



REVIEW

The different roles of social learning in vocal communication

VINCENT M. JANIK* & PETER J. B. SLATER†

*Biology Department, Woods Hole Oceanographic Institution

†School of Biology, University of St Andrews

*(Received 6 October 1999; initial acceptance 19 November 1999;
final acceptance 2 February 2000; MS. number: RV-14)*

While vocal learning has been studied extensively in birds and mammals, little effort has been made to define what exactly constitutes vocal learning and to classify the forms that it may take. We present such a theoretical framework for the study of social learning in vocal communication. We define different forms of social learning that affect communication and discuss the required methodology to show each one. We distinguish between contextual and production learning in animal communication. Contextual learning affects the behavioural context or serial position of a signal. It can affect both usage and comprehension. Production learning refers to instances where the signals themselves are modified in form as a result of experience with those of other individuals. Vocal learning is defined as production learning in the vocal domain. It can affect one or more of three systems: the respiratory, phonatory and filter systems. Each involves a different level of control over the sound production apparatus. We hypothesize that contextual learning and respiratory production learning preceded the evolution of phonatory and filter production learning. Each form of learning potentially increases the complexity of a communication system. We also found that unexpected genetic or environmental factors can have considerable effects on vocal behaviour in birds and mammals and are often more likely to cause changes or differences in vocalizations than investigators may assume. Finally, we discuss how production learning is used in innovation and invention, and present important future research questions.

© 2000 The Association for the Study of Animal Behaviour

Social learning and, in particular, imitation has received considerable research attention in recent years (Heyes & Galef 1996; Box & Gibson 1999). One reason for this is the assumption that exhibiting true motor imitation is an indication that an animal has a theory of mind (Whiten & Ham 1992). Since motor patterns can be learned socially or individually, researchers have taken great care to develop a detailed theoretical framework, within which they define different forms of individual and social learning and discuss the methodology required to show each one (Whiten & Ham 1992; Heyes 1998). However, one form of social learning, vocal learning, has been studied intensively for a long time from a different perspective: to elucidate the interactions between nature and nurture.

Much of the work on vocal learning has been on birds, a group in which many versatile vocal learners can be

Correspondence: V. M. Janik, Woods Hole Oceanographic Institution, Biology MS 34, Woods Hole, MA 02543, U.S.A (email: vjanik@whoi.edu). P. J. B. Slater is at the School of Biology, University of St Andrews, Fife KY16 9TS, U.K.

found. In contrast to other motor imitation in animals, the occurrence of vocal learning in songbirds is so obvious that very little effort has focused on defining what constitutes vocal learning and classifying the forms that it may take. This lack of a clear framework has carried over into studies on mammals, where the evidence for vocal learning is less clear and studies have concentrated more on learning when to use signals rather than how to produce them. This difference of perspective has often led to confusion over terms or overinterpretation of results. But, even in songbirds, many aspects of song are not learned (Marler & Pickert 1984; Marler & Sherman 1985; Podos 1996; Nowicki et al. 1999) and learning abilities can differ between populations of the same species (Kroodsma & Canady 1985; Nelson et al. 1996).

The variety of ways that social learning may influence vocal communication, and of possible research foci, shows that a clear theoretical framework is needed to be able to compare species and to understand the factors influencing the occurrence of learning in animal communication. In this review, we give such a theoretical

Table 1. Definitions of different forms of social learning in animal communication

Form of learning	Definition
Contextual learning	An existing signal is associated with a new context as a result of experience with the usage of signals by other individuals
Usage learning	An existing signal is produced in a new context as a result of experience with the usage of signals by other individuals
Comprehension learning	A receiver comes to extract a novel meaning from a signal as a result of experience with the usage of signals by other individuals
Production learning	Signals are modified in form as a result of experience with those of other individuals. This can lead to signals that are either similar or dissimilar to the model

framework for the study of how social learning influences vocal communication and outline the methods that can provide evidence that it occurs. Social learning has been defined in many different ways (e.g. Whiten & Ham 1992), all of which require the presence of another individual. Many forms of learning qualify as social under these definitions. However, even if we apply a much narrower definition that includes only learning in which information that is necessary for the learning process is extracted from the behaviour of another individual, almost all learning about communication signals qualifies. Thus, apart from cases in which experience with one's own voice alone influences vocal behaviour, as occurs in practising, any learning of communication behaviour is a form of social learning. Note that an interaction is not necessary, since learning can occur by observation alone, and also that we consider learning from a simulated individual, as with tape recordings, to be social learning as well.

We are aware that some of the distinctions we make are in reality difficult to draw at this stage. However, we think that a detailed theoretical framework is helpful to focus research effort and we hope that our suggestions will lead to fruitful discussions. We do not discuss factors that influence the effectiveness of learning or the extent to which learning can influence vocal behaviour. Food aversion learning in rats, for example, is possible only with certain conditioning stimuli but not with others (Garcia & Koelling 1966). Similar constraints may also exist in learning that affects communication, but too little is known to discuss the possibility here. Furthermore, since vocal learning itself has been demonstrated only in birds and mammals, we aim our discussion and examples at these groups. Reviews of the evidence for vocal learning in these taxa have been provided elsewhere (Catchpole & Slater 1995; Janik & Slater 1997).

THE FORMS OF SOCIAL LEARNING IN COMMUNICATION

For our discussion of different forms of learning we use the terms message and meaning as defined by Smith (1977). There is a clear distinction between contextual learning and production learning in animal communication (Janik & Slater 1997). Contextual learning occurs when a pre-existing signal comes to be associated with a new context as a result of experience with the signals of

other individuals. This form of learning can occur in both the signaller and the receiver (Table 1). If a signaller produces a pre-existing signal in a new context this signal can take on a new message. We also refer to this form of contextual learning as usage learning, since the individual learns to use the same signal in different contexts to encode different messages. Similarly, a receiver can learn to associate a pre-existing signal with a new context. Again, this is contextual learning, but here we call it comprehension learning since the receiver learns to comprehend that a pre-existing signal is given in a new context and thus can extract a new meaning from it. Learning to recognize individuals from their signals is a form of such comprehension learning. Here, the identity of the signaller is the context with which specific signal characteristics are associated. Context can therefore mean different things in contextual learning. It can be the identity, age, sex or social status of a signaller. It can be a reference context, for example if monkeys give a leopard alarm call in the presence of a human hunter. It can be a new signal context, for example arranging signals in new combinations or sequences. Such new combinations can then convey new messages. Thus, one way that animals capable of contextual learning can increase their message repertoire is by rearranging pre-existing signals into new sequences (see below).

