

*Museum of Zoology and Department of Biology, University of Michigan, Ann Arbor,
Michigan*

Randomization Analyses: Mimicry, Geographic Variation and Cultural Evolution of Song in Brood-Parasitic Straw-Tailed Whydahs, *Vidua fischeri*

Robert B. Payne, Jean L. Woods, Mark E. Siddall & Cynthia S. Parr

Payne, R. B., Woods, J. L., Siddall, M. E. & Parr, C. S. 2000: Randomization analyses: mimicry, geographic variation and cultural evolution of song in brood-parasitic straw-tailed whydahs, *Vidua fischeri*. *Ethology* **106**, 261–282.

Abstract

Although bird song has been an important model for investigating questions of behavior development, cultural evolution and population differentiation, the quantitative methods of analysis have been problematic. Here we develop and apply quantitative randomization methods to test hypotheses about these processes in a natural population of birds. Songs of the African brood-parasitic straw-tailed whydahs (*Vidua fischeri*) and songs of their host species, the purple grenadier (*Granatina ianthinogaster*), were compared in audiospectrograms for similarity to test the following hypotheses: Whydahs mimic the songs of their host species, they have local song dialects, neighboring males match their song themes, local males match the songs of local hosts, remote populations have different songs according to their geographic distance, and songs undergo cultural evolution over time across generations. Randomization analyses were completed using (1) Mantel matrix statistics and (2) tree-based measures employing Sankoff optimization of Manhattan matrices and approximate randomizations. Our results provide evidence for song mimicry, local song dialects, matching song themes between neighboring males, song matching of local whydah mimics and grenadier song models, correspondence of song differences and geographic distance, and cultural continuity with change in song traditions within a local population. These randomization methods may be useful in other studies of animal communication, and they are sufficiently general for use both with distance matrices derived either from naturalistic impressions of song similarity as in our example or from acoustic measurements.

Corresponding author: Robert B. Payne, Museum of Zoology, University of Michigan, Ann Arbor, MI 48109–1079, USA. E-mail: rbpayne@umich.edu

Introduction

Bird songs are of interest not only as a parallel for human language and cultural evolution, but also in their significance in determining species limits and evolutionary relationships, in their importance in social interactions, and as time markers of restricted periods of learning (Nicolai 1964, 1973; Marler 1970; Mundinger 1980; Payne 1981, 1985, 1996; Payne et al. 1981; Kroodsma & Baylis 1982; Kroodsma et al. 1985; Slater 1986; Beecher et al. 1996; Kroodsma & Miller 1996). In all these instances, it is critical to detect similarities and differences among songs. Because of the differences among species in song structure, no single set of acoustic criteria allows a naive investigator to identify a particular song or to describe all its detail, nor is there a reason to describe a particular song in full detail. Yet, an experienced naturalist has no difficulty determining songs of different species or one song from another within a species. And though acoustic parameters can be measured, the visual representations of the same bird songs may be at least as consistent in conveying the distinctiveness of the songs. Even though the corresponding song themes among individual birds may vary, the overall similarities and differences in songs may be apparent in a visual representation of a complex vocal signal. In some studies where visual classifications of audiospectrograms of songs were compared with results of multivariate analysis of acoustic parameters of these songs, the results agreed well (Clark et al. 1987; Nowicki & Nelson 1990; Helweg et al. 1998). In other studies, independent observers agreed between themselves in their classifications of audiospectrograms (Waser & Marler 1977; Payne et al. 1981; Baptista & Petrinovich 1986; Hughes et al. 1998). These studies encouraged us to explore the use of visual comparisons of songs in a songbird whose behavior is of both special and general interest to test the development, geographic variation and cultural evolution of song, and to develop analytical methods that would allow us to make inferences about the significance of the patterns observed.

Straw-tailed whydahs (*Vidua fischeri*) are brood-parasitic finches in East Africa (Zimmerman et al. 1996). Male whydahs mimic the songs of their host species, the purple grenadier (*Granatina ianthinogaster*), which rear the young whydahs along with their own young. But in spite of experiments on their song development (Nicolai 1973), it is not clear whether in natural conditions they copy the songs of their individual foster parents or the songs of older whydahs in the same area. However, in the related and closely studied village indigobirds (*V. chalybeata*), males match their mimicry songs with their neighbors, and these local song dialects correspond with the set of neighboring birds that interact socially over a few km². The males modify each of their songs from year to year, each male revising its songs in the same way as its neighbors, in a continuing process of cultural evolution (Payne 1985, 1990). In both the indigobirds and whydahs, the mimicry songs in males and the mating and species-specific parasitism preferences in females that are attracted to these mimicry songs are responsible for the origin and maintenance of the different species (Nicolai 1964; Payne et al. 1998, in press). Finally, because straw-tailed whydahs (together with shaft-tailed whydahs, *V.*

regia) are the sister group to the indigobirds (Klein & Payne 1998), it is of interest to determine their song development and variation in order to trace which patterns of song behavior and development are shared between species because of common ancestry, and which are adaptations to the lifestyle of each species.

We carried out a field study to determine the local and regional geographic variation in song mimicry in straw-tailed whydahs and the matching of songs among local neighboring male whydahs. We sampled local neighborhoods and remote populations to determine the geographic scale of song matching. We also tested whether songs change over time in a process of cultural evolution by comparing populations over space and by sampling songs across years. We used two analytical methods of randomization to address the following hypotheses: (1) Whydah songs mimic the songs of their host species, (2) whydahs within a region have songs more similar than whydahs in different regions, (3) whydah songs vary with geographic distance, (4) whydah songs are more similar to each other than they are to songs of their local grenadier hosts, and neighboring male whydahs match their songs with each other, and (5) whydah song continues as a cultural tradition across generations but undergoes divergence in a process of cultural evolution. We also used the results of these randomization methods to compare their application in studies of behavior.

Methods

Field Populations and Song Recordings

Fieldwork with straw-tailed whydahs and purple grenadiers was carried out by R. B. Payne and L. L. Payne for three weeks in 1988 in Kenya on the west side of Lake Baringo (0°38'N, 36°05'E). R. B. Payne had recorded songs of whydahs and grenadiers in 1976 in the same locality. In both years, the whydahs and grenadiers were observed and their songs recorded within 10 km of Kampi-ya-Samaki (=Kampi) from Loruk south to Chemeron. Songs of whydahs and grenadiers were also recorded at Kacheliba (1°30'N, 35°00'E) in western Kenya in 1976, and 14 km SW of Voi (at 3°26'S, 38°42'E) and Olorgesailie (1°34'S, 36°27'E) in southern Kenya in 1967 (Fig. 1). In addition, captive grenadiers, thought to have originated near Dodoma, Tanzania, were recorded. Within each study area, we recognized individual male whydahs by variations in the wear of their long tail feathers, we determined the nearest neighboring male whydahs, and we plotted the position of each bird on a field map. Distances between birds were measured from maps (Bartholomew & Son. 1993) and in the Lake Baringo region from 1 : 50 000 scale maps and on-site surveys. We recorded the song of each male whydah over a period of days or weeks during 46 h of focal-animal observation in 1988 and during shorter periods in the other years.

