

2011

# Assessing the Phylogenetic and Cultural Content of Learned Song

Nicole Creanza

Follow this and additional works at: [http://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations](http://digitalcommons.rockefeller.edu/student_theses_and_dissertations)



Part of the [Life Sciences Commons](#)

---

## Recommended Citation

Creanza, Nicole, "Assessing the Phylogenetic and Cultural Content of Learned Song" (2011). *Student Theses and Dissertations*. Paper 146.



Assessing the Phylogenetic and Cultural Content of Learned Song

A Thesis Presented to the Faculty of  
The Rockefeller University  
in Partial Fulfillment of the Requirements for  
the degree of Doctor of Philosophy

by  
Nicole Creanza  
June 2011



## Assessing the Phylogenetic and Cultural Content of Learned Song

Nicole Creanza, Ph.D.

The Rockefeller University 2011

In the oscine songbirds, song is learned by a juvenile from a tutor of the same species, in a pattern that is analogous to human language, and likewise has the potential to change over time by cultural evolution. The similarities between human languages have been studied for centuries, but historically the relationships between the songs of birds of different species have been seen as too divergent to be useful. Using a computational analysis of song databases coupled with genetic phylogenies, I have shown that there is indeed a significant correlation between genetic distance and song similarity in the oscines. For a subset of Emberizid species, it was possible to reconstruct a genus-level phylogeny using the song syntax of these birds. For one member of this subset, I traced the cultural evolution of song properties, both over time and across distances. In addition, I examined the role of innate predispositions in song learning by hybridizing two species of Estrildid finches and examining the vocal output of hybrids in comparison with normally raised and cross-fostered birds. This work represents the first large-scale evolutionary analysis of learned song and the first clear demonstration of the relationship between genetic relatedness and song similarity.

## Acknowledgements

I am deeply indebted to my Faculty Advisory Committee members, Jim Hudspeth, for supporting and shaping my work, and Joel Cohen, for teaching me the benefits of hand-to-hand combat with the data. To Fernando Nottebohm, thank you for your willingness to chat over tea and for pushing me to dig deeply into my ideas. To Ofer Tchernichovski, thank you for taking the time to teach me about sound analysis and for serving as my external reviewer. Thanks to the Rockefeller University Dean's office. I am grateful to the Field Research Center staff, past and present, for making my work there seem like a vacation. Special thanks to Sharon Sepe for her finch matchmaking skills and constant devotion to the birds. Many thanks to the Nottebohm lab members, especially to Wan-chun Liu for chipping sparrow data and discussions and to Clare Walton for guidance. To Desmond Fugar, for devoting his summers to song analysis. To Jason Schwarz, for long-term collaborations. To the Cornell Lab of Ornithology, especially Jack Bradbury and Tammy Bishop, for donating many hours of recordings. To Joel Cracraft, for helpful conversations and references. To Christine Queitsch and David Haig, for fostering my love of evolutionary biology. To my family and friends, always.

## Table of Contents

|   |     |
|---|-----|
| Acknowledgments.....                                    | iii |
| Table of Contents.....                                  | iv  |
| List of Figures.....                                    | v   |
| List of Tables.....                                     | vi  |
| 1. Introduction.....                                    | 1   |
| 2. Song analysis of Emberizid sparrows.....             | 12  |
| 3. Genetic phylogeny of Passeriformes.....              | 38  |
| 4. Interfamily song analysis.....                       | 48  |
| 5. Cultural evolution of song in chipping sparrows..... | 61  |
| 6. The effects of hybridization on learned song.....    | 87  |
| 7. Discussion.....                                      | 110 |
| References.....   | 115 |