While contextual learning refers only to learning the context in which existing signals are used, production learning refers to instances where the signals themselves are modified in form as a result of experience with those of other individuals (Janik & Slater 1997). It is this production learning to which the expression 'vocal learning' has traditionally been applied, for example, in studies of bird song. But here, to avoid confusion, we use the term production learning. It includes the production of novel signals that were developed to differ from those of specific individuals as well as those that match them. The distinctions between usage, comprehension and production learning apply to all forms of communication.

PRODUCTION LEARNING IN VOCAL COMMUNICATION

Like contextual learning, production learning can also be split up into different forms, but here the distinction is not between signaller and receiver but between different

Table 2. Vocal systems involved in production learning

System	Role in production learning
Respiratory system	Changing of sound parameters (e.g. duration and amplitude) by altering the state of respiratory muscles
Phonatory system	Changing of sound parameters (e.g. absolute frequency and modulation) by altering the state of the sound-producing structure (e.g. the syrinx or the larynx)
Filter system	Changing of sound parameters (e.g. relative energy distribution) by altering the state of filtering structures

systems that are involved in sound production, the contribution of which may be changed by learning (Table 2). In birds and mammals the sound production mechanism has at least two components: the respiratory system, which provides airflow under pressure, and a phonatory system, consisting of the sound production apparatus (in most cases the larynx or the syrinx) where a sound originates. In many cases a third component, the filter system, has substantial influence on the signal. This component includes all filtering or resonance structures between the phonatory organ and the outside world. Learning can modify sound production at any of these stages.

Each of these three components controls different aspects of a sound. The duration and amplitude of a signal can be changed by the respiratory system alone. The phonatory system has control over the fundamental frequency that is produced and its modulations. Finally, modification of the filter system can lead to changes in the relative energy distribution in the vocalization. Even though these different contributions seem simple, it can be difficult to assign a certain parameter change to one of these systems. For example, a novel tonal component in a vocalization can be produced by rapid amplitude modulation through the respiratory system, by an adjustment in the production apparatus or by changes in the filtering apparatus. Complex amplitude modulations often need an elaborate coordination between respiratory and phonatory systems, or can be caused by either system on its own. Similarly, changes in air pressure provided by the respiratory system can lead to changes of a few Hertz in fundamental frequency (Hsiao et al. 1994). Thus, to identify which of these systems has been altered by production learning is not an easy task. Nevertheless, we think it is important to distinguish between these different motor systems since they represent very different ways in which sounds can be modified.

There may even be different forms of production learning involved using separate brain areas. While in humans the primary motor cortex and a connection from it to the nucleus ambiguus are necessary for the production of learned vocalizations, such as speech or the humming of tunes, it is not for the production of vocalizations like crying or laughing (Groswasser et al. 1988). Nonhuman primates lack the direct connection between the primary motor cortex and the nucleus ambiguus, and lesions of the motor cortex, as well as of the main structures with projections to it (the cerebellum and the ventrolateral

thalamus) and from it (the putamen), have no effect on their vocalizations (review in Jürgens 1992). To date, there is no evidence for phonatory production learning in nonhuman primates. However, rhesus monkeys, *Macaca mulatta*, have conditional control over the duration and amplitude of their calls (Sutton et al. 1973). This suggests that production learning that alters the duration or amplitude of a call uses different areas to store or control vocalizations from the ones used for speech production in humans. To our knowledge, it is not known whether humans with a destroyed motor cortex maintain their ability to alter voluntarily the duration or amplitude of the vocalizations that they can produce. If they do, it would indicate that different forms of production learning exist. Changing the general duration and amplitude by learning to alter the state of the respiratory system may be an evolutionarily early form of production learning that is not dependent on the primary motor cortex in mammals, while more complex learning of temporal and frequency structure depends on the primary motor cortex and a newly evolved connection to the nucleus ambiguus (for a description of a similar connection between the telencephalon and vocal motor neurons in songbirds and its absence in nonsongbirds see Wild 1994).

The distinction of such different forms of learning can lead to interesting insights into possible pathways of the evolution of production learning. Even for usage learning, for example, the respiratory system and, in particular, triggering of the on- and offset of a vocalization must be under conditional control, so we may expect animals capable of changing usage also to be able to alter the duration and amplitude of their vocalizations. This is indeed the case for animals that have been tested for these forms of learning (e.g. cat, *Felis catus*: Molliver 1963; rhesus monkey: Sutton et al. 1973). Here, the modifications in duration or amplitude always seem to occur in parts of the signal that show very little frequency modulation. For complex changes in frequency parameters, however, conditional control over the phonatory system or the filter system, as well as close coordination with the respiratory system, is required, and this suggests that greater control over the vocal apparatus has evolved. To date, we know that many mammals and many birds are capable of usage learning; some species, such as cats and rhesus monkeys, are also capable of altering the duration and amplitude of their calls, while only a few, such as songbirds, humans and bottlenose dolphins, *Tursiops truncatus*, are able to modify the

frequency as well as the temporal and amplitude structure of their calls (reviews in Adret 1993; Catchpole & Slater 1995; Janik & Slater 1997).

Each form of learning can increase the complexity of a communication system and the more forms of learning present the greater the system's openness and flexibility. Even though either usage learning or production learning may lead to a richer message repertoire, a species that shows both usage learning and production learning can theoretically evolve a more complex communication system than one that shows only usage learning. A similar argument applies to comprehension learning. If an animal can learn to give an alarm call in the absence of a predator (usage learning), or can copy an alarm call of another species (production learning), it might benefit from increased access to food abandoned by other individuals in response to that call. With such deceptive signals the receiver will have increased uncertainty as to whether the information is reliable, but the repertoire of different messages that can be transmitted is no greater. Only if the receiver is capable of comprehension learning can a signal become established in the communication system as having a new meaning.

Our terminology is more precise than that used recently by some researchers, who have used the term vocal learning rather generally for any learning involving vocalizations including contextual learning (e.g. Snowdon et al. 1997; Elowson et al. 1998; Pepperberg et al. 1999). This terminology seems somewhat confusing, since it conflates several very different forms of learning (Table 1). We also think it is important to point out that even production learning may involve several different processes at markedly different levels of complexity. We hope that our terminology will help to clarify these processes so that the exact phenomenon involved in each group of animals can be determined and the distribution of production learning more precisely defined.