Songs were recorded using either a Uher 4000-L tape recorder and Uher M514 microphone, or a Marantz PMD221 cassette recorder and Sennheiser ME-80 microphone mounted in a Sony 330 mm parabolic reflector. The sample recorded at Lake Baringo in 1988 was 7138 songs from 15 male whydahs. The sample

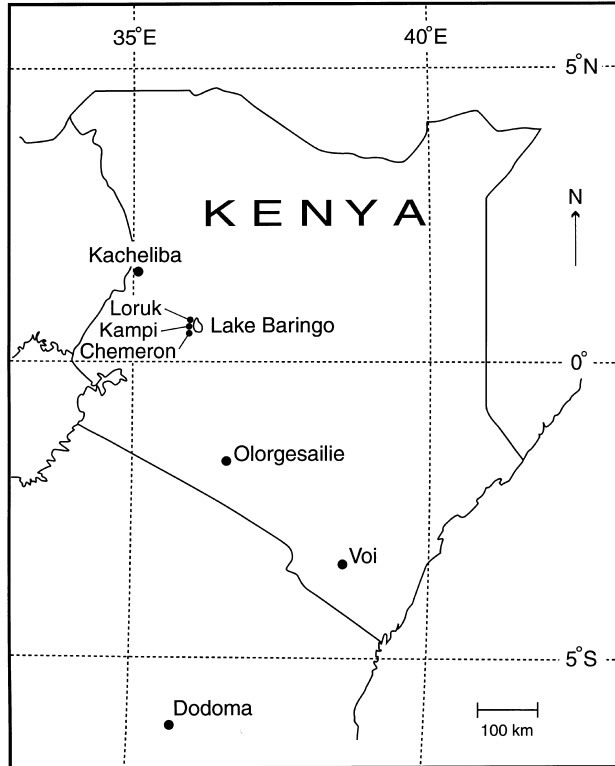


Fig. 1. Map of recording localities of straw-tailed whydahs and purple grenadiers

recorded in the earlier years was 940 songs from 18 male whydahs. All songs recorded were audiospectrographed. The sample available and used for visual examination was thus the entire sample that was tape recorded in all years.

For the 1988 Lake Baringo recordings, songs were first visually inspected with a Uniscan II (Multigon Industries, Mt. Vernon, NY) and then were documented on 35 mm film strips of all 1988 whydah songs, using a PAR-4512 real-time analyser (Princeton Applied Research, Princeton, NJ), Tektronix oscilloscopes (Tektronix, Beaverton, OR) and a Grass kymograph camera (Grass Instruments, West Warwick, RI; Hopkins et al. 1974; Payne 1985). The film strips were examined for matching songs and groups of phrases within the songs. Because we found that songs were often repeated in nearly identical structure within a bout of a few hundred songs for each bird, we used a match-to-sample method of determining song types and phrase types. We photocopied each successive song that differed from any previous one in the film strips. We arranged the photocopies in plastic envelopes and physically moved the array of photocopies along the film strips until we either found one that matched, or we added a copy to the array. A similar approach is often applied to avian vocalizations for species that have standardized

songs and song elements (e.g. Slater et al. 1980; Payne et al. 1981, 1998; McGregor & Krebs 1982; Marler & Pickert 1984; Payne 1985, 1996; Farabaugh et al. 1988; Lynch et al. 1989; Beecher et al. 1996; Hughes et al. 1998; Hansen 1999).

In the present study, spectrograms of representative songs first identified on the film strips were made with a Kay DSP-5500 Sona-graph (Kay Elemetrics, Pine Brook, NJ) for the 1988 songs. Kay spectrograms were also generated for all songs recorded in the earlier years. We saw no evidence that our view of whydah song repertoires based upon the 1988 data could not also be applied to these other years.

Selection of Song Samples

Straw-tailed whydahs include in their song repertoire the mimicry of the songs and calls of the host species as well as chatters and complex nonmimetic vocalizations that are unlike those of the host (Nicolai 1969, 1973). For male whydahs the songs we selected were the longest and most complex songs, with a variety of whistled notes interspersed with short soft notes, the 'loud song' or 'Rufstrophe' as described by Nicolai (1973); and for grenadiers they were the corresponding songs of males and the most complex songs of females. We focused on the mimicked loud songs because they are thought to be learned (Nicolai 1973) and so are appropriate for a test of hypotheses related to song learning. We also focused on these songs because they are complex in acoustic structure, and we reasoned that if neighboring birds have similar complex songs, this is more likely to be the result of song copying and learning than any similarity that would be indicated by simpler vocalizations. Finally, we regard these songs as most likely to correspond both to a functional class of songs used to attract females (and so to be functionally equivalent among the whydahs) and to a cultural lineage. In the same way, the corresponding or homologous songs can be traced among populations of other species of *Vidua* both by their functional context and by their structural detail (Payne 1979, 1985). We found that the mimicry loud songs of a male whydah varied with recombinations of elements, and the songs generally fell into three versions or themes. When more than one variation of a song theme appeared in the film strips, we used the modal variant and we avoided shorter versions of whydah songs where longer versions were available (whydah mimicry of loud song is often shorter than the loud song as given by grenadiers).

Spectrograms of mimicked loud songs were selected from birds for which we had the most complete and acoustically clean recordings for their spatial category and (grenadiers) their sex. The total number of songs to be compared ($n = 23$) was determined from our estimates of (1) the minimum number needed to compare each of the hypotheses of interest in at least one set of songs, and (2) the number of audiospectrograms that an observer could compare and score within two hours. Pairwise comparison of songs increases with the number of songs, in the form $n \text{ pairs} = [(n \text{ songs} - 1) \exp 2]/(2)$, or 242 song dyads for 23 songs. So, including an additional six songs would involve 392 song dyads, or 1.6 times the number of song dyads from the 23 songs. We found that an observer's enthusiasm in scoring songs decreased after about two hours.

We selected song samples to provide a diversity of geographic distances between birds and at least one example to compare for each hypothesis. This sample allowed comparison of songs within a bird, between neighbors, variation within a few km and variation across a large geographic region, the songs of the brood parasite and its host within localities and across localities, and the variation across years.

We sampled whydah songs as follows: (1) The two themes of mimicked loud song for the male with the largest sample of songs in the Lake Baringo area in 1988 and the most frequent mimicry loud song theme for his nearest neighbor were selected (both at Chemeron); the samples allowed a comparison of song diversity within a male and between neighboring males. (2) The three most diverse versions of mimicry loud songs from Loruk and from Kacheliba were selected, with each song from a different male; the samples allowed comparisons of songs of neighboring males. These songs were from three of the four males recorded at Kacheliba (excluding the male with the fewest songs recorded) and from three of the six males recorded at Loruk (including the three males at the north end of the sample area, and excluding a male that arrived late in the season and his two neighbors). (3) The three most different versions of mimicked loud song were selected from the male at Kampi in 1976 (for comparison with later year songs at Lake Baringo) and the male at Voi; the Lake Baringo and Voi males allowed comparison of songs within an individual, only one male was recorded in each of these sites. Whydahs were not seen at Kampi in 1988, where the habitat had degraded in the years since 1976. (4) Whydahs ($n = 7$) at a site 5 km south of Kampi at Lake Baringo were excluded in order to limit the song sample. We excluded this sample because it was located between the Chemeron site and the Kampi and Loruk sites, and the Kampi and Loruk sites were the same distance apart (5 km) as Kampi and 5 km south of Kampi, so the sample would not add to the range of geographic distances that we had sampled already. (5) The most frequent mimicry loud song theme at Olorgesailie (where only one male was recorded) was selected for comparison of loud songs with the other localities. For the grenadiers, we selected seven songs to represent the two areas in which we also had the local songs of the whydah, captive birds from another area to represent geographic variation over a broader scale, two variants of song from one bird (male from Chemeron; one was recorded at this site), and a song from two females (Chemeron and captive; no female song was recorded at Olorgesailie).

Song Comparisons, Similarity Scores and Distance Matrices

There are multivariate methods to describe variations on sets of acoustic measures (Clark et al. 1987; Gaunt et al. 1994; Khanna et al. 1997), and multivariate scores have been applied successfully to compare vocal behavior of birds with simple calls and songs with few kinds of song elements or notes (Payne 1978; Trainer 1983; Shackell et al. 1988; Gaunt et al. 1994; Bell et al. 1998). The reason that multivariate descriptions of acoustic measurements were not indicated as appropriate in our analysis is that the songs of these whydahs and grenadiers

were much more complex than the simple songs where the major features can be recognized and comparable units can be measured. For example, not all songs had whistles that corresponded in shape or in number or position within a song and thus the elements of song were incommensurable in the songs of whydahs and grenadiers. On the other hand, the overall features of pattern of complex song that are used in simplified verbal or onomatopoeic descriptions of these songs suggest an auditory pattern (for the grenadier, one such pattern of notes that vary in time and frequency is a high thin 'cheerer cheet tsee-tsee sur-chit' and another is 'chit-cheet-tsereea-eee tsit tsit' with a long whistle ('tsee' or 'tsereea') that glides upwards or down in pitch, Zimmerman et al. 1996).