## List of : ictures

|  |     |
|--|-----|
| 1-1. An illustration of a species phylogeny.....                                     | 10  |
| 1-2. An early historical linguistic tree.....  | 11  |
| 2-1. Song spectrogram analysis.....  | 27  |
| 2-2. Calculating spectrogram overlap.....  | 28  |
| 2-3. DNA Barcode phylogeny of the Emberizidae.....                                   | 29  |
| 2-4. Phylogeny of the analyzed subset of Emberizidae.....                            | 31  |
| 2-5. Emberizidae phylogeny with song illustrations.....                              | 32  |
| 2-6. Neighbor-joining tree of syllable syntax distances.....                         | 33  |
| 2-7. Neighbor-joining tree of normalized syllable property distances.....            | 34  |
| 2-8. Correlation between genetic distance and song distance.....                     | 35  |
| 2-9. Neighbor-joining tree of factor analysis results.....                           | 36  |
| 3-1. Phylogenetic tree topology of birds included in the song analysis.....          | 47  |
| 4-1. Species selected for song analysis.....   | 54  |
| 4-2. Genera included in the song analysis.....                                       | 56  |
| 4-3. Song phonology tree for the phylogenetic song analysis.....                     | 57  |
| 4-4. Song syntax tree for the phylogenetic song analysis.....                        | 58  |
| 4-5. Neighbor-joining tree reconstructed from song factors.....                      | 59  |
| 5-1. Distribution of chipping sparrow syllable duration by longitude.....            | 78  |
| 5-2. Distribution of chipping sparrow number of syllables by longitude.....          | 79  |
| 5-3. Distribution of chipping sparrow data by longitude and latitude.....            | 80  |
| 5-4. Heat map of chipping sparrow syllable durations by latitude and longitude.....  | 81  |
| 5-5. Heat map of chipping sparrow number of syllables by latitude and longitude..... | 82  |
| 5-6. Cultural analysis of chipping sparrow syllables by longitude.....               | 83  |
| 5-7. Cultural analysis of chipping sparrow syllables by year.....                    | 84  |
| 6-1. Mouth markings and begging calls of zebra finch, silverbill, and hybrids.....   | 103 |
| 6-2. Plumage of zebra finch, silverbill, and hybrids.....                            | 104 |
| 6-3. One-on-one tutoring in zebra finches, silverbills, and hybrids.....             | 105 |
| 6-4. Neighbor-joining tree of average syllable properties.....                       | 106 |
| 6-5. Neighbor-joining tree of individual syllable properties.....                    | 107 |
| 6-6. Neighbor-joining tree of song syntax.....                                       | 108 |
| 6-7. Group tutoring in zebra finches, silverbills and hybrids.....                   | 109 |

## List of Tables

|  |    |
|--|----|
| 2-1. Grouping of Emberizid syllable properties by factor analysis.....   | 37 |
| 4-1. Grouping of the family syllable properties by factor analysis.....  | 60 |
| 5-1. Significant associations between syllables, date, and location..... | 85 |



## Part 1. Introduction

### *Evolution of language*

A simple line drawing, found in Charles Darwin's Notebook B, is one of the most famous images in biology (Figure 1-1). This sketch is the first species phylogeny, drawn as a branching tree to represent Darwin's idea that current species are the gradual result of descent with modification from ancestral species. Thus, two organisms on the tree share a common ancestor that can be traced back to where their two branches meet. This great intellectual leap, however, built upon prior work by historical linguists such as August Schleicher, who pioneered the classification of languages by relatedness, representing them as a genealogical tree (Figure 1-2). His family-tree theory (*Stammbaumtheorie*) of linguistic change over time was published in 1853, six years before *On the Origin of Species*, but other *Stammbaum* approaches arose in linguistics as early as 1808 (Atkinson and Gray 2004). Insisting that credit be given where it is rightfully due, Schleicher quickly reacted to reading *On the Origin of Species* in 1863, publishing an open letter citing the links between Darwin's new theory and existing linguistic concepts (1863).