EVIDENCE FOR CONTEXTUAL LEARNING

Many studies have used either total isolation or lesioning of tissue to investigate whether an aspect of behaviour was influenced by learning. If the ontogeny of a behaviour pattern does not change in an isolation experiment, it gives good evidence that social learning is of little importance. Similarly, a lack of change in behaviour patterns after lesioning indicates that the lesioned structures are not crucial for the production of these patterns. This has been used to determine which parts of the brain are crucial for learning to occur. However, if changes or differences can be observed in such experiments it is not clear whether this is due to the lack of learning experiences. Other reasons why changes may be observed are the lack of more general social interactions in isolation experiments or the side-effects of a surgical procedure.

Like production learning (see below), contextual learning is demonstrated best with conditioning experiments. If an individual is able to produce a call in response to a conditioning stimulus it is capable of usage learning. Comprehension learning can be demonstrated if the animal can be conditioned to associate a played back call

with a new context, for example shows an aversion reaction if a playback call that did not elicit aversion before has been paired with an unpleasant stimulus in the experiment. Such experiments have shown that contextual learning can be found in many bird and mammal species (reviews in Adret 1993; Janik & Slater 1997). There may be two different stages in contextual learning. An animal may either associate any call with a new context, that is, just produce a random call in response to a conditioning stimulus rather than a specific one, or it could learn to associate a specific call with a specific stimulus. To our knowledge this difference has not been investigated systematically. Additionally, animals may be easier to condition with certain calls than with others depending on how they are used in the wild, or may only be able to associate call types with specific contexts. An alternative to conditioning is to use cross-fostering experiments. While Owren et al. (1993) found that call usage was not influenced by experience in cross-fostered rhesus and Japanese macaques, *M. fuscata*, Seyfarth & Cheney (1997) showed that comprehension learning was clearly possible.

Observational studies are often harder to interpret since it is difficult to be sure that the changes observed are not just due to maturational processes. Seyfarth & Cheney (1986) and Hauser (1989) presented very detailed observational data on the development of call usage and comprehension in vervet monkeys, *Cercopithecus aethiops*. They reported that young vervets showed a loose context specificity in their use of alarm calls and 'wrr' vocalizations. However, their reactions to such signals and their use became more refined with age. This suggests that contextual learning was involved but does not prove it since maturational effects cannot be excluded. A study comparing call usage in different groups of vervet monkeys, however, showed that infants in groups that are exposed more often to intergroup encounters showed the appropriate use of the intergroup 'wrr' much earlier (Hauser 1989). Similarly, infants that lived closer to swamps reacted much earlier to alarm calls of the superb starling, *Spreo superbus*, a species that prefers swamp habitat, than did those living in dry woodlands (Hauser 1988). Thus, observational data often need a control group to demonstrate contextual learning as is provided by this comparison of groups with different exposure levels. An exception is where animals learn to adapt their vocal behaviour to new stimuli in their environment that they cannot have adapted to genetically. An example are siamang, *Hylobates syndactylus*, duets in which the exact timing of the calls in response to the partner develops over time (Maples et al. 1989). This is possible only if the animals are capable of usage learning.

THE DISTINCTION BETWEEN PRODUCTION LEARNING AND CONTEXTUAL LEARNING

Given the potential increase in complexity of the communication system with each extra form of learning, it is important to know the learning capabilities of each species. If a single novel call is found and maturational processes can be excluded, the only question is whether

the call has been in the animal's repertoire before (see below). In a sequence of signals, however, as well as each signal, the arrangement of signals in time, or the syntax, can be heavily influenced by learning (Hultsch & Todt 1989; Podos et al. 1999). In a bird, for example, such a sequence can either be a sequence of elements that constitutes a song, or one of the songs that makes up a song sequence. Since such learning concerns only the sequential order of vocalizations, we consider it to be an example of contextual learning in which signals can be put into new sequences. This can, of course, occur at the same time as production learning of signal types. Even if the exact order of songs in a sequence of bird song, for example, is sometimes considered to be a by-product of the learning process, it nevertheless carries information on where the sequence has been learned. Thus, the sequential order of signals still carries a specific message. But how do we recognize what constitutes one signal? If a sequence consists of several elements it is not easy to decide whether an apparently novel vocalization arises because of a recombination of existing elements into a new sequence (contextual learning) or because of learning a new element sequence anew as one signal (production learning).

An apparently simple example is from animals that use different rhythmical patterns to communicate. Sperm whales, *Physeter catodon*, use click codas that consist of separate clicks arranged in stereotypic, rhythmical patterns (Watkins & Schevill 1977). The fact that a sperm whale has been found to match the click repetition rate of a human sonar signal (Backus & Schevill 1966) suggests that they are capable of contextual learning. Contextual learning would be sufficient to explain the variety of codas found. Whales can produce a new coda by emitting the same clicks in a new rhythmical pattern, thus changing the temporal arrangement of clicks but not the structure of the click vocalization itself. An example of an addition of a novel and structurally different element to a sequence of elements has been found in chimpanzee, *Pan troglodytes*, pant hoots (Marshall et al. 1999). Here, a new element that had previously occurred frequently on its own was incorporated into pant hoots by several group members after the introduction into the group of an animal that had this form of pant hoot in its repertoire.

But are these cases as clear cut as it seems? Did those chimpanzees simply take an existing element and put it into the element sequence that constitutes a pant hoot? Or did they learn the whole sequence again as a new signal through production learning? And do sperm whales use production or usage learning to develop new click patterns? The question here is whether the whole sequence or each separate element is the minimal unit of production (MUP; Barlow 1977) that can be recombined to form a novel signal. If the MUP is a whole sequence of elements and not just a single element, animals may learn different sequences containing the same elements separately, effectively learning the same elements twice. In that case we would conclude that the new sequence has been acquired by production learning and not contextual learning. However, if each separate element is a MUP, new sequences can be achieved by rearranging existing

MUPs in time through contextual learning. Slater (1983) argued that repeated learning of the same signal can occur in animals. From a detailed sequential analysis of chaffinch, *Fringilla coelebs*, song he concluded that chaffinches can learn certain song types twice and, as a result, have the same song type stored twice in their brains. However, he considered this to be a by-product of the learning mechanism and not a way to increase the message repertoire.