We suspected that these complex auditory patterns as visualized in the audio-spectrograms can be recognized with agreement among observers and that the observers can agree in their assessment of similarity among these complex patterns. We reasoned that although different naturalists may focus upon different features of song, much as naturalists have independently attributed different words as mnemonics to describe the songs of birds, each naturalist can detect pattern features that are useful in comparing songs. While the method introduces its own biases, it avoids biasing the choice of acoustic parameters to be measured and compared. In our results, we found a high degree of correspondence between observers using this method of complex pattern recognition.

Others have validated the method of visual pattern recognition of songs. Visual classifications of audiospectrograms correspond closely with those of multivariate analysis of acoustic parameters. Clark et al. (1987) determined that a visual classification of song elements of swamp sparrows *Melospiza georgiana* (Marler & Pickert 1984) gave the same results as a computer-based method for sound comparison and averaging, and Nowicki & Nelson (1990) determined that visual classification of the simple song elements of black-capped chickadees (*Parus atricapillus*) agreed with results of a cross-correlation computer-based method and with a principal components analysis of acoustic measurements of the same song elements. Similarly, a visual classification of song elements of humpback whales (*Megaptera novaeangliae*) agreed well with the results of acoustic parameters in a resonance model of a self-organized neural network (Helweg et al. 1998). On the other hand, the generally available computer-based methods for pattern recognition are insensitive to time and frequency shifts of the elements, and are particularly insensitive to re-ordering of elements in complex bird songs (Khanna et al. 1997), and human visual pattern recognition performed better than the available computer-based methods in other studies (Janik 1999). Moreover, the choice of specific features to measure can affect the outcome of song studies, as Kroodsma et al. (1985) found for geographic variants of songs of white-crowned sparrows (*Zonotrichia leucophrys*), and Shackell et al. (1988) found for songs of neighboring American redstarts (*Setophaga ruticilla*).

For our study, a matrix of scores was determined by pairwise comparisons of the sample songs. DSP audiospectrograms for the 23 song exemplars were presented independently to six observers who were experienced as naturalists, were able to identify local birds to species by their songs, and had experience with

audiospectrograms of other songbirds (Fig. 2). Each was instructed to assess the similarity of these songs pairwise, rating each pair by the following ranking scheme: '1 = either very similar or identical, 2 = similar, 3 = somewhat similar, 4 = a little bit similar, or 5 = not similar.' The other instruction was to ignore song length. Each observer was informed that the sample included songs by grenadiers and song mimicry by whydahs, but not the proportion of grenadiers and whydahs, the time or place of recording, or whether a bird was represented by more than one song. Identity of birds was concealed by code numbers unrelated to locality and species, and the observers worked independently to score a set of similarity between songs. Observers varied in their knowledge of the hypotheses under test: Three were informed and three were naive.

We used these pairwise scores, as assessed by independent observers, rather than pairwise scores determined from a set of measured acoustic parameters. In principle, either method can generate dyadic comparisons that can be analysed in randomization methods. Because randomization tests do not depend on the method of construction of the score matrices, we expect the same results of randomization tests to follow in acoustic parameters. Because we scored song dyads that were more similar with the lower score and song dyads that were more different with the higher score, we refer to the score matrices as 'distance matrices' to facilitate verbal comparisons.

Randomization Methods

Quantitative analyses of songs were carried out with the statistical package SAS (System for Solaris, version 6.11, Cary, North Carolina, USA), PAUP* 4.0 (Swofford 1999), a Mantel matrix randomization program written by J. L. Woods, and contingency randomization programs written by M. E. Siddall.

The Mantel statistic is a measure of association between the ordered elements in two matrices (Manly 1997). We computed Mantel statistics to compare the scores of the observers. We also used the statistics to compare the probability that the data on song similarities were predicted to a significant degree by each of our specified hypotheses. In these tests, the hypothesis matrix and each observer's song distance matrix were used to calculate a Mantel statistic. The significance of the statistic was evaluated with a Monte Carlo randomization procedure, in which the rows and columns of one matrix were permuted randomly and the Mantel statistic was re-calculated. This permutation was repeated 5000 times to create a distribution of Mantel statistics. The one-tailed probability of the observed Mantel statistic was evaluated with this distribution and was the probability that the two original matrices were more similar than would be expected by chance. Iterated randomization avoids the problem of nonindependence of song comparisons: Within a single matrix each bird is represented many times, once in comparison with each other bird; so the correlation scores are not independent, and statistical inference in a single matrix is problematic (Schnell et al. 1985; Shackell et al. 1988; Gaunt et al. 1994).

