Why songbirds learn songs: an arms race over ranging?

Eugene S. Morton
National Zoological Park, Washington DC 20008, USA

ABSTRACT

Learning songs, as practised by the oscine passerines (songbirds), has provided subject matter for neurophysiologists and molecular and field biologists. Here, I review attempts to explain the evolutionary origin of song learning so that why they learn songs can be integrated with, and made complementary to, advances in our understanding of the mechanisms of song learning at the neurophysiological level. Ranging theory, that posits a strong relation between song function and perception, demography and energetic constraints, provides the most inclusive answer. Ranging, where a listener compares an incoming song to those in its memory, thereby assessing the degree of sound degradation it contains, provides the source of selection to favor learning songs only when a strong and persistent singer/listener dichotomy in selfish interests exists. When this dichotomy circumvents the normal dominance accruing to individuals with prior residency, an evolutionary arms race between singers and listeners produces the selection necessary for song learning to evolve.

INTRODUCTION

Why do songbirds learn songs, why does song repertoire size vary between and within species, why do individuals in some populations share all song types while in other populations they do not, and why is female singing so rare in temperate regions while common in tropical ones? Researching these questions, and others, relies on the ease with which song can be recorded and reproduced, manipulated, and stored. When used in playbacks, songs are unparalleled as natural stimuli. As a consequence, bird song research embraces nearly all levels of biological inquiry, from molecular to whole organism to population. There is great potential for bird song research to fully integrate these levels and provide a holistic view of a complex behavior and its evolution. Here, I try to describe why, not how, birds learn songs. Why the system of learning songs for territorial use evolved might provide some additional questions for neurophysiological research or, at least, help engender the holistic view of the system that we desire.

A SHORT HISTORY OF EXPLANATIONS FOR SONG LEARNING

The common function of song, territorial defense, does not provide an obvious answer to why the learning of them should be favored by natural selection. Song learning is typical of many families of birds termed oscine passerines. Their close relatives, the rest of the perching bird order (Passeriformes) do not learn songs, even though many sing in defense of territories just like songbirds. It is doubtful that these nonoscine passerines learn songs because their songs vary little geographically, and attempts to find evidence for learning have failed (see Kroodsma and Baylis, 1982; Kroodsma, 1984, 1985, 1989). Kroodsma and Konishi (1991) showed that a deafened nonoscine (eastern phoebe, Sayornis phoebe) still learned normal songs and lacked cell clusters like those in the forebrain song nuclei of songbirds.

Past attempts to explain why birds learn songs have stimulated research but have not resulted in a widely accepted synthesis. They are also few. Marler (1960) argued that learning enhances speciation because songs can change rapidly, producing species isolation. Nottebohm (1972) added that co-adapted gene complexes might be conserved by assortative mating among birds sharing the same song dialect (a song dialect is a population-restricted song type or types, see Marler and Tamura, 1962) and that, in turn, dialects, might enhance speciation. Speciation, however, is abundant in nonoscine passerines, and dialects, while they are one result of the ability of oscines to learn songs, are not sufficiently widespread to be a cause of it. Furthermore, song learning does not appear to have contributed to the diversity of oscines by restricting gene flow (reviewed in Baptista and Trail, 1992). More recently, Nottebohm (1991) suggested that the dampening of sounds by the stapedius muscle of the inner ear, in order to protect the ear against
overly loud self-generated sounds, could lead to auditorily guided vocal flexibility. The work of Grassi et al. (1990) on crowing in roosters (Gallus gallus), however, suggested that the stapedius muscle does not protect the inner ear so much as it provides modulation of auditory feedback so that feedback during development occurs at a specific frequency and amplitude. Although his title suggests an analysis of the origin of vocal learning, Slater (1989) only speculates about current advantages that might maintain vocal learning. And, there is nothing new in the most recent book on bird song, (Catchpole and Slater, 1995).

Hansen (1979) suggested that learning provides a filter so that only songs that transmit best are sung. Perhaps, if we find out why songbirds need such a filtering mechanism more than other birds, we can answer the question of the origin of song learning?