In *The Descent of Man and Selection in Relation to Sex* (1871), Darwin makes explicit a connection between changes in languages and evolution in organisms that might logically be extended to bird songs, recognizing that both phenomena could be products of his descent with modification theory:

The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel.... Languages, like organic beings, can be classed in groups under groups; and they can be classed either naturally according to descent, or artificially by other characters. (91-92)

As seen in Figure 1-2, for more than 150 years researchers in historical linguistics have studied mutually unintelligible languages and constructed a phylogeny based on the similarities between words and grammatical structures. By comparing modern languages in light of these phylogenetic relationships, remarkable predictions have been made about the vocabulary and grammar of the common ancestor to the Indo-European languages. These predictions are not just an academic curiosity: reconstructing proto-Indo-European words has shed light on the lifestyle of its ancient speakers and on the way languages change over the course of history. Given the similarity of words for farm tools, domesticated plants and animals, and local flora in the extant languages, the speakers of this proto-language were putatively members of a farming culture living in the steppes north of the Black Sea about five to seven millennia ago (Renfrew 1999). The naturalist Gerhard Thielcke found this line of thought to be worthwhile in birds, wondering “[h]ow did the common ancestor of the twin treecreeper species (*Certhia brachydactyla* and *Certhia familiaris*) sing when they were still one species?” and visually comparing songs from various species in the genus *Parus* (tits) to find similarity in the song structure (1976, pp. 142-146).

Modern studies of linguistic change have identified four levels at which modification can take place: words can be added or replaced, the pronunciation of words can shift, the meaning of words can change, and the grammar can be altered (Cavalli-Sforza and Feldman 1981). While human language and the song of oscine birds are both vocalizations that are culturally transmitted predominantly during a sensitive learning period in juveniles, the parallels between birdsong and language are not complete. Birdsong syllables are analogous to words in the language evolution metaphor: changes in the

phonology (sounds) of the syllables are similar to shifts in pronunciation, and alterations in the syntax (pattern) of the syllables are comparable to grammatical changes in word order. However, there is no evidence in the literature that the syllables in a bird's song have symbolic meaning as words do, so the other aspects of linguistic change, the addition of a word with a new meaning and the shift in word definitions, are not readily applicable. In addition, the syllable repertoire sizes of songbirds probably peak around 2,000, whereas the number of words in the English language is at least two orders of magnitude higher. The comparative study of birdsong might be less complex than that of language in some ways, but the lack of symbolic meaning might make it difficult to trace the ancestry of particular syllables. Additionally, the radiation of the oscine songbirds began in the Eocene epoch, 34 to 56 million years ago (Barker *et al.* 2004), so the songs of these birds have had five to ten thousand times longer to accumulate changes than the proto-Indo-European languages. In the work that follows, I attempted to determine whether it is possible on a large scale to gather evolutionary information from learned song, to correlate song diversity with genetic distance, and to reconstruct phylogenetic relationships from behavioral data.

### *Evolution of vocal learning*

Avian communication can be divided into two categories: songs, which function intraspecifically in territory defense and mate attraction, and shorter vocalizations termed calls, which tend to carry information about the bird and its environment (predator, food, alarm, contact) and can serve interspecifically in some cases. Birds can use aspects of their songs to differentiate themselves from other species (Boughey and Thompson 1976) and to

identify themselves as fit individuals within their species. Humans can also discriminate bird species using song, both for casual identification and for systematic classification. Song can aid in taxonomy when discrepancies exist in species boundaries, such as a species with different plumage morphs that react (with aggression in males and copulation solicitation in females) to playbacks of the same song (Lanyon 1969, Thielcke 1976, Payne 1986) or, as in the case of the Willow and Alder Flycatchers, separate species with remarkably similar plumage but different innate songs (Kroodsma 1984).

The study of vocal learning in oscine songbirds is nearing the half-century mark. The early observations that male chaffinches can learn a song by imitating adult males (Thorpe 1958) were followed by the discovery that learned song was a robust behavioral output of a neural circuit that could be experimentally manipulated (Nottebohm et al. 1976). In general, these birds require both a song model and auditory feedback to develop normal vocalizations; isolates and deafened birds produce aberrant song (Konishi 1965; Price 1979). In the most closely related group of non-learners, the suboscines, deafened Eastern phoebes can develop normal song, indicating that the behavioral program is innate (Kroodsma and Konishi 1991).

