Social Influences on Vocal Development*, ed. CT Snowdon, M Hausberger, pp. 208–33. Cambridge, England: Cambridge University Press.
- Van Parijs SM, Schusterman RJ. 2003. Aquatic Mammals: Special Issue on Animal Vocal Communication, ed. SM Van Parijs, RJ Schusterman, *Aquatic Mammals* 29: 1–319.
- West MJ, King AP. 1985. Social guidance of vocal learning by female cowbirds: Validating its functional significance. *Ethology* 70: 225–35.
- West MJ, King AP. 1996. Social learning: Synergy and songbirds. In *Social Learning in Animals: The Roots of Culture*, ed. CM Heyes, BG Galef, pp. 155–78. San Diego, CA: Academic Press.
- Wilbrecht L, Nottebohn F. 2003. Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Review* 9: 135–48.
- Yamaguchi S, Myers RE. 1972. Failure of discriminative vocal conditioning in rhesus monkey. *Brain Research* 37: 109–14.