In the case of the chimpanzee pant hoot we would agree with Marshall et al. (1999) that it is a case of contextual learning, since the element frequently occurred on its own and since pant hoots are often interrupted and not finished, suggesting that each element is a separate unit of production. Similarly, in a sperm whale, a gradual change in click repetition rate before it matched that of the human sonar has been observed (Backus & Schevill 1966). This again suggests that each click is a separate unit supporting the idea that the change is due to contextual learning. However, the question of what is a minimal unit of production is not trivial and it is crucial for our definition of production learning. Methods that have been used to determine MUPs in birds include the use of a flashlight to see where a song can be interrupted (Cynx 1990; Riebel & Todt 1997) and acoustic analysis to determine sequential relationships between elements (Podos et al. 1992). MUPs can be very different from fundamental units of perception depending on the context. Adult male song sparrows, *Melospiza melodia*, for example, seem to perceive whole songs as fundamental units (Searcy et al. 1999), but young birds often combine elements from different tutors or song types to form new songs (Marler & Peters 1987; Beecher 1996). In song sparrows, MUPs can either be single elements or groups of elements that always occur together (Podos et al. 1992). Thus, it needs great care to decide whether an apparent rearrangement of elements is actually a novel MUP the production of which has been learned, or whether it arises from a change in syntax of existing MUPs achieved by contextual learning. In complicated cases the question of whether animals learn different combinations of elements separately, or whether they rearrange existing elements, may require neurobiological studies looking at how these units are stored in the brain.

As we can see from this discussion, contextual learning can lead to very complex communication patterns. Marler & Nelson (1992) argued that one songbird, the swamp sparrow, *Melospiza georgiana*, may use contextual learning extensively for the development of its song patterns. In this species, element variability is low while syntactical variability is not (Marler & Pickert 1984). Thus, to develop their song types, birds may recombine a pre-existing limited number of elements into new sequences. Marler & Nelson used the term action-based learning to emphasize the attrition process, in which an individual decides which elements to retain rather than copying elements anew. In our terminology this would qualify as usage learning since the bird learns only when to use what element or song type. We use the more general term 'usage learning' here since 'action-based'

suggests that individuals have to use patterns actively themselves in this process. However, animals could also learn when to use what element by observing other individuals using those signals. Nelson & Marler (1994) used another term, 'selection-based learning'. This term emphasizes the selection from an existing repertoire. Again, this is a specific form of usage learning in which the bird selects from a rich repertoire. The more general term 'usage learning', however, also includes cases where only one signal is present and only the context in which to use it is learned. The initial acquisition of the elements that are used in these forms of learning might or might not involve production learning. The bird species in which action- or selection-based learning has been described seem to memorize elements heard early in life and select later which ones to use and in what sequences. Thus, they use production learning for the initial acquisition and usage learning later on in the selection process.

EVIDENCE FOR PRODUCTION LEARNING

The Problem

To demonstrate production learning in practice is not always easy. The main question is how we determine that a particular sound is a new signal, that is, one that has not been in the individual's repertoire before. In cases where animals mimic atypical sounds, such as a harbour seal, *Phoca vitulina* (Ralls et al. 1985), or a grey parrot, *Psittacus erithacus* (Todt 1975), that produces human speech, the evidence is relatively clear. However, in most cases, the distinction is not a trivial one. Jakobson (1941) argued that even humans are born with all the sounds that they use later in life. Even though humans are clearly capable of production learning, for example copying patterns of frequency modulation, this may not be so in other animals. If usage learning is taking place, in which an animal selects signals only from a pre-existing repertoire without production learning being involved, studies on adult animals will be hard to interpret. An apparently new signal could just be one that has never been used before in adulthood. Thus, young and inexperienced animals are better subjects for the study of production learning. In theory, an animal could, of course, still have a signal in its pre-existing repertoire that accidentally resembles the experimental stimulus. But this problem can be avoided by using a wide range of signals or ones that are rather atypical for that species.

A problem here is to decide how strict we should be in our requirements of the evidence for production learning. Should a learner really be able to copy a completely atypical signal? While some animals have tremendous flexibility in their learning abilities, all sound production organs nevertheless have limits in what they can produce. But, even within the broad range set by such anatomical limitations, many animals have only a small set of signals in their vocal repertoire. They often cannot be conditioned to alter signals even slightly. However, it is likely that production learning, like contextual learning, is not an all-or-nothing trait and that there are cases in which

some parameters but not others can be influenced socially. Such cases could indicate possible pathways in the evolution of learning. In the first stages of the evolution of production learning, we would expect animals to produce signals that, while clearly recognizable as one signal type shared among individuals, differ in subtle learned aspects, such as having consistent differences in absolute fundamental frequency. This pattern can be found in greater horseshoe bats, *Rhinolophus ferrumequinum* (Jones & Ransome 1993) even though experiments with more varied sounds are lacking. To show learning of subtle differences is difficult since such parameters are usually also influenced by the affective state of the animal. Thus, differences between individuals or between contexts could also stem from differences in internal states without any influence of learning. Size differences can also influence such signal parameters consistently. Careful controls are thus necessary to show that production learning is really taking place.

Experimental Data

The best experimental evidence comes from studies in which animals have been conditioned to copy stimulus sounds or have been reared with playbacks of particular sounds. If production learning occurs the subject should be found to copy the stimulus. For a clear demonstration, the sounds used should be as atypical for that species as possible (see above). However, a demonstration that animals can alter single parameters gradually within a signal type is equally convincing. This can be shown best by using conditioning procedures. For an excellent example of such a conditioning study see Manabe et al. (1997).

If animals are reared with sound playbacks, at least two experimental groups are needed, each one reared with a different stimulus. A group that receives no playback is not a valid control since some form of acoustic stimulation might be necessary for vocal development even if production learning is not involved. It could also be problematic, both technically and ethically, to isolate an infant from its mother. In some cases a cross-fostering study may be the best way of investigating how learning influences vocal behaviour, provided that the foster parent and the real parent are sufficiently different in their vocalizations to demonstrate an influence on vocal development (Owren et al. 1992).

Observational Data

To show that production learning occurs from observational data is much more difficult. However, sometimes these are the only source of information available. Unfortunately observational data can almost never exclude the occurrence of usage learning based on a pre-existing repertoire. However, it may show this to be an unlikely explanation. Two cases in which production learning has been demonstrated simply by observation are in humpback whales, *Megaptera novaeangliae*, and greater horseshoe bats. Humpback whales produce elaborate songs during the breeding season (Payne & McVay 1971). These

songs change gradually over one season and are not repeated later in life (Payne et al. 1983; Payne & Payne 1985). But all whales in one population sing the same song at any given time (Guinee et al. 1983). This synchronous change of song in all individuals can be explained only by production learning. In the greater horseshoe bat, infants produce calls with the same fundamental frequency as their mothers (Jones & Ransome 1993). This frequency tends to decrease with age, so that infants born to the same female when she was young have higher fundamental frequencies than those born to her when she was older. Again, other than production learning, the only alternative explanation for these phenomena seems to be that the individuals learn to select from a repertoire of calls that was already present. However, after studying the changes in humpback whale song over 19 years, Payne & Payne (1985) concluded that changes are progressive and not cyclical. Even after such a long period they could still find new elements or units in songs. In the horseshoe bat, the calls are used in mother–infant recognition and learning seems to result in matching of frequencies. It seems unlikely that each bat has the same call stored with many different fundamental frequencies from which it selects the one most similar to that of its mother. However, further studies are certainly needed on the versatility of production learning in bats.