THE DEVELOPMENT OF RANGING THEORY

A new perspective on the origin of song learning, called Ranging Theory (Morton, 1982, 1986, 1996; McGregor, 1994) was stimulated by the finding that a bird, hearing the song of a conspecific, adjusted its responses dependent upon the amount of degradation in the song (Richards, 1981). Richards played a tape of “near” and “far” songs on the territories of wild Carolina wrens (Thryothorus ludovicianus), 25m away from each bird. “Near” and “far” songs differed only in that “far” songs contained reverberations and other sources of signal degradation that accrue with passage through 50m of woods. The songs were equalized for amplitude. He obtained different responses. The birds sang after hearing the degraded songs, just as they respond to songs from neighbors a long distance off their territory. Instead of singing, they rushed to attack the speaker when they heard the undegraded “near” songs. Richards concluded that Carolina wrens used the degradation in the song to range the distance to its source and to respond accordingly.

Responses of wrens to hearing song was followed by a study of the singer’s “goals”, the other side of the coin. There are two possible interpretations of Richard’s work, one, that singers provide songs that degrade predictably to listeners, always providing good distance cues so that neighbors do not expend energy needlessly and respond only to real threats of intrusion (Richards, 1981). This might be predicted by a tit for tat game (e.g., Godard 1993). The second interpretation predicts that singers would not provide distance information to listeners except in situations where the singer benefits by doing so. In any situation, singers should use songs that degrade as little as possible so that they sound close to listeners, too close to ignore.

To test these predictions, Gish and Morton (1981) recorded Carolina wren songs from birds in deciduous forests in Maryland and 2,200 km south in subtropical palmetto hummocks in Florida. The song did not differ in frequency range or mean frequency so frequency-dependent attenuation could not contribute to degradation, regardless of the song’s native habitat. Following Richards, they broadcast 50 songs and re-recorded them through 50m of habitat in three sites, the same temperate deciduous forest and the same subtropical palmetto hummock the songs came from and a third site where none of the songs were native.

Gish and Morton compared the energy/time distribution in each undegraded song with its distribution after travelling through 50m of habitat. This was accomplished by converting the recorded song into an amplitude/time trace using a sound level recorder. Each trace consists of a series of peaks, each peak representing an element in a song. Gish and Morton (1981) compared the distribution of these peaks before and after the songs were re-recorded and converted to an index of how much change occurred.

The results of this study showed that Carolina wren songs are physically structured to retain their source characteristics during propagation. The set of Florida songs had a lower mean change index values in the Florida test sites (0.208) than in the deciduous forest site (0.220) while the deciduous forest songs had lower index values in their native site (0.198) than when tested in Florida (0.236). Change index values for the “neutral” sites showed no difference in degradation in the two song groups. Songs native to the test area degraded less than songs foreign to the area. Clearly, Carolina wrens songs are adapted to degrade as little as possible, substantiating Hansen’s (1979) idea that song learning could provide a means to adapt song to habitat acoustic conditions. Finally, in a laboratory study, we showed that song learning wrens use song degradation in their choice of songs to learn (Morton et al., 1986).

We have two sets of data on song, one on the responses to song showing that degradation is a factor (Richards, 1981) and one showing that the songs do not degrade predictably to the listener (Gish and Morton, 1981). This suggests that listeners might use general aspect of degradation to estimate distance. Further evidence proved this was not the case in Carolina wrens (Shy and Morton, 1986a) and other species tested (McGregor et al., 1983; McGregor and Krebs, 1984; Sorjonen, 1983; McGregor and Falls, 1984; Shy and Morton, 1986b; Morton and Young, 1986).
These playback studies showed that the perception of distance using degradation in signals is based upon whether or not the listener has the perceived signal in its memory. This feature is essentially to ranging, the ability to use degradation to estimate distance from singer (Morton, 1982, 1986). General features of degradation (e.g., reverberations, differential frequency attenuation, etc.) do not account for the field data reported. Most of the studies used playbacks of songs to show that degradation became an effective contributor to responses only if the focal individual had the song type it heard in its memory. Generally, the playbacks were done in the center of a focal bird’s territory, and the researchers compared responses to undegraded and artificially degraded songs played from the same position. Shy and Morton (1986a) used playbacks from within and without territories, thus providing natural degradation for the birds to assess. These studies have shown that the bird must have the signal in its memory before it is able to assess degradation and make responses appropriate to the apparent distance from which the signal originated.