To examine spatiotemporal variation in songs and other hypotheses, we used

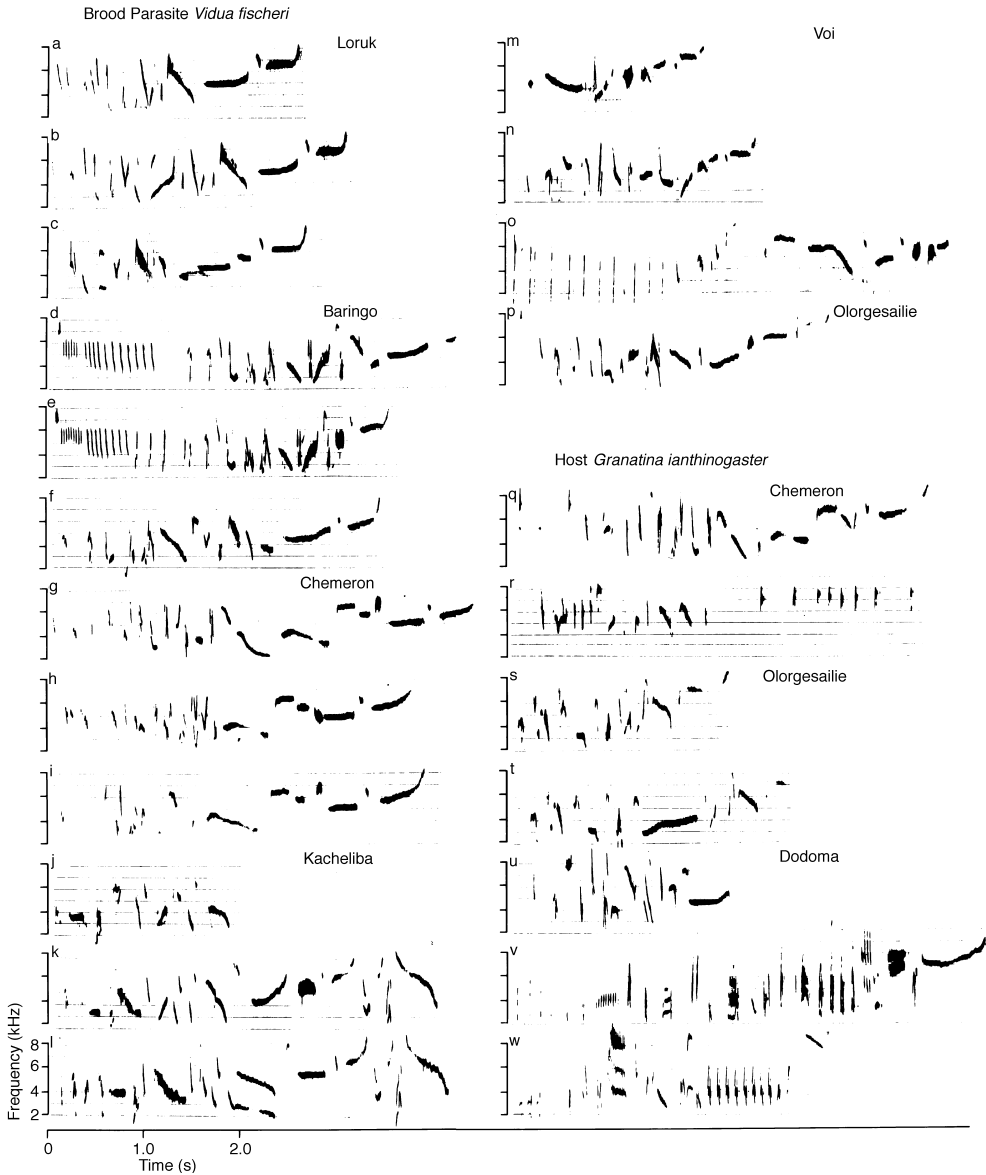


Fig. 2: Audiospectrograms of 23 songs of straw-tailed whydahs (a–p) and their host species the purple grenadier (q–w). whydahs, a–c, Loruk (three males in 1988); d–f, Baringo (three songs of one male in 1976); g, h, Chemeron (two songs of one male in 1988); i, Chemeron (a second male in 1988); j–l, Kacheliba (three males); m–o, Voi (three songs of one male); p, Ologesailie in 1967; grenadiers, q, Chemeron, male; r, Chemeron, female; s, t, Ologesailie (two songs of one male in 1976); u–w, captives ex Tanzania (Dodoma) (u, v; two males; w, female)

two randomization approaches. First, the Mantel method was used to compare the song distance matrices with each other (distance scores as determined by different observers). We then used this method to compare these song distance matrices with map distances of the song sites. In addition, we used the method to test whether songs of the host were more alike, and songs of whydahs were more alike, than songs of hosts were to songs of whydahs. To test this we used a hypothesis matrix (Schnell et al. 1985; Gaunt et al. 1994), where the host–host or whydah–whydah song pairs were given a score of 0 (similar) and host–whydah song pairs were given a score of 1 (different). This hypothesis matrix then was compared to each song distance matrix using the randomization procedure described above. In the same way, we tested whether songs differed between the 1976 and 1988 samples (1976–76 and 1988–88 song pairs were scored 0, and 1976–88 song pairs were scored 1). The similarity of host and parasite songs within an area was compared to the similarity among areas using a third hypothesis matrix, where song pairs from the same site had a score of 0 and song pairs from different sites had a score of 1. In these tests, we assumed that the mechanisms which generate the distributions in the assessed song differences, and the distributions in the song differences and their map distances, were independent (Manly 1997). For these comparisons we used either the full data set of 23 songs, or a restricted data set of 19 songs (excluding songs of female and captive male grenadiers), or other restricted data sets as defined later, according to the hypothesis of interest.

In a parallel analytic approach, Manhattan-distance-based tree measures were used to determine the optimal solution to relatedness of songs based on the information in song distance matrices from each observer. These solutions were determined in unconstrained tree searches in PAUP* (Swofford 1999) under a Sankoff optimization criterion (Sankoff & Rousseau 1975; Sankoff et al. 1976). The degree of nonrandomness of any agreement found was determined by comparing the optimal tree length (L_o) from the original data with the tree length found when observers' scores were randomized among birds (L_r). The frequency with which tree lengths as short or shorter than with the original data were found was the one-sided p-value of random association of scores. Modifications of this procedure were employed for testing the departure from randomness for specific hypotheses. The significance of differences between song model (grenadier) and mimic (whydah) was determined by searching for the optimal Sankoff tree under the constraint that model songs and mimic songs represent separate sets and comparing the score of that tree with optimal trees obtained from 100 randomly determined constraining sets equivalent in size to the observed sets (i.e. 7 and 16 songs each). An equivalent procedure was used for a data set restricted to songs in the Lake Baringo area with a constraint of temporal partition between 1976 and 1988 songs, followed by random constraining partitions of the same sizes (i.e. 3 and 6 songs each). The association of song and geographic distances also was reassessed in the Sankoff procedure by finding the shortest spanning tree for pairwise geographic separations among birds, searching for the ensemble fit of observers' song-distance matrices using this geographic tree as a constraint, and determining the significance by re-evaluation of this fit with birds randomly reallocated to geographic sites.

The generation of random associations of scores and cases was applied to the two sets of matrices to be compared for the hypotheses under test, one where the primary matrices were observed sets (Mantel randomizations) and the other where the primary matrices were constrained in a series of prescribed sets (Sankoff optimizations). Because these randomizations and the primary matrices are quite different mathematically, agreement in results would give us confidence in drawing conclusions about song behavior.

The difference in perspective between the Mantel and Sankoff approaches is that the Mantel approach considers each pairwise similarity (or dissimilarity) of song type to be independent of the next one, whereas the Sankoff approach sees these as interrelated. That is, in the Sankoff approach, the differences between songs A, B and C are not seen as independent, and part of the difference between A and B may also be part of the difference between A and C without fully explaining the difference between B and C. It is not, however, necessary to be able to specify exactly what part of these differences explain the others, only the degree of difference, and that is what the Sankoff method takes into account. Similarly with geography, where if we know the distance from point X to Y and from X to Z, then we then know something about Y to Z (i.e. that it is not greater than the sum of X to Y and X to Z). So the Sankoff approach, arguably, may carry more information content than the Mantel approach. However, the Mantel approach is more conservative. There is no way to combine the different observers' scores in the Mantel analyses. The matrices must be assessed pairwise (or averaged with a huge loss of power). But the Sankoff procedure allows all of the scores to be used together in a composite analysis, thus increasing the power of the question (and this implies a higher composite significance than any of the parts).

We tested each hypothesis against a null model using the randomization methods described above. We note that the hypotheses we test are complementary rather than mutually exclusive, and we credit more than one biological process for the observed patterns of song similarities.

Results

Agreement between Observers' Scores

Each pair of the six matrices (one for each observer), when compared pairwise with 5000 randomizations, was significantly associated (Mantel statistics, all $p < 0.01$, range $0.32 < r^2 < 0.49$). The results indicate a high degree of correspondence in song comparisons by the observers. The details in differences among observers reflect their criteria of scoring similarity among categories. Two observers used the information that the songs were conspecifics or mimics to approach a median score of 3, whereas others were more sceptical and tended to score songs as less similar than 3. For the two categories of most similar (1 and 2, 'very similar or identical' and 'similar') there was very good agreement, with each individual applying this score to 14–30 cells. Of the 253 total cells in the pairwise triangular matrix $[(23 \times 23) - 23] / 2$ cells, 60 cells were scored 1 or 2, with 1–6 observers so

scoring these cells (mean, 2.25 observers), and 9 of 60 cells were so scored by 5 or 6 observers. In all, 24% of the cells were scored as 'similar' or 'very similar or identical.' We repeated the procedure with the restricted set of 19 songs, excluding the female and captive male grenadiers of uncertain origin, to compare a more homogeneous set of songs. Within this set, we obtained consistent and parallel results ($p < 0.001$, $0.49 < r^2 < 0.64$ for all pairs of observers).

We also determined the minimum number of steps needed to describe the relationships among songs in a minimal tree using the Sankoff procedure. Using each observer as a character (a total of 6 characters for the 23 songs), we found 10 equally optimal minimal-length trees. Each tree had an observed length $L_o = 323$ steps and the trees had a consistency index $CI = 0.944$, which indicated a high degree of agreement among the observers. To test whether this agreement was different than one would expect by chance, the assignments were randomly reallocated to songs for all six observers over 100 iterations and the minimal lengths L_r then were computed for each iteration. The results $L_r = 359-372$ were all longer than the number of steps in the original tree, indicating a significant ($p < 0.01$) nonrandom agreement among observers.