One phenomenon frequently taken as evidence for production learning is variation in vocalizations between groups or locations. This stems partly from the fact that such variation is common in birds that learn their songs. Thus, it can be indicative of production learning and deserves further analysis. However, such variation is far from being clear evidence for learning. Such differences might result for several other reasons, including genetic variation, founder effects and adaptation to environmental as well as social factors through contextual learning. Genetic variation is often discounted even though nothing is known about the relevant genetics of the individuals under study and how they affect vocalizations. However, there are several examples where variation in vocalizations is known to have a genetic basis. In the northern bobwhite quail, *Colinus virginianus*, differences in the call structure between familial lines are clearly caused by genetic differences and not by learning (Baker & Bailey 1987). Medvin et al. (1992) presented similar evidence for cliff swallows, *Hirundo pyrrhonota*. As for geographical variation, the difference in repertoire size and the style of delivery of song between marsh wrens, *Cistothorus palustris*, from California and New York (Kroodsmas & Canady 1985), and the difference in the structure of squirrel monkey, *Saimiri sciureus*, isolation calls from two different populations (Lieblich et al. 1980), also have a genetic basis. In both cases individuals showed the population-specific pattern even if reared in auditory isolation. Founder effects can occur if groups in different locations stem from very small core groups. In such a case genetic differences are likely and cannot be excluded as an explanation for vocal variations. In northern elephant seals, *Mirounga angustirostris*, such a case has even been documented in the wild after this species had been severely depleted by hunting and

then recolonized different breeding islands (Le Boeuf & Petrinovich 1974). Founder effects are also likely to cause differences in vocalizations between holding facilities.

Hansen (1979) hypothesized that learning could help to adjust signals to the sound transmission characteristics of the habitat. However, genetic factors, production learning or contextual learning could all help to achieve such adaptation if environmental factors do indeed influence the effectiveness of vocalizations. Animals that are only capable of contextual learning could adapt to environmental factors by using primarily those call types from their repertoire that transmit best and are thus most effective at eliciting responses. Unless extensive data sampling is conducted, this could easily result in the assumption that certain call types are site specific. However, even though several studies have shown geographical variation in the usage of call types or site-specific calls (reviews in Catchpole & Slater 1995; Janik & Slater 1997), we have found no bird or mammal study that has demonstrated conclusively which mechanism was responsible for these differences. Note that, even in animals that are capable of production learning, adaptations of vocal behaviour to specific sites might nevertheless be genetic or arise through contextual learning.

Studies on variation in vocalizations of captive animals or those that have contact with humans are even more difficult. Differences in vocalizations between two holding facilities, for example, can be caused by caretakers. If animals interact with humans, involuntary conditioning effects can occur (Knight 1984; Adams et al. 1987). These need not stem from obvious differences in treatment but could be caused by factors such as preferences of a caretaker for certain behaviour patterns or simply from the experience the animals had with particular caretakers in the past. Such conditioning effects would indicate the occurrence of contextual but not of production learning. Furthermore, negative experiences could result in different internal states if certain humans are present and these could have profound effects on vocalizations. The human influence could be considered an environmental or a social factor. Social interactions between animals may have similar effects. A dominant individual in one group that prefers certain calls could also influence the call type distribution within a group and account for differences between groups.

Other phenomena that can arise through production learning but present no evidence for it are changes in calls during ontogeny and infant babbling. Changes during ontogeny occur in almost all animals by maturational processes and simple growth. It is helpful to describe such vocal phenomena in detail, and they deserve further investigation, but interpretations should be cautious if causes are not clear (see Jorgensen & French 1998 for a good example of a careful interpretation).

INNOVATION AND INVENTION

In many birds, copying from others is so obvious that researchers can recognize easily whether a song was learnt. However, new elements similar to, but not exactly the same as, a model vocalization often arise. Marler &

Peters (1982) have used the term 'invention' to refer to new signals that are not modifications of any of the model vocalizations and 'improvisation' to describe songs that are similar to those heard but sufficiently different to be classified as new types. We agree with the term 'invention' but prefer 'innovation' to describe signals that are modifications of model signals, since 'improvisation' in music refers to modifications that are not necessarily based on previous themes, are made up on the spot and do not last. This is very different from the lasting changes built into a bird's repertoire to which Marler & Peters apply the term.

Payne (1997) used 'improvisation' ('innovation' in our terminology) as a synonym for copying mistakes, thus suggesting that it has no adaptive significance. We think that the terms 'innovation' and 'invention' are more appropriately used for phenomena that have an adaptive value. For example, if an increase in the variability of song confers an increased reproductive success because females use variability or novelty in mate choice, innovation or invention could have evolved to arrive at such an increased variability. The term 'copying mistake' or 'error' on the other hand suggests that the outcome is either selectively neutral or disadvantageous. Both mechanisms could, of course, be responsible for an observed novel signal, and it will require careful experimenting to determine which one is at work.

How do innovation and invention fit in with our terminology? Marler & Peters (1982) described innovations as either elemental (new acoustic elements) or combinatorial (temporal rearrangements of existing acoustic elements). Thus, innovations can be achieved either by production or by contextual learning. Invention, on the other hand, can involve social learning but does not need to do so. Only if experience with vocalizations of other individuals was used to arrive at a signal that differs from theirs would production learning have been used to invent that signal. If such experience is of no relevance, invention does not involve social learning. Thus, two different mechanisms can lead to invented vocalizations: one with and one without production learning. Experiments in which young animals are exposed to vocalizations that range over certain frequency bandwidths or use certain modulation patterns could help to show whether they use production learning to invent novel signals, that is, develop signals that tend not to overlap in acoustic space with the ones they have heard. To demonstrate that invention without production learning is taking place, young animals would have to be isolated from auditory stimuli and still be found to develop individually specific vocalization patterns that cannot be explained by differences in the morphology of their vocal tracts. Invention is generally difficult to demonstrate since the researcher also has to exclude the possibility that seemingly new vocalizations are arrived at by drastic innovation. In birds, observation of the development of a new song element through subsong could help to identify cases of drastic innovation.