THE ORIGIN OF SONG LEARNING AND RANGING

It is not surprising that birds can range song distance for it provides them with a means to avoid wasting energy. Indeed, ranging may be a general ability of all birds, not just songbirds (Morton, 1986). Perceivers should adjust their responses to singing to the threat the singer poses to the listener’s territorial or reproductive interests and they should not expend more time and energy than is necessary to defend their interests. One possibility, therefore, is that ranging evolved because it served this interest for perceivers: ignoring songs is important to their energetic balance when it allows uninterrupted foraging, mate guarding, etc., without incurring costs. Perceivers’ interests are served by continuing to do what is in their best interests after hearing a song. Ranging evolved because it served this function for perceivers, but such ranging did not favor the evolution of song learning.

What may be important to the origin of the oscine song learning system is the question of why selection favored singers to take countermeasures against perceiver ranging? Singers in small species of birds, and most passerines are small, have a low energy storage capability relative to their daily energy needs. Such species are severely constrained by energy balance and so they are particularly likely to develop an evolutionary arms race (sensu Dawkins and Krebs, 1979) over ranging with listeners.

Song learning enables singers to one-up listeners under certain circumstances that favor unrangeable songs. Territorial instability will generate selection that favors singers whose locations are not always rangeable, those with songs not in the memories of listeners. In such a situation, hearing conspecific songs not in their repertoires causes a bird to stop what it was doing and expend energy and time checking out the location of the singer. This disruption may function to enhance pre-existing dichotomies in territory energetics between the haves and the have-nots (Morton, 1986). Selection may be particularly intense where territories differ in food abundance and song output is food-limited.

The evolutionary response of listeners was to ratchet up their ability to learn songs so as to be able to range more song types. While this may have reduced the ability of singers to disrupt it also selected for larger numbers of song types in the singer's quiver. These repertoires of songs gave singers a means to disrupt listeners with unshared songs and shared songs for threatening neighboring territory occupants. My definition of repertoires includes both number of songtypes and the complexity of syllable structures that make up songtypes. Both these aspects of songs can vary, but, apparently, not independently due to constraints on memory size (Morton, 1982). A singer threatens a neighbor by matching the songtype the neighbor just sang, thereby allowing it to range and backoff. This function of matched countersinging harks back to the pre-arms race situation. Nothing new here that would underpin the origin of learning. Threat as a function, take alone, favors single songs, not repertoires, and would not underpin learning. More on this below. Again, it must be that disrupting neighbors and newcomers in unstable neighborhoods favored the arms race that began with singers.

Alternatively, repertoires and song learning might have increased attention to singing. Simple attendance to songs by listeners may benefit singers by lowering the likelihood that neighboring males will attempt extra-pair copulations with the singer’s mate, lowering territorial defense costs through tonic communication (Schleidt, 1973), attracting mates or females seeking extra-pair copulations, and by generally managing the behavior of perceivers in a manner beneficial to singers (Owings and Morton, in preparation; Morton and Page, 1992).

But I believe the evidence suggests that song learning evolved as a countermeasure to the ability of perceivers to range. Selection favoring song learning as a countermeasure is derived from the demographic situation where neighbors are turning over rapidly, when territorial boundaries are not known to newcomers and when, therefore, an individual must expend more energy to maintain them against newcomers than against old neighbors (e.g., Godard, 1991). Song learning for singers.
therefore, evolved as a countermeasure to increase the cost of song reception to perceivers by forcing them to respond or attend to unrangeable songs, those not shared, in geographic regions characterized by unstable neighborhoods. These regions are most likely to be those with a temperate climate, particularly in the northern hemisphere.

Where perceiver turnover is low, such as in dialect-singing oscines and tropical subspecies, singing is a reliable indicator of signaller quality (Zahavi, 1977, 1987), and little or no selection pressure on singers to use countermeasures to ranging exist. Indeed, selection has favored singers who use the ranging ability of listeners against them: they threaten by producing songs rangeable to everyone in the population (Morton, 1986). These songbirds use song in the same manner that nonlearning nonoscine passerines do. It is not accidental that dialect-songbirds share the same stable or tropical climates which host most of the nonlearning nonoscine species (Morton, 1986).