The majority-rule consensus tree of those optimal solutions (Fig. 3) illustrates several features of song behavior in the whydahs. All observers agreed that certain

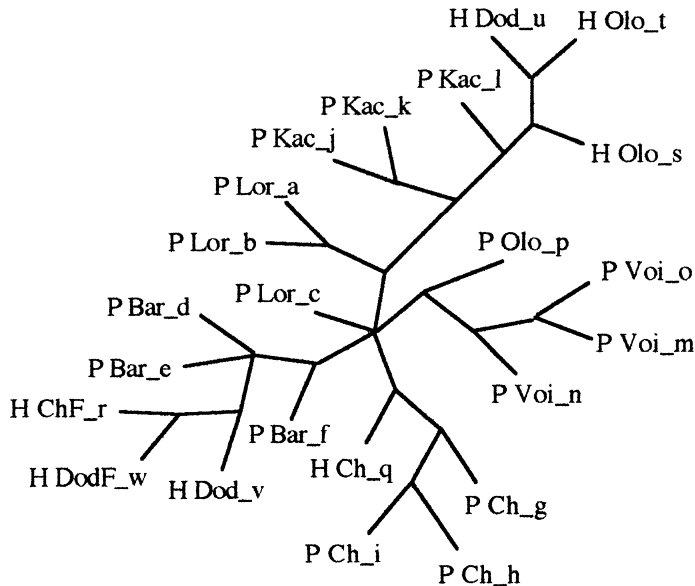


Fig. 3: Sankoff tree analysis of audiospectrograms of songs of brood-parasitic straw-tailed whydahs and their host species, the purple grenadier, as compared by six observers. H, host; P, parasite; Voi = Voi; Olo = Olorgesailie; Bar = Baringo (1976); Lor = Loruk (1988); Ch = Chemeron (1988); Kac = Kacheliba; F = female (others are male); lower-case letters indicate the songs in Fig. 2

songs were more like each other than any one was like another song, in the following sets:

- (1) PCh_g + PCh_i,
- (2) PCh_g, PCh_h + PCh_i,
- (3) PCh_g, PCh_h, PCh_i + HCh_q,
- (4) PVoi_m, PVoi_n + PVoi_o, and
- (5) PKac_j + PKac_k.

These sets indicate (1, 5 above) that neighboring male whydahs share songs with each other, whereas males in different geographic regions do not, (4) that songs of a male are more similar to other mimicry songs in his repertoire than to songs of males in different regions (2) that neighboring males have similar songs, and the similarity among males is greater than the similarity between some songs in the repertoire of an individual male, and (3) that local whydahs have songs that are more similar to their local host species, the grenadier, than they are to whydahs or grenadiers in other neighborhoods.

Finally, in set (PBar_d, PBar_e, PBar_f), most observers agreed that the songs of a male whydah (PBar) were more similar to each other than to other whydahs. All observers scored two of the songs (PBar_d and PBar_e) as more similar to the other than either was to any other songs. The third song was considered by all observers to be more like one or more songs of the Loruk birds, which was the neighborhood closest to the 1976 recording site at Kampi, and one likely to share a recent cultural ancestry with the Kampi neighborhood. Two observers scored PBar_f song as most similar to a grenadier song, tied for score (with PBar_f) with another whydah song. The few other pairs of songs scored high (score 1 or 2) by the observers indicate that the songs of the Baringo male are similar to other whydah songs, not only in the nearby Kampi and Chemeron areas but also at Kacheliba and Ologesailie over a broad geographic region.

Regional and Local Variation in Song

To determine whether songs vary among geographic regions, we used the Mantel statistic to compare the song difference matrices of each observer with the geographic map matrix. Each observer scored songs as more different when the songs were from a more remote site (for 23 songs, all $p < 0.01$, range $0.001 < p < 0.008$, $0.13 < r^2 < 0.24$; for the 19 songs excluding female and captive male grenadiers, all but one had $p < 0.01$, range $0.001-0.007$; $0.14 < r^2 < 0.31$; the other had $p = 0.20$, $r^2 = 0.08$). The probability that all six by chance would estimate a trend in the expected direction is 0.016 (binomial test), indicating that the increase in song difference with geographic distance was significant.

In the Sankoff procedure, we tested whether model grenadier and mimic whydah songs were more similar to each other within an area than were model and mimic in different areas. We determined the length of the tree that was

consistent with geographic distance among birds. We then permuted the geographic assignments of individual birds and their songs, and reassessed the best fit in the geographically constrained model (Fig. 4). In 100 random iterations of geographic area, the permutations for the restricted set of 19 songs gave tree lengths $L_r = 243$ – 258 steps, and all of these trees were longer than the original tree length $L_o = 242$ where distances were not randomized ($p < 0.01$). The results suggest that songs of host and parasite together are more different with increasing geographic distance, and that they are more similar within an area than among areas. This is consistent with the interpretation that whydahs mimic the local songs of their hosts.

When songs were restricted to those in the area at Lake Baringo, Mantel comparisons yielded mixed results. Excluding the song of female grenadier, four observers had a significant association of song difference and map distance ($n = 10$ songs, $p < 0.03$, $r^2 > 0.25$) and two had no close association ($p > 0.20$, $r^2 < 0.13$). In a Sankoff analysis, randomized songs gave tree lengths $L_r = 112$ – 137 steps, all of them longer than the original tree length $L_o = 102$ ($p < 0.01$), indicating a significantly higher similarity among songs within a local neighborhood than among neighborhoods. We attribute the difference in results to the greater power of the Sankoff test to differentiate between the null hypothesis and the hypothesis of a nonrandom negative association of song similarity with geographic distance.

The results show regional differences in song. Mimicry songs near Lake Baringo differed from mimicry songs in more remote populations. At Olorgesailie, songs ended with three whistles rising stepwise in pitch, each punctuated by a short

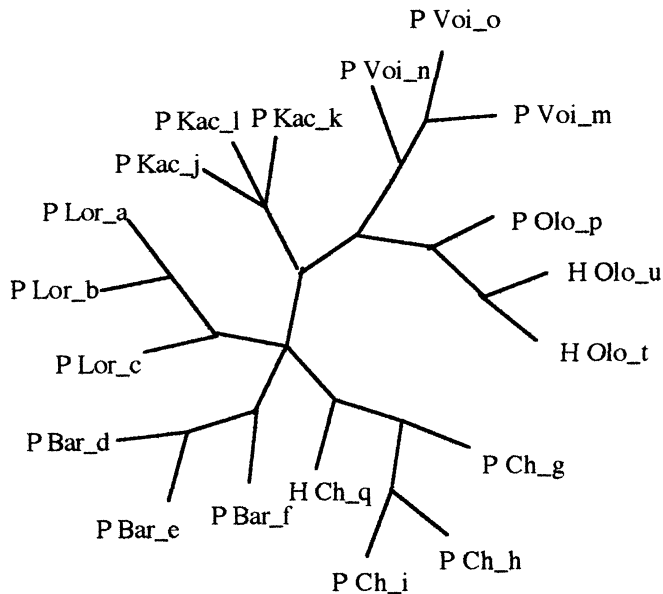


Fig. 4. Sankoff tree analysis of songs of whydahs and purple grenadiers, constrained to the geographic distances between singing birds. For legend to labels, see Fig. 3 caption

higher note. At Voi, songs were similar to those at Lake Baringo, but they differed in detail. At Kacheliba, songs ended with a descending note rather than an ascending note (Fig. 2). We conclude that the songs of whydahs and the songs of their host grenadiers vary over geographic distance, both on a local scale of neighboring song dialect populations and on a large scale of several hundred km.

Song Mimicry

The audiospectrograms (Fig. 2) support the interpretation of song mimicry (Nicolai 1964, 1973) by the whydahs of their host species, and we address the question of song mimicry further by comparing the song similarity and diversity within and between parasite and host, and the geographic variation in songs of parasite and host. The issue of song mimicry and its developmental origin includes two questions: (1) Do the parasites and hosts differ from each other in their songs, as expected if the whydahs copy their songs from other whydahs rather than copying the songs of their foster parents? (2) Are parasites and hosts more similar to each other within a local population than expected on the basis of song variation across a broad geographic region?