Even though innovation and invention are difficult to demonstrate and species that use them must have very specific constraints on sound production to ensure that

novel vocalizations are recognizable to conspecifics, there are some examples where it is clearly advantageous. Bottlenose dolphins seem to use them to develop novel signals for individual recognition (Tyack 1997). Kroodsma and his colleagues hypothesized that innovation and invention in grey catbirds, *Dumetella carolinensis* (Kroodsma et al. 1997) and North American sedge wrens, *Cistothorus platensis* (Kroodsma et al. 1999) are related to a nomadic lifestyle and leads to a lack of geographical variation in songs. This might facilitate communication between individuals that were reared at different locations. Further studies will need to investigate what selective pressures lead to the use of invention or innovation in the acquisition of signal repertoires.

CONCLUSIONS

Any attempt to determine whether and in what form learning is responsible for a change in vocalizations will encounter difficulties. For example, it could be that animals demonstrate different forms of learning at different times in their developments and therefore vary in the mechanisms through which they acquire new sounds. However, even if we are not capable of distinguishing between such phenomena at this stage, it is important to be aware of the diversity of possible mechanisms. We think that a clear theoretical framework can help to identify research needs and direct research efforts.

Most studies on social learning in vocal communication have been concerned with production learning in birds. However, even here some areas have received less attention than others. Because production learning is a fascinating subject and the ability to learn is often so obvious in birds, few studies have looked at genetic influences on song or call variation between individuals of the same species. In many species it would be interesting to know whether and, if so, how parental genes influence similarities between siblings in learning abilities and in the fine structure of their songs or calls. Similarly, little is known about contextual learning in birds. Duetting species, for example, can show an amazing coordination in their song patterns and we know little about how this is achieved. Finally, the topic of innovation and invention has great significance in relation to the role of variability and individuality of song or call patterns. We have argued (Janik & Slater 1997) that production learning may have evolved to facilitate individual or group recognition. In that case it would be used to arrive at signals that differ from specific others in the environment rather than producing accurate copies. To demonstrate this use of production learning we need to determine what mechanism is involved in the production of a novel signal. This is also very relevant to many mammalian species.

Topics in mammalian production learning that we feel have received rather little attention are the possibility of learned gradual parameter changes within call types as found in bats and the whole possibility of learning involving filter changes. Studies that use conditioning of animals can help to elucidate both topics. For example, to date, there is no evidence for learning of frequency

parameters in nonhuman primates (Janik & Slater 1997). It is difficult to draw conclusions from negative results, but it is noticeable how such learning has been clearly demonstrated in other groups with similar experiments and much less research effort. However, some data suggest that chimpanzees may have an ability to match frequency parameters of conspecifics (Mitani & Brandt 1994; Mitani & Gros-Louis 1998), but further studies on such gradual parameter changes are needed to verify this claim and to determine whether it is caused by an adjustment of the phonatory system or by changes in the air pressure provided by the respiratory system causing subtle changes in fundamental frequency (Hsiao et al. 1994). Furthermore, primates have formants in their calls that vary with context (Owren et al. 1997). Formants are a feature that encodes most of the information in human language, yet very little work has been done on filter learning in nonhuman primates, even though it has been shown that they can be conditioned to alter jaw movement and that this is based on a direct projection from the motor cortex (Luschei & Goodwin 1975).

To date, filter learning has been demonstrated only in humans and song sparrows (Nowicki et al. 1992). However, vocal tract filtering is common in other birds and mammals too (e.g. bats: Hatley & Suthers 1988) and whether it can be influenced by learning certainly deserves further investigation. The relationship between usage learning, comprehension learning and production learning of temporal and amplitude parameters is another interesting topic for further studies. Primatologists often point out similarities between human and nonhuman primate vocal development focusing on contextual learning (Seyfarth & Cheney 1997; Snowdon et al. 1997). However, many mammals (e.g. Salzinger & Waller 1962; Molliver 1963; Lilly 1965; Schusterman & Feinstein 1965; Burnstein & Wolff 1967; Lal 1967) and birds (Lane 1960; Todt 1975; Spector et al. 1989) are capable of both forms of contextual learning. As pointed out earlier, contextual learning as far as usage is concerned requires a certain level of control over the respiratory system and it is probably closely linked to respiratory production learning. The similarities that we find between non-human primates and humans may, in fact, be abilities that all mammal and bird species that use vocal signals in their communication have in common. We suspect that these forms of learning are relatively widespread in birds and mammals and can even perhaps be found in some other taxa. Therefore, we need a broader research approach with more studies concentrating on neglected groups and species.

Finally, animals can produce nonvocal acoustic signals with different parts of their bodies, for example, the lip smacks of primates and the drum signatures of kangaroo rats and woodpeckers. These signals can, of course, also be influenced by learning. Many arguments in our discussion of ways to demonstrate production learning also apply to these signals, and indeed to communication signals transmitted through other modalities. Furthermore, there can be complex interactions between different signal channels that we miss if we concentrate only on vocal signals. Lip smacking in primates has been

proposed as a precursor to language evolution (Andrew 1998). Since the main information in human language is encoded by the filter system of which the lips are a part, this is an interesting possibility. However, to our knowledge no study has tried to investigate the influence of learning on such nonvocal acoustic signals.

These are just some of the research questions that currently appear of special interest. We hope that this theoretical framework for studying the influence of learning in vocal communication will help to focus research on areas that need further exploration and to stimulate further discussions on the role of learning in animal communication.

Acknowledgments

We thank K. Riebel and P. Tyack for helpful comments on the manuscript. V.M.J. was funded by a BASF-Postdoktorandenstipendium of the German National Merit Foundation. This is contribution number 10086 from the Woods Hole Oceanographic Institution.