SOME PREVIOUS PREDICTIONS OF RANGING THEORY AND OUTCOMES

Songbirds that have dialects were predicted to respond most to their own dialect because song functions only to threaten via ranging, not to disrupt, for them. They can't range other-dialect songs but there has been no selection favoring their caring about these anyhow. They only care about defending boundaries, not about disrupting newcomers that don't know the local territories or enhancing differences in territorial quality. This prediction has been supported. Dialect species tend to ignore conspecific songs or calls dissimilar to their own (Baker et al., 1981; Petrinovich and Patterson, 1981; Payne, 1981). In contrast, individuals in species with non-dialect systems, repertoires or individualistic single songs, respond stronger to unmemorized or new conspecific songs (Morton, 1982, 1986; McGregor et al., 1983; McGregor and Krebs, 1984; Shy and Morton, 1986). These energetic responses are costs added by singers to perceivers, when both song learning and ranging are in effect.

The ranging hypothesis offers an explanation for many, often disparate, functions of bird song as well as the general trends already discussed. Playback experiments have supported several earlier predictions (Morton, 1986). For example, in addition to oscines, which are generally small species, evolutionary arms races should also produce song learning in other small-bodied, non-oscine taxa. This was confirmed for a hummingbird (Calypte anna) (Mirskey, 1976; Baptista and Schuchmann, 1990), which is energy limited (Stiles, 1971). The presence of song dialects in lekking hermit hummingbirds (e.g., Phaethornis longuemareus, P. superciliosus) suggests that song learning may be widespread in hummingbirds (Snow, 1968). Ranging has recently been shown in a non-oscine passerine, the dusky antbird (Cercomacra tyrannina), supporting the prediction that ranging evolved before song learning did in the passerines (Morton and Derrickson, in press).

Other predictions need testing with playback experiments and field research: Dialects should be favored in warm climates where the same individuals defend territories on a lifelong basis (stable neighborhoods), whereas birds living in unstable neighborhoods are not as likely to share songs. These songs, unshared with neighbors, should reduce their ability to range (e.g., Morton, 1982; Shy and Morton, 1986a). Species with single songs that differ from those of nearby territory holders should also be difficult to range (but see Morton and Young, 1986). An example is found in the white-crowned sparrow (Zonotrichia leucophrys), whose well-studied populations inhabiting the warm climate of California have dialects (Marler and Tamura, 1962) but whose migratory populations of temperate eastern North America have distinctive individualistic songs (Austen and Handford, 1991). Another prediction was that song output is maximized so as to advertise food availability, with males "overexerting" themselves when singing until a mate is acquired. Accordingly, females should prefer to mate with territorial males that spend the most time in song, when singing and foraging cannot occur together (in species with short term pair bonds formed only for breeding, not those with year-long pairbonds). This prediction received experimental support (Gottlander, 1987; Radesater et al., 1987; Reid, 1987; Alatalo et al., 1990; Arvidsson and Neergaard, 1991; Hoi Leitner et al., 1995). Furthermore, in the permanently paired Carolina wren, Strain and Mumme (1988) confirmed that song output is limited by food availability.

FUTURE DIRECTIONS

I began this essay with the suggestion that bird song research encompasses many levels of biological interest. Ranging theory provides many questions that can be answered through laboratory approaches. For example, we do not know how distance perception through ranging is accomplished neurophysiologically. The process may be similar to bat echolocation but, instead of sending out a known signal and using echo arrival times to assess distance, the bird assesses the amount of degradation (approximating its distance from the source) in another individual's incoming signal with its undegraded memorized version of the same signal. The bird might use motor inputs to the neural tissue, activating its memory of what its own song sounds like when produced.
Williams and Nottebohm (1985) suggest that syringeal (the syrinx is the sound producing organ in birds) hypoglossal motor neurons respond selectively to natural song elements heard by the zebra finch (*Poephila*). Thus listening or singing finches have similar neural circuitry and it should be possible for a listener to convert a song heard into the motor commands necessary to reproduce the same sounds. If this is the case for ranging, only stored songs will allow an assessment of degradation. Margoliash (1983) provides evidence supporting this notion. He found that the auditory response properties of units in telencephalic nucleus in white-crowned sparrows (*Zonotrichia leucophrys*) exhibited considerable selectivity for the individual's own song. Furthermore, he found that song-specific units in wild-caught birds showed intradialect selectivity. It thus seems reasonable that the songs a bird hears are compared to song(s) stored in a portion of the brain termed the HVc and that these autogenous (self) songs serve as the reference component in ranging (see also, Margoliash and Fortune, 1992; Margoliash et al., 1994). The Gambel's white-crowned sparrow (Z. l. gambelii), offers an opportunity to test the generality of autogenous song sensitivity because, in this population, dialects are not found and self songs are individualistic (Austen and Handford, 1991). I predict that response to autogenous songs are lessened here.