(1a) In the individual Mantel comparisons, the songs were not clearly more similar within a species than between species. For four observers there was no significant trend for the two species to differ in their songs, and for the other two the results were marginally significant ($n = 19$, $p = 0.04$, $r^2 = 0.006$).

(2a) A more sensitive test of whether the whydahs mimic the songs of their host is whether they are more similar within a population than between populations. To test for local mimicry, the songs of parasite whydahs and host grenadiers were compared within and between the two areas (Lake Baringo and Olorgesailie) where song samples were available for both species. In the individual Mantel comparisons, three observers scored the songs of parasites and hosts more similar within an area than between areas ($n = 13$, $p < 0.03$, $r^2 > 0.19$), and the trends observed were in the same direction in the other three tests ($r^2 = 0.08$, $p = 0.06$, 0.10 , 0.20).

(1b) To ascertain whether the songs of parasite and host differ in Sankoff optimization, we constrained a tree to a parasite-host split and found a tree length $L_o = 332$. In 100 randomized iterations we found tree lengths $L_r = 333$ – 377 ($p < 0.01$). In contrast to the results of the Mantel comparisons, the whydahs had songs more similar to conspecific whydahs than to their host species.

(2b) To test for the occurrence of parasite mimicry of the local songs of their host, we constrained a tree to include the songs of birds at Lake Baringo (PBar, PLor, HCh, PCh) to be in one set, and the birds at Olorgesailie (HOlo, POlo) to another set. In the Sankoff optimization we found a minimum $L_o = 173$. This was less than in 100 randomizations based on the same number of songs for each set with tree lengths $L_r = 173$ – 201 ($p = 0.04$). As in the Mantel test, we conclude that songs of parasite and host vary together from region to region.

The Mantel and Sankoff tests gave the same results in showing a significantly greater similarity between local songs of host and parasite than expected from the

variation in song recorded over a broader geographic region. Only the Sankoff test showed a significant difference between host and parasite, suggesting that the songs of the whydahs were derived from their host species and then were modified by the whydahs in a series of accumulations of improvisations on the host theme in a process of cultural evolution. The different result may reflect the greater power of the Sankoff test.

Two main findings appear. First, whydah songs mimic the songs of their local host grenadiers. Within areas where songs of both whydah and grenadier were recorded, mimicry songs of whydahs were more similar to the local songs of their host grenadiers than to whydahs or grenadiers in another population (Figs 2, 3 and 4). Second, mimicry songs vary within a male whydah (Fig. 2d–f, g–h, m–o). A mimicry song of a male and his neighbors sometimes was more like mimicry song in another song neighborhood than like the male's other mimicry songs. The variation among mimicry songs of a male whydah and his neighbors may parallel the song variation among the grenadiers, though we do not know the amount of song variation among grenadiers.

Cultural Evolution: Song Change over Time

To test whether songs differed across years, we compared observed map distances among mimicry songs at Lake Baringo (Kampi in 1976, Loruk and Chemeron in 1988) with Mantel matrix randomizations for these mimicry songs. We found a marginally significant association between sample year and song distance for three observers ($n = 8$, $p = 0.05$, $0.29 < r^2 < 0.33$) and a nonsignificant association in the same direction for the other three observers ($p > 0.1$, $r^2 < 0.24$). A binomial test comparing the chance that all observers would find the same trend ($p = 0.02$) indicates that this trend for change across years is significant.

In a parallel Sankoff test, we determined the number of steps required to describe the minimal tree for whydah songs in the Lake Baringo area when compared between different years. We found a minimum $L_o = 81$. This was marginally lower than those found in 100 permutations ($p = 0.05$), but was significantly lower than them after 1000 permutations ($p = 0.034$).

We conclude that songs within a year are more alike than are songs 12 yr apart in time in the same area. Because these birds at sites within 10 km of each other at Lake Baringo are likely to be members of a single interdispersing population, the songs appear to have changed more in time than they differ in space on a scale of 10 km. We do not know the date of divergence of these songs into distinct dialects in the whydahs. Continuous real-time field data are necessary to calibrate the rate of song divergence. In a field study of indigobirds (Payne 1985), a population on one side of a ridge had split into two different song neighborhoods within 2 yr of the time the birds colonized the valley on the other side of the ridge. In the same study, song themes that continued from year to year changed in the detail and naturalist observers were not able to score them as the same over a period of 5 yr unless information was available on the intermediate songs in the intervening years.

Discussion

Song Mimicry and its Development

The songs of a *Vidua* brood parasite and its host species are similar (Nicolai 1964), yet the songs of a male do not match precisely the songs of his foster parents, either in the field or in experiments where the time scale resembles that of foster parental care in natural conditions. Because each singing male *Vidua* has only one set of foster parents, but has more than one loud mimicry song theme, the variations among males could be due to either improvisation, learning error, or song sharing.

In the best studied host–parasite association to date, the host red-billed firefinches (*Lagonosticta senegala*) are parasitized and their songs mimicked by village indigobirds (Payne 1973, 1985, 1990). Each firefinch has one song theme and is individualistic. In contrast, each indigobird has a repertoire of firefinch songs (3 or 4) and all males in a local neighborhood match the same set of firefinch songs. These male indigobirds copy their songs from other male indigobirds, as opposed to their recalling and singing the songs of the individual foster parents that reared them. In experimental studies, the males copy their songs from other male indigobirds (Payne 1985; Payne et al. 1998). When males are forced to live with their foster father for much longer than the time when they would become independent and disperse in nature, they copy the song of their foster father; but when they are separated from their foster father at a normal age, they do not sing the song of their foster father. Instead, they sing the mimicry songs of other adult indigobirds that they hear (Payne et al. 1998).

In our field study, male straw-tailed whydahs have the same songs as neighboring male whydahs, rather than each male having its own songs as expected if it copied its own individualistic foster grenadier. The variations in song might be due to differences in the songs of their individual foster parents, but in our captive birds each male grenadier varies less in his song than the songs vary in the repertoire of a single male whydah. Neither improvisation nor learning error is likely to lead to the same songs being shared by different males. The mimicry loud songs of male whydahs were more alike than songs of the whydahs were like the songs of grenadiers in the same area (Fig. 2; randomization results). That is, the observed pattern of neighboring male whydahs that matched mimicry songs with each other at Lake Baringo suggests that neighboring males learn from each other.

In an experimental study, Nicolai (1973) reported that male straw-tailed whydahs developed the individualistic songs of their own foster parents. However, in that study the young whydahs were kept together with their foster parents for an unnaturally long period of time (12 months), well after the age when the young whydahs would disperse in the field, about 2 wks after fledging as it is in the indigobirds (Payne et al. 1998). On the other hand, Nicolai found that yearling whydahs learned mimicry song from older adult *Vidua* males: Although no adult male whydahs were present in the experiment, the yearlings copied the mimicry songs of the normal host of the indigobirds *V. chalybeata* from the indigobird mimics of firefinches, after the whydahs were more than 12 months old and first heard the indigobird songs. Nicolai's observations on song development in captive

straw-tailed whydahs are consistent with our idea, derived both from the experimental work with indigobirds and our field observations, that under natural conditions of independence and dispersal the whydahs do not simply copy the songs of their individual foster parents, rather the whydahs normally learn their mimicry songs from other adult *Vidua*.

Scale and Rate of Cultural Evolution

The time scale of change in whydah song could include both innovations by a male from year to year and also accumulations of copy errors in song learning across generations. No direct evidence of change within a male is available as the birds were not marked and observed from year to year. However, demographic comparison suggests that generations of birds were involved, insofar as the life expectancy of small finches in Africa is low (0.5 annual survival, Morel 1973; Woodall 1975; Payne & Payne 1977), and the 12-year period at Baringo (1976–88) was sufficient time for changes to accumulate in a process of cultural evolution.