Most studies of birdsong focus on the subtle differences between the songs of a single species. However, interspecific song comparisons may uncover a wealth of evolutionary information. The songs of oscines are learned by a new generation of birds each breeding season, and these songs have been shown in many species to be an indicator of male success. Selection favors those birds with the song best fit to the environment and to the preferences of females. Learned song is transmitted culturally, much like human

language, but the sexual selection pressures on song could be stabilizing over time or might lead to rapid behavioral changes in a population.

Vocal learning has been linked to an adaptive radiation in the songbird lineage. Evolving relatively recently (60-65 million years ago, Barker et al. 2004), the oscine clade contains roughly 4,500 of the 10,000 bird species. Presumably, vocal learning could facilitate speciation if song could undergo sexual selection. For this model to work, some aspect of song must be a reliable indicator of male fitness so that choosing a mate on the basis of his song would have a genetic benefit for the female. For example, many species of vocal learners have regional song dialects; a female might be able to determine by a male's dialect that he was raised locally or migrated there in his first breeding season, which could potentially inform her that the male was well suited to his environment. If this kind of behavior happened consistently, then a preference for the local dialect could lead to song-based reproductive isolation and eventually speciation. However, not all dialect variations are new species in the making – some have correlated to genetic differences (in this case, between subspecies) (Kroodsma and Canady 1985) and some have not (Nottebohm and Selander 1972).

Song dialects could potentially correspond to adaptation to certain environments, but several researchers have proposed a more general link between a bird's song and its natural history. Certain song properties correlated with habitat characteristics in small groups of species, particularly linking the density of the vegetation in a bird's habitat (or, alternatively, the other species of birds sharing the habitat) to acoustic song features such as average frequency and trill rate (Chappuis 1971, Morton 1975, Nottebohm 1975,

Wasserman 1979, Handford 1981). With a computational approach, I can test these predictions across more species than was previously possible and determine whether they are universal or limited to certain locations, clades, or song types.

#### *Systematic information in birdsong*

Until molecular phylogenies became feasible for large-scale studies, evolutionary relationships between species were most often inferred by examining morphological characters and determining a parsimonious arrangement. Occasionally, scientists also probed the evolutionary history of behavioral features within a group of species. For example, Konrad Lorenz hypothesized that the innate breeding behaviors of the Anatid ducks should change over time in a similar way to physiological characteristics, and he successfully correlated these behaviors to the accepted species phylogeny of the time (Lorenz 1941). Debates on the merits of this kind of analysis have persisted in the intervening decades. Several researchers have found the evolutionary relationships between behaviors to closely associate with morphological phylogenies (de Queiroz and Wimberger 1993, Prum 1994). Studying the lekking behavior in manakins, Prum states “vertebrate reproductive social behavior has an evolutionary history, and [it] is not determined solely by adaptive individual plasticity to current conditions” (1994). The pace of cultural evolution can be more rapid than that of genetic evolution, and determining whether behaviors are inherited from the same ancestral state could be problematic if rapid changes in behavior contribute to chance convergence (Atz 1970). Reviewing avian systematics and the vocalizations of songbirds, Robert Payne (1986) suggested that learned behaviors might not

reflect phylogeny since they are learned instead of inherited, but he conceded that the relative contributions of learning and inheritance are not clear. However, with some exceptions, traits that are culturally transmitted should still reflect the evolutionary history of the organism, since these are usually passed on from a conspecific and are subject to natural selection.