References

- Adams, R. A., Lengas, B. J. & Bekoff, M. 1987. Variations in the avoidance responses to humans by black-tailed prairie dogs (*Cynomys ludovicianus*). *Journal of Mammalogy*, **68**, 686–689.
- Adret, P. 1993. Vocal learning induced with operant techniques: an overview. *Netherlands Journal of Zoology*, **43**, 125–142.
- Andrew, R. J. 1998. Cyclicity in speech derived from call repetition rather than from intrinsic cyclicity of ingestion. *Behavioral and Brain Sciences*, **21**, 513–514.
- Backus, R. H. & Schevill, W. E. 1966. Physeter clicks. In: *Whales, Dolphins, and Porpoises* (Ed. by K. S. Norris), pp. 510–528. Berkeley: University of California Press.
- Baker, J. A. & Bailey, E. D. 1987. Sources of phenotypic variation in the separation call of northern bobwhite (*Colinus virginianus*). *Canadian Journal of Zoology*, **65**, 1010–1015.
- Barlow, G. W. 1977. Modal action patterns. In: *How Animals Communicate* (Ed. by T. A. Sebeok), pp. 98–134. Bloomington: Indiana University Press.
- Beecher, M. D. 1996. Birdsong learning in the laboratory and field. In: *Ecology and Evolution of Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 61–78. Ithaca: Comstock Publishing Associates.
- Box, H. O. & Gibson, K. R. 1999. *Mammalian Social Learning: Comparative and Ecological Perspectives*. Cambridge: Cambridge University Press.
- Burnstein, D. D. & Wolff, P. C. 1967. Vocal conditioning in the guinea pig. *Psychonomic Science*, **8**, 39–40.
- Catchpole, C. K. & Slater, P. J. B. 1995. *Bird Song: Biological Themes and Variations*. Cambridge: Cambridge University Press.
- Cynx, J. 1990. Experimental determination of a unit of song production in the zebra finch (*Taeniopygia guttata*). *Journal of Comparative Psychology*, **110**, 267–275.
- Elowson, A. M., Snowdon, C. T. & Lazaro-Perea, C. 1998. Infant 'babbling' in a nonhuman primate: complex vocal sequences with repeated call types. *Behaviour*, **135**, 643–664.
- Garcia, J. & Koelling, R. 1966. Relation of cue to consequence in avoidance learning. *Psychonomic Science*, **4**, 123–124.

- Groswasser, Z., Korn, C., Groswasser-Reider, J. & Solzi, P. 1988. Mutism associated with buccofacial apraxia and bihemispheric lesions. *Brain and Language*, **34**, 157–168.
- Guinee, L. N., Chu, K. & Dorsey, E. M. 1983. Changes over time in the songs of known individual humpback whales (*Megaptera novaeangliae*). In: *Communication and Behavior of Whales* (Ed. by R. Payne), pp. 59–80. Boulder: Westview Press.
- Hansen, P. 1979. Vocal learning: its role in adapting sound structures to long-distance propagation and a hypothesis on its evolution. *Animal Behaviour*, **27**, 1270–1271.
- Hatley, R. S. & Suthers, R. A. 1988. The acoustics of the vocal tract in the horseshoe bat, *Rhinolophus hildebrandti*. *Journal of the Acoustical Society of America*, **84**, 1201–1213.
- Hauser, M. D. 1988. How infant vervet monkeys learn to recognize startling alarm calls: the role of experience. *Behaviour*, **105**, 187–201.
- Hauser, M. D. 1989. Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalizations. *Journal of Comparative Psychology*, **103**, 149–158.
- Heyes, C. M. 1998. Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, **21**, 101–148.
- Heyes, C. M. & Galef, B. G. 1996. *Social Learning in Animals: the Roots of Culture*. San Diego: Academic Press.
- Hsiao, T. Y., Solomon, N. P., Luschei, E. S., Titzel, I. R., Liu, K., Fu, T. C. & Hsu, M. M. 1994. Effects of subglottic pressure on fundamental frequency of the canine larynx with active muscle tension. *Annals of Otolaryngology, Rhinology and Laryngology*, **103**, 817–821.
- Hultsch, H. & Todt, D. 1989. Memorization and reproduction of songs in nightingales (*Luscinia megarhynchos*): evidence for package formation. *Journal of Comparative Physiology A*, **165**, 197–203.
- Jakobson, R. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.
- Janik, V. M. & Slater, P. J. B. 1997. Vocal learning in mammals. *Advances in the Study of Behavior*, **26**, 59–99.
- Jones, G. & Ransome, R. D. 1993. Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proceedings of the Royal Society of London, Series B*, **252**, 125–128.
- Jorgensen, D. D. & French, J. A. 1998. Individuality but not stability in marmoset long calls. *Ethology*, **104**, 729–742.
- Jürgens, U. 1992. On the neurobiology of vocal communication. In: *Nonverbal Vocal Communication: Comparative and Developmental Approaches* (Ed. by H. Papoušek, U. Jürgens & M. Papoušek), pp. 31–42. Cambridge: Cambridge University Press.
- Knight, R. L. 1984. Responses of nesting ravens to people in areas of different human densities. *Condor*, **86**, 345–346.
- Kroodsma, D. E. & Canady, R. A. 1985. Differences in repertoire size, singing behavior, and associated neuroanatomy among marsh wren populations have a genetic basis. *Auk*, **102**, 439–446.
- Kroodsma, D. E., Houlihan, P. W., Fallon, P. A. & Wells, J. A. 1997. Song development by grey catbirds. *Animal Behaviour*, **54**, 457–464.
- Kroodsma, D. E., Liu, W.-C., Goodwin, E. & Bedell, P. A. 1999. The ecology of song improvisation as illustrated by North American sedge wrens. *Auk*, **116**, 373–386.
- Lal, H. 1967. Operant control of vocal responding in rats. *Psychonomic Science*, **8**, 35–36.
- Lane, H. L. 1960. Control of vocal responding in chickens. *Science*, **132**, 37–38.
- Le Boeuf, B. J. & Petrinovich, L. F. 1974. Dialects of northern elephant seals, *Mirounga angustirostris*: origin and reliability. *Animal Behaviour*, **22**, 565–663.
- Lieblich, A. K., Symmes, D., Newman, J. D. & Shapiro, M. 1980. Development of the isolation peep in laboratory-bred squirrel monkeys. *Animal Behaviour*, **28**, 1–9.
- Lilly, J. C. 1965. Vocal mimicry in *Tursiops*: ability to match numbers and durations of human vocal bursts. *Science*, **147**, 300–301.
- Luschei, E. S. & Goodwin, G. M. 1975. Role of monkey precentral cortex in control of voluntary jaw movements. *Journal of Neurophysiology*, **38**, 146–157.
- Manabe, K., Staddon, J. E. R. & Cleaveland, J. M. 1997. Control of vocal repertoire by reward in budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **111**, 50–62.
- Maples, E. G., Haraway, M. M. & Hutto, C. W. 1989. Development of coordinated singing in a newly formed siamang pair (*Hylobates syndactylus*). *Zoo Biology*, **8**, 367–378.
- Marler, P. & Nelson, D. 1992. Neuroselection and song learning in birds: species universals in a culturally transmitted behavior. *Seminars in the Neurosciences*, **4**, 415–423.
- Marler, P. R. & Peters, S. 1982. Subsong and plastic song: their role in the vocal learning process. In: *Acoustic Communication in Birds, Vol. 2: Song Learning and its Consequences* (Ed. by D. E. Kroodsma, E. H. Miller & H. Ouellet), pp. 25–50. New York: Academic Press.
- Marler, P. & Peters, S. 1987. A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*, a case of age-limited learning. *Ethology*, **76**, 89–100.
- Marler, P. & Pickert, R. 1984. Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Animal Behaviour*, **32**, 673–689.
- Marler, P. & Sherman, V. 1985. Innate differences in singing behaviour of sparrows reared in isolation from adult conspecific song. *Animal Behaviour*, **33**, 57–71.
- Marshall, A. J., Wrangham, R. W. & Arcadi, A. C. 1999. Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour*, **58**, 825–830.
- Medvin, M. B., Stoddard, P. K. & Beecher, M. D. 1992. Signals for parent-offspring recognition: strong sib-sib call similarity in cliff swallows but not barn swallows. *Ethology*, **90**, 17–28.
- Mitani, J. & Brandt, K. 1994. Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology*, **96**, 233–252.
- Mitani, J. C. & Gros-Louis, J. 1998. Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour*, **135**, 1041–1064.
- Molliver, M. E. 1963. Operant control of vocal behavior in the cat. *Journal of the Experimental Analysis of Behavior*, **6**, 197–202.
- Nelson, D. A. & Marler, P. 1994. Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences U.S.A.*, **91**, 10 498–10 501.
- Nelson, D. A., Whaling, C. & Marler, P. 1996. The capacity for song memorization varies in populations of the same species. *Animal Behaviour*, **52**, 379–387.
- Nowicki, S., Marler, P., Maynard, A. & Peters, S. 1992. Is the tonal quality of birdsong learned? Evidence from song sparrows. *Ethology*, **90**, 225–235.
- Nowicki, S., Peter, S., Searcy, W. A. & Clayton, N. 1999. The development of within-song type variation in song sparrows. *Animal Behaviour*, **57**, 1257–1264.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M. & Cheney, D. L. 1992. 'Food' calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. *Behaviour*, **120**, 218–231.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M. & Cheney, D. L. 1993. Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, **26**, 389–406.
- Owren, M. J., Seyfarth, R. M. & Cheney, D. L. 1997. The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cynocephalus ursinus*): implications for production processes and functions. *Journal of the Acoustical Society of America*, **101**, 2951–2963.