The geographic scale of variation in bird songs ranges from broad continental regions in species such as in pine grosbeaks *Pinicola enucleator* (Adkisson 1981) and red crossbills *Loxia curvirostra* (Groth 1993), to populations of 20–200 birds in village indigobirds *Vidua chalybeata* (Payne 1985) and white-crowned sparrows *Zonotrichia leucophrys nuttalli* (Baptista 1975), or local neighborhoods of 2–6 birds in indigo buntings *Passerina cyanea* (Payne 1996) and song sparrows *Melospiza melodia* (Beecher et al. 1996). In some cases the macrogeographic scale suggests different species, whereas in others the microgeographic scale points more towards individual birds styling their songs after a certain neighbor. The ecological and social significance of such variation in song themes varies with historical, environmental, dispersal and social factors (Payne 1981, 1996; Kroodsma & Baylis 1982; Adret-Hausberger 1988). In *Vidua* finches, the indigobird males that live within a few square km and visit each other share their song repertoires. Also, females visit the males with the same set of songs and choose one of these males as a mate (Payne 1985; Payne & Payne 1997). The microgeographic scale of song variation in the straw-tailed whydahs over a few km (the distance between Loruk, Kampi and Chemeron) is about the same scale as that in the indigobirds, where birds within an area of about 10 km² shared their song types with each other, and birds only a few km beyond had another set of song types (Payne 1985). The social organization of these two species may be similar as well.

The change in songs of straw-tailed whydahs over time in East Africa is similar to that observed in more detail in village indigobirds in Zambia, where songs undergo a continuing process of change both within an individual male and between the generations of males that culturally transmit a song tradition within a population (Payne 1985). The change in songs in the whydahs is also consistent with a process of demographic replacement through extinction and immigration or dispersal between populations.

Rates of cultural evolution within a song lineage may be rapid, as in village indigobirds where all males modify each of their 24 song themes every year (Payne

1985). Continuity within a tradition sometimes has been assumed when songs were compared from one year to a much later year (Slater 1986). However, song themes may undergo local extinctions when birds disappear before their songs are copied, and when the old songs are replaced by immigrants that introduce new songs. Estimates of rates of cultural change may be exaggerated if local extinctions and the immigration by birds and replacement by songs from other populations is superimposed on a more gradual accumulation of change within a song tradition (Payne 1985, 1996). As straw-tailed whydahs were seen at Lake Baringo in the intervening years (Stevenson 1981), it is likely that cultural change has occurred within a continuing song lineage.

Randomization Models and Bird Songs

As expected from the earlier general comparison of the Mantel and Sankoff approaches, the Sankoff optimization showed statistically significant differences in some comparisons where the Mantel test did not. We interpret this difference in results as the difference in the power of the two models. First, we note that the hypothesis matrices in the Mantel tests were constrained to binary codes, whereas the Sankoff test was able to retain the graded difference between the elements, and this gradation contributed to the power difference of the two tests. Second, the Sankoff procedure allows the scores of all observers to be used together in an analysis, and this increases the power of the question (hence the higher composite significance than any of the parts in our Results). The drawback of this approach is that even if the composite shows a strong relationship, it does so at the expense of knowing whether any observers disagreed with each other (if this is of interest). If the question of interest is whether observers agree on an issue, then the Mantel approach is appropriate, and if one is interested in whether the collective information of observers speaks to some empirical pattern or process (in spite of any disagreements between observers), then the Sankoff approach is the more appropriate. Because our goal was to test biological processes, we consider the Sankoff approach to be the better method.

In another application of randomization methods, Shackell et al. (1988) measured a set of acoustic features to compare the single song elements of neighboring birds, used principal components analysis to describe song variation within a population, and used randomization tests to compare the correlation in observed PCA scores of song elements between neighboring birds with the correlations in randomized matrices. They found that some song elements were more similar between neighbors than expected by chance.

The randomization methods applied to the whydah songs were useful in confirming that different observers recognize the audiospectrogram patterns and assess the same similarities between complex songs. Randomization methods also are useful where the songs are compared with measured acoustic parameters to compare the similarity among neighboring birds (Gaunt et al. 1994), and the methods should be just as useful for complex sets of acoustic parameters. More importantly, the methods are useful in testing multiple biological hypotheses about social, spatial and temporal variations in bird song.

Acknowledgements

The National Science Foundation supported field studies in 1967 and 1976 and the National Geographic Society supported field studies in 1988. Helen R. Woods and the Map Room of the U.S. Library of Congress, Washington, D.C., provided maps. For assessment and scores of song audiospectrograms we thank Janet Hinshaw, Alec R. Lindsay and Laura L. Payne. For comments on the manuscript we thank Salvatore Cerchio, Alec R. Lindsay, Laura L. Payne, and the reviewers and editor.

Literature Cited

- Adkisson, C. S. 1981: Geographic variation in vocalizations and evolution of North American pine grosbeaks. *Condor* **83**, 277–288.
- Adret-Hausberger, M. 1988: Species specificity and dialects in starlings' whistles. In: *Acta XIX Congressus Internationalis Ornithologici*. National Museum of Natural Sciences, Ottawa, pp. 1585–1597.
- Baptista, L. F. 1975. Song dialects and demes in sedentary populations of the white-crowned sparrow (*Zonotrichia leucophrys nuttalli*). *Univ. Calif. Publ. Zool.* **105**, 1–52.
- Baptista, L. F. & Petrinovich, L. 1986: Song development in the white-crowned sparrow: Social factors and sex differences. *Anim. Behav.* **35**, 1359–1371.
- Bartholomew, J. & Son. 1993: *The Times Atlas of the World*. Times Books, London.
- Beecher, M. D., Stoddard, P. K., Campbell, S. E. & Horning, C. L. 1996: Repertoire matching between neighbouring song sparrows. *Anim. Behav.* **51**, 917–923.
- Bell, D. A., Trail, P. W. & Baptista, L. F. 1998: Song learning and vocal tradition in Nuttall's white-crowned sparrows. *Anim. Behav.* **55**, 939–956.
- Clark, C. W., Marler, P. & Beeman, K. 1987: Quantitative analysis of animal vocal phonology: An application to swamp sparrow song. *Ethology* **76**, 101–115.
- Farabaugh, S. M., Brown, E. D. & Veltman, C. J. 1988: Song sharing in a group-living songbird, the Australian magpie. Part II. Vocal sharing between territorial neighbors, within and between geographic regions, and between sexes. *Behaviour* **104**, 105–125.
- Gaunt, S. L. L., Baptista, L. F., Sánchez, J. E. & Hernandez, D. 1994: Song learning as evidenced from song sharing in two hummingbird species (*Colibri coruscans* and *C. thalassinus*). *Auk* **111**, 87–103.
- Groth, J. G. 1993: Evolutionary differentiation in morphology, vocalizations, and allozymes among nomadic sibling species in the North American red crossbill (*Loxia Curvirostra*) Complex. *Univ. Calif. Publ. Zool.* **127**, 1–143.
- Hansen, P. 1999: Long-term stability of song elements in the yellowhammer *Emberiza citrinella*. *Bioacoustics* **9**, 281–295.
- Helweg, D. A., Cato, D. H., Jenkins, P. F., Garrigue, C. & McCauley, R. D. 1998: Geographic variation in South Pacific humpback whale songs. *Behaviour* **135**, 1–27.
- Hopkins, C. D., Rosetto, M. & Lutjen, A. 1974: A continuous sound spectrum analyzer for animal sounds. *Z. Tierpsychol.* **34**, 313–320.
- Hughes, M., Nowicki, S., Searcy, W. A. & Peters, S. 1998: Song-type sharing in song sparrows: Implications for repertoire function and song learning. *Behav. Ecol. Sociobiol.* **42**, 437–446.
- Janik, V. M., 1999: Pitfalls in the categorization of behaviour: A comparison of dolphin whistle classification methods. *Anim. Behav.* **57**, 133–143.
- Khanna, H., Gaunt, S. L. L. & McCallum, D. A. 1997: Digital spectrographic cross-correlation: tests of sensitivity. *Bioacoustics* **7**, 209–234.
- Klein, N. K. & Payne, R. B. 1998: Evolutionary associations of brood parasitic finches (*Vidua*) and their host species: Analyses of mitochondrial restriction sites. *Evolution* **52**, 299–315.
- Kroodtsma, D. E. & Baylis, J. R. 1982: Appendix: A world survey of evidence for vocal learning in birds. In: *Acoustic Communication in Birds*, Vol. 2. (Kroodtsma, D. E. & Miller, E. H., eds). Academic Press, New York, pp. 311–337.
- Kroodtsma, D. E. & Miller, E. H., eds., 1996: *Ecology and Evolution of Acoustic Communication in Birds*. Cornell Univ. Press, Ithaca.