Features of the song that are conserved between species can be classified by whether they are shared by an outgroup species or present in descendants of a common ancestor but not in close outgroups. Given enough song characters that are present in some members of the group and not in the outgroup, a phylogeny can be inferred, such as for the innate vocalizations of African honeyguides, which are non-passerine birds (Payne 1986). Similarly, the vocalizations of herons were inspected for phylogenetic and ecological influence (McCracken and Sheldon 1997). This kind of analysis has rarely been extended to the songs of vocal learners; one study mapped several song characteristics onto a phylogeny in the *Dendroica* warblers (Payne 1986). Another study, in the vein of Gould's theories on ontogeny and phylogeny of morphological development (1977), focused on the parallels between phylogeny and song ontogeny, finding developmental constraints on repertoire size and variability to be evolutionarily conserved. The general opinion of birdsong researchers, as summarized by Payne, seems to be "that song may be more useful at the species level than at higher levels" (1986). Indeed, learned song is often more complex than innate behaviors and thus presents a greater challenge to characterization. However, the fact that oscine song is transmitted culturally rather than genetically should

not preclude a thorough consideration of song in light of the evolutionary history of these birds.

*Advantages of the birdsong system to study the evolution of behavior*

In the field, song studies have yielded information about regional dialects, species boundaries, sexual selection, fitness, and learning. However, learned song has many characteristics that make it well suited to a large-scale interspecies study. One monophyletic group of vocal learners comprises nearly half of known bird species, and the songs of many of these 4,500+ species are performed as repeated bouts of syllables, facilitating comparisons. Song is produced for the same purposes (territory defense and mate attraction), with the same organ (the syrinx), and in a comparable frequency range for all singers. Birdsong is also a behavior that can be studied objectively and computationally. Many animal behaviors need to be manually scored by trained observers, usually looking for one aspect of behavior, and records kept by one researcher might prove useless to another lab. In contrast, birdsong can be readily recorded in nature without disturbing the animals and then digitized and studied in the lab, and such studies avoid the expense and questionable predictive power of housing animals, making inbred lines, and creating artificial situations in which the animals can behave. Song recordings made by amateurs in harsh conditions can be just as valuable as those made by scientists at field research centers, and a recording from the 1950s can be easily digitized and mathematically analyzed. The long history of recordings by scientists and birdsong enthusiasts could enable the



investigation of song change over time, and the wide distribution of recording locations could facilitate a study of song differences by region.

To probe the evolutionary content of learned vocalizations in songbirds, I used a two-fold approach. First, I performed a computational analysis of existing song recordings. To determine the extent of the evolutionary information contained in this behavior, I examined the properties of the syllables themselves and the syntactical pattern of the song as a whole. I constructed both genetic phylogenies and trees of song relatedness and compared the song differences to the genetic distance. Within one species, I also tested for cultural changes over time and across distances. Birdsong is a fascinating system for studying the evolution of behavior, especially given the evolutionary relevance of birdsong, the genetic and cultural components of the behavior, and the vast and underutilized repositories of recording data. As a second avenue into this study, I hybridized two species of birds, the zebra finch (*Taeniopygia guttata*) and the silverbill finch (*Lonchura malabarica*), and raised them with either parent species. In addition, I raised juveniles of each species with an adult of the other species. From this set of experiments, I attempted to probe the genetic constraints on song learning by testing how well two species can learn one another's song and whether hybrids can perform better at this task than either purebred species. I also tested whether hybrids show any species-specific predispositions in song imitation or species-specific learning strategies. By combining these two contrasting approaches to the study of song evolution, I hope to shed light on the evolution of learned behaviors and the species constraints therein.