There is still controversy, however, about whether listeners can learn to range songs they hear from neighbors but do not sing themselves. McGregor (1992) felt that his study of great tit (*Parus major*) responses to song degradation and song “familiarity” (McGregor and Krebs, 1984) proved that neighbors could range each others songs even if they do not share them. They compared responses to three categories of songs: those found only in the test male (OWN), those found in both the test male and one of his neighbors (OWN + NEIGHBOR), and those found only in neighbors (NEIGHBORS). Only 2 of 8 measures showed significant heterogeneity and they concluded that their three categories of familiar song had little effect on degradation discrimination. They concluded that songs of neighbors can be ranged even if the test male does not sing them (McGregor, 1991). But, to see if birds respond similarly to songs that they share with their neighbors, and to songs sung by neighbors but not by themselves, one should compare only the categories OWN + NEIGHBORS and NEIGHBORS, not their responses to OWN songs too, which are obviously in the bird’s memory.

When this is done, one finds for undegraded songs that the birds respond significantly stronger to OWN + NEIGHBOR than to NEIGHBOR songs and that the discrimination between these two categories is very diminished in response to degraded songs (Eyal Shy, unpublished manuscript). In other words, birds can sing unarrangeble songs and disturb listeners that do not share that song type, even, as in the McGregor and Krebs study, between neighbors familiar with each other through vocal interactions over several months. Ranging theory predicts that singing unshared songs should cause more disruption before territorial boundaries have become well known and when climate and season provide the greatest energetic stress (Morton, 1982; 1986).

The great tit is probably not a good species in which to show disruption anyway, because individuals are able to learn songs throughout life, or, at least, after the age of two years (McGregor and Krebs, 1988). This study casts doubt on the main conclusion of an earlier study (McGregor and Avery, 1986), one cited as evidence against ranging. The McGregor/Avery study suggested that old males learn songs of new neighbors for discriminating between these new neighbors and strangers, but not for singing. That is, birds learned songs for discrimination but not for performing. Obviously, if this is generally true, then birds could range the songs of neighbors, even if they do not perform or sing, these songs. But the later study shows that they learn to sing new songs of new neighbors to threaten them through matched countersinging (McGregor and Krebs, 1988). This supports predictions of function from ranging theory, not learning theory developed by psychologists (performance versus learning for discrimination, Dickinson (1980) as referenced by McGregor and Avery (1986)).

I conclude by suggesting how field studies of song ranging might integrate with another level of inquiry into bird song. Because the perception of degradation involves precise time assessment, it is probably no coincidence that birds are superior to most mammals only in this area of peripheral hearing ability. Time interval assessment takes place peripherally, in the ear itself, rather than in the central nervous system. Budgerigars (*Melopsittacus undulatus*), for example, can resolve sounds separated by as little as 1–2 ms. Humans, by contrast lose sensitivity to sound happening faster than 5–6 ms apart (Dooling, 1982).

**ACKNOWLEDGMENTS**

I thank a legion of people who have worked with me on ranging and other acoustic studies, especially S. Gish, E. Shy, K. Derrickson, M. Van der Voort, and K. Young. Martin Moynihan and B. J. M. Stutchbury provided a critical review and other support. The research was supported by Scholarly Studies Grant 1235S109 from the Smithsonian Institution.
REFERENCES


Morton, E. S. 1986. Predictions from the ranging hypothesis
for the evolution of long distance signals in birds. 