- Kroodtsma, D. E., Baker, M. C., Baptista, L. F. & Petrinovich, L. 1985: Vocal 'dialects' in Nuttall's white-crowned sparrow. *Current Ornithol.* **2**, 103—133.
- Lynch, A., Plunkett, G. M., Baker, A. J. & Jenkins, P. F. 1989: A model of cultural evolution of chaffinch song derived with the meme concept. *Am. Nat.* **133**, 634—653.
- Manly, B. F. J. 1997: *Randomization, Bootstrap and Monte Carlo Methods in Biology*. 2nd edn. Chapman & Hall, London.
- Marler, P. 1970: Birdsong and speech development: Could there be parallels? *Amer. Sci.* **58**, 669—673.
- Marler, P. & Pickert, R. 1984: Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Anim. Behav.* **32**, 673—689.
- McGregor, P. K. & Krebs, J. R. 1982: Song types in a population of great tits (*Parus major*): Their distribution, abundance and acquisition by individuals. *Behaviour* **79**, 126—152.
- Morel, M.-Y. 1973: Contribution à l'étude dynamique de la population de *Lagonosticta senegala* L. (Estrildides) à Richard-Toll (Sénégal). *Mémoires du Muséum National d'Histoire Naturelle, Serie A., Zoologie* **78**, 1—156.
- Mundinger, P. C. 1980: Animal cultures and a general theory of cultural evolution. *Ethol. Sociobiol.* **1**, 183—223.
- Nicolai, J. 1964: Der Brutparasitismus der Viduinae als ethologisches Problem. *Z. Tierpsychol.* **21**, 129—204.
- Nicolai, J. 1969: Beobachtungen an Paradieswitwen (*Steganura paradisaea* L., *Steganura obtusa* Chapin) und der Strohwitwe (*Tetraena fischeri* Reichenow) in Ostafrika. *J. Ornithol.* **110**, 421—447.
- Nicolai, J. 1973: Das Lernprogramm in der Gesangsausbildung der Strohwitwe *Tetraena fischeri* Reichenow. *Z. Tierpsychol.* **32**, 113—138.
- Nowicki, S. & Nelson, D. A. 1990: Defining natural categories in acoustic signals: Comparison of three methods applied to 'chick-a-dee' call notes. *Ethology* **86**, 89—101.
- Payne, R. B. 1973: Behavior, mimetic songs and song dialects, and relationships of the parasitic indigobirds (*Vidua*) of Africa. *Ornithol. Monogr.* **11**, 1—333.
- Payne, R. B. 1978: Microgeographic variation in songs of splendid sunbirds *Nectarinia coccinigaster*: Population phenetics, habitats, and song dialects. *Behaviour* **65**, 282—308.
- Payne, R. B. 1979: Song structure, behaviour, and sequence of song types in a population of village indigobirds, *Vidua chalybeata*. *Anim. Behav.* **27**, 997—1013.
- Payne, R. B. 1981: Population structure and social behavior: models for testing the ecological significance of song dialects in birds. In: *Natural Selection and Social Behavior*. (Alexander, R. D. & Tinkle, D. W., eds). Chiron Press, New York, pp. 108—120.
- Payne, R. B. 1985: Behavioral continuity and change in local song populations of village indigobirds *Vidua chalybeata*. *Z. Tierpsychol.* **70**, 1—44.
- Payne, R. B. 1990: Song mimicry by the village indigobird (*Vidua chalybeata*) of the red-billed firefinch (*Lagonosticta senegala*). *Vogelwarte* **35**, 321—328.
- Payne, R. B. 1996: Song traditions in indigo buntings: origin, improvisation, dispersal, and extinction in cultural evolution. In: *Ecology and Evolution of Acoustic Communication in Birds*. (Kroodtsma, D. E. & Miller, E. H., eds). Cornell Univ. Press, Ithaca, pp. 198—220.
- Payne, R. B. & Payne, K. 1977: Social organization and mating success in local song populations of village indigobirds, *Vidua chalybeata*. *Z. Tierpsychol.* **45**, 113—173.
- Payne, R. B. & Payne, L. L. 1997: Field observations, experimental design, and the time and place of learning in bird songs. In: *Social Influences on Vocal Development* (Snowdon, C. & Hausberger, M., eds). Cambridge Univ. Press, Cambridge, pp. 57—84.
- Payne, R. B., Payne, L. L. & Woods, J. L. 1998: Song learning in brood parasitic indigobirds *Vidua chalybeata*: Song mimicry of their host species. *Anim. Behav.* **55**, 1537—1553.
- Payne, R. B., Payne, L. L., Woods, J. L. & Sorenson, M. D., in press: Imprinting and the origin of parasite-host species associations in brood parasitic indigobirds *Vidua chalybeata*. *Anim. Behav.*, in press.
- Payne, R. B., Thompson, W. L., Fiala, K. L. & Sweany, L. L. 1981: Local song traditions in indigo buntings: cultural transmission of behavior patterns across generations. *Behaviour* **77**, 199—221.
- Sankoff, D. & Rousseau, P. 1975: Locating the vertices of a Steiner tree in arbitrary space. *Math. Prog.* **9**, 240—246.
- Sankoff, D., Cedergren, R. R. & Lapalme, G. 1976: Frequency of insertion-deletion, transversion, and transition in the evolution of 5S ribosomal RNA. *J. Mol. Evol.* **7**, 133—149.

- Schnell, G. D., Watt, D. J. & Douglas, M. E. 1985: Statistical comparison of proximity matrices: Applications in animal behaviour. *Anim. Behav.* **33**, 239—253.
- Shackell, N. L., Lemon, R. E. & Roff, D. 1988: Song similarity between neighboring American redstarts (*Setophaga ruticilla*): A statistical analysis. *Auk* **105**, 609—615.
- Slater, P. J. B. 1986: The cultural transmission of bird song. *Trends. Ecol. Evol.* **1**, 94—97.
- Slater, P. J. B., Ince, S. A. & Colgan, P. W. 1980: Chaffinch song types: Their frequencies in the population and distribution between repertoires of different individuals. *Behaviour* **75**, 207—218.
- Stevenson, T. 1981: *The Birds of Lake Baringo*. Sealpoint, Nairobi.
- Swofford, D. L. 1999: PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Trainer, J. M. 1983: Changes in song dialect distributions and microgeographic variation in song of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*). *Auk* **100**, 568—582.
- Waser, M. S. & Marler, P. 1977: Song learning in canaries. *J. Comp. Physiol. Psychol.* **91**, 1—7.
- Woodall, P. F. 1975: On the life-history of the bronze mannikin. *Ostrich* **46**, 55—86.
- Zimmerman, D. A., Turner, D. A. & Pearson, D. J. 1996: *Birds of Kenya and Northern Tanzania*. Christopher Helm, London.

Received: January 13, 1999

Initial acceptance: May 9, 1999

Final acceptance: September 4, 1999 (J.-G. Godin)