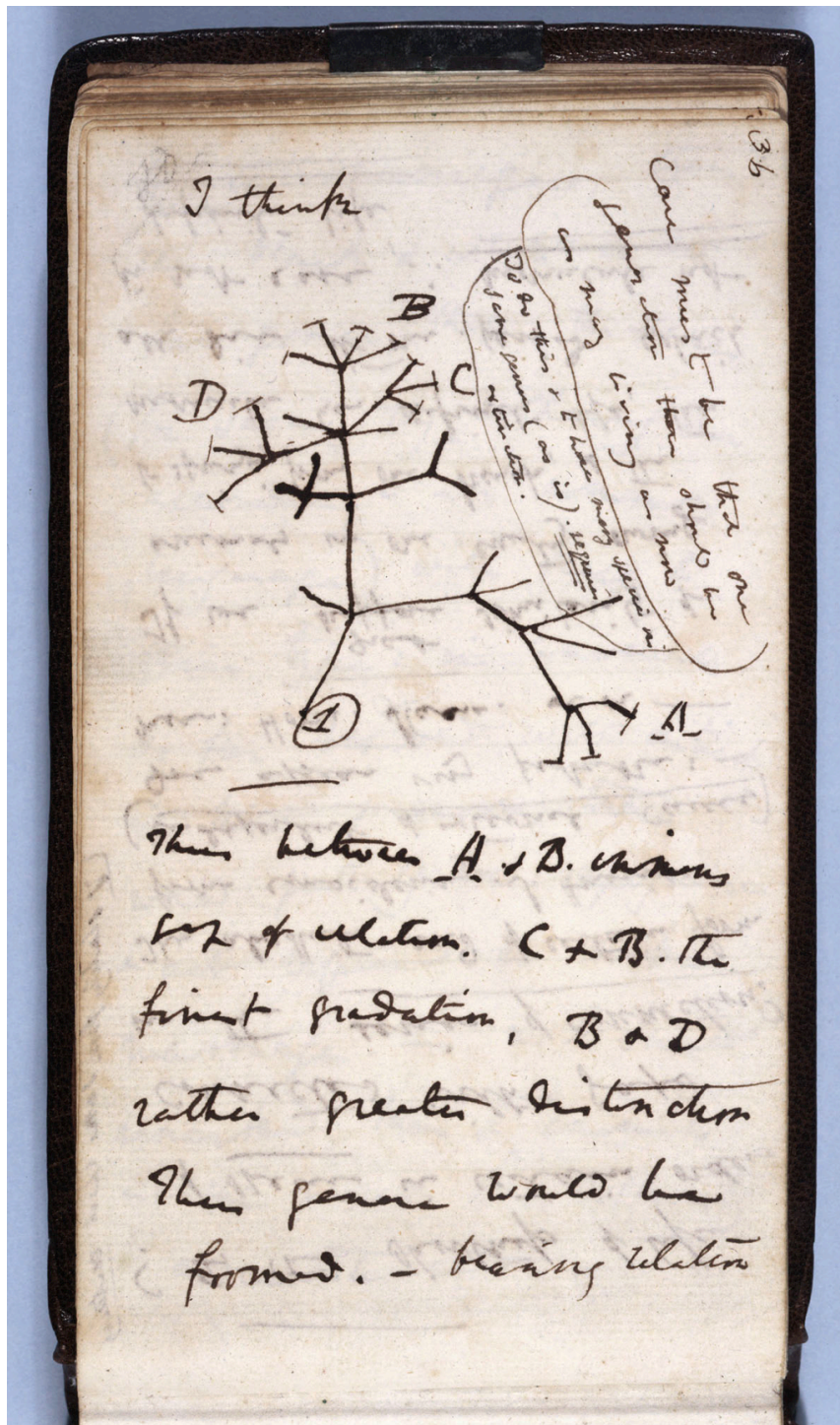


Figure 1-1. The first species phylogeny. This illustration of a species phylogeny by Darwin in his Notebook B (1837) marks the first time that the relationships between species were depicted as diverging from a common ancestor in this way.