- Payne, K. & Payne, R. 1985. Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, **68**, 89–114.
- Payne, K., Tyack, P. & Payne, R. 1983. Progressive changes in the songs of humpback whales (*Megaptera novaeangliae*): a detailed analysis of two seasons in Hawaii. In: *Communication and Behavior of Whales* (Ed. by R. Payne), pp. 9–57. Boulder: Westview Press.
- Payne, R. B. 1997. Song traditions in indigo buntings: origin, improvisation, dispersal, and extinction in cultural evolution. In: *Ecology and Evolution of Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 198–220. Ithaca: Cornell University Press.
- Payne, R. S. & McVay, S. 1971. Songs of humpback whales. *Science*, **173**, 585–597.
- Pepperberg, I. M., Gardiner, L. I. & Luttrell, L. J. 1999. Limited contextual vocal learning in the grey parrot (*Psittacus erithacus*): the effect of interactive co-viewers on videotaped instruction. *Journal of Comparative Psychology*, **113**, 158–172.
- Podos, J. 1996. Motor constraints on vocal development in a songbird. *Animal Behaviour*, **51**, 1061–1070.
- Podos, J., Peters, S., Rudnicki, T., Marler, P. & Nowicki, S. 1992. The organization of song repertoires in song sparrows: themes and variations. *Ethology*, **90**, 89–106.
- Podos, J., Nowicki, S. & Peters, S. 1999. Permissiveness in the learning and development of song syntax in swamp sparrows. *Animal Behaviour*, **58**, 93–103.
- Ralls, K., Fiorelli, P. & Gish, S. 1985. Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, **63**, 1050–1056.
- Riebel, K. & Todt, D. 1997. Light flash stimulation alters the nightingale's singing style: implications for song control mechanisms. *Behaviour*, **134**, 789–808.
- Salzinger, K. & Waller, M. B. 1962. The operant control of vocalization in the dog. *Journal of the Experimental Analysis of Behavior*, **5**, 383–389.
- Schusterman, R. J. & Feinstein, S. H. 1965. Shaping and discriminative control of underwater click vocalizations in a California sea lion. *Science*, **150**, 1743–1744.
- Searcy, W. A., Nowicki, S. & Peters, S. 1999. Song types as fundamental units in vocal repertoires. *Animal Behaviour*, **58**, 37–44.
- Seyfarth, R. M. & Cheney, D. L. 1986. Vocal development in vervet monkeys. *Animal Behaviour*, **34**, 1640–1658.
- Seyfarth, R. M. & Cheney, D. L. 1997. Some general features of vocal development in nonhuman primates. In: *Social Influences on Vocal Development* (Ed. by C. T. Snowdon & M. Hausberger), pp. 249–273. Cambridge: Cambridge University Press.
- Slater, P. J. B. 1983. Sequences of song in chaffinches. *Animal Behaviour*, **31**, 272–281.
- Smith, W. J. 1977. *The Behavior of Communicating*. Cambridge, Massachusetts: Harvard University Press.
- Snowdon, C. T., Elowson, A. M. & Roush, R. S. 1997. Social influences on vocal development in New World monkeys. In: *Social Influences on Vocal Development* (Ed. by C. T. Snowdon & M. Hausberger), pp. 234–248. Cambridge: Cambridge University Press.
- Spector, D. A., McKim, L. K. & Kroodsma, D. E. 1989. Yellow warblers are able to learn songs and situations in which to use them. *Animal Behaviour*, **38**, 723–725.
- Sutton, D., Larson, C., Taylor, E. M. & Lindeman, R. C. 1973. Vocalization in rhesus monkeys: conditionability. *Brain Research*, **52**, 225–231.
- Todt, D. 1975. Social learning of vocal patterns and modes of their application in grey parrots (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie*, **39**, 178–188.
- Tyack, P. L. 1997. Development and social functions of signature whistles in bottlenose dolphins *Tursiops truncatus*. *Bioacoustics*, **8**, 21–46.
- Watkins, W. A. & Schevill, W. E. 1977. Sperm whale codas. *Journal of the Acoustical Society of America*, **62**, 1485–1490.
- Whiten, A. & Ham, R. 1992. On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. *Advances in the Study of Behavior*, **21**, 239–283.
- Wild, J. M. 1994. The auditory–vocal–respiratory axis in birds. *Brain, Behavior and Evolution*, **44**, 192–209.