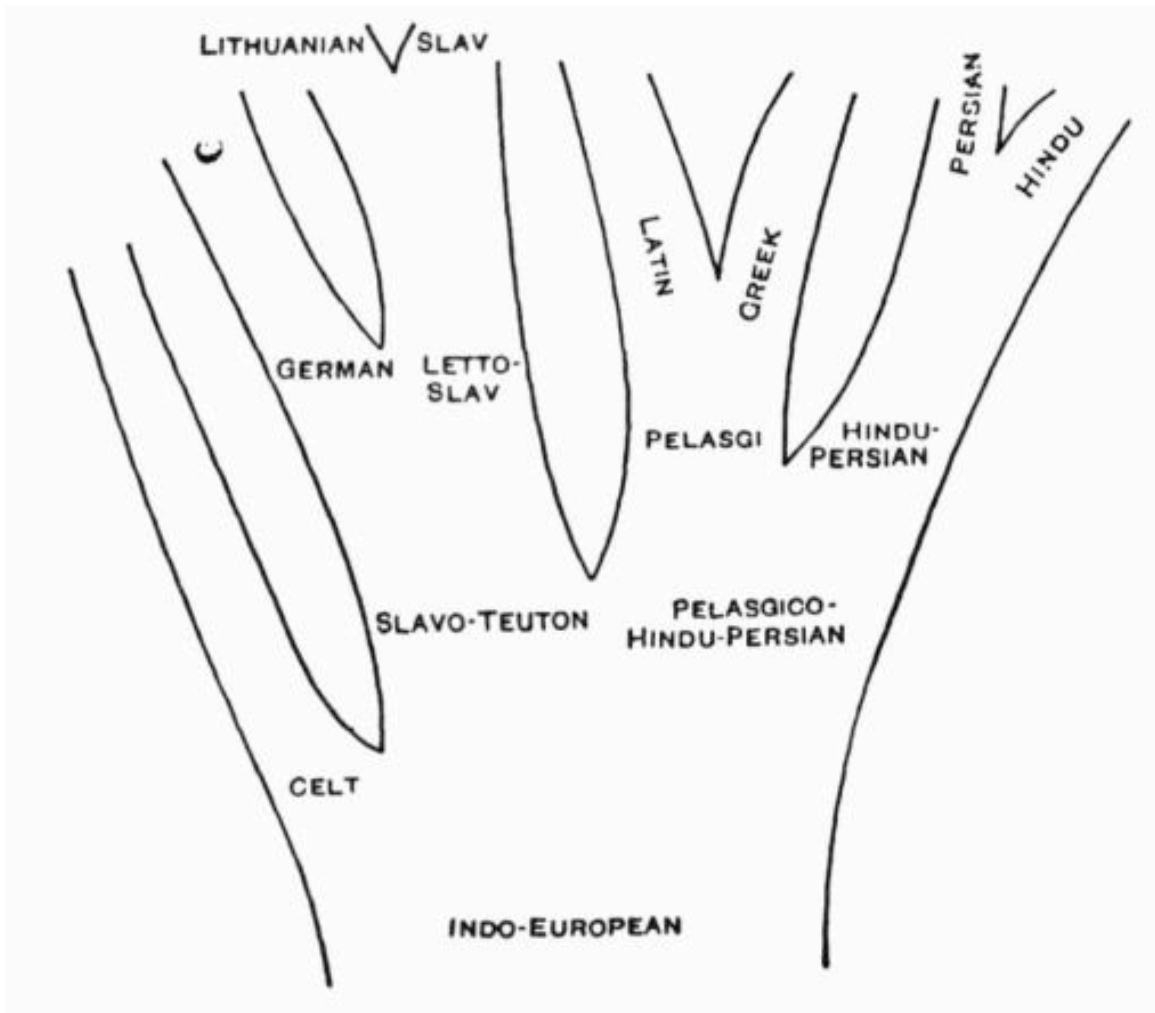


Figure 1-2. An early historical linguistic tree. Prior to Darwin's work on species, there was a long history of representing the relationships between languages with a tree-branching diagram. This language tree predates the *Origin of Species*, and Darwin was almost certainly familiar with ones like it (Schleicher 1853).

## Part 2. Song analysis of Emberizid sparrows

To begin testing the putative connections between species relatedness and song similarity, I needed to develop metrics of song analysis that were suitable for interspecific comparisons. The existing song analysis software is well suited to examining the fine structural changes that take place during song learning and development, but the fidelity of song imitation between pupil and tutor within one species was not possible to test given the database of recordings I used. Instead, I developed computational techniques for extracting syllable and song features from existing recordings and for determining the syntax of these songs in a way that made interspecific comparisons meaningful. The results of these analyses and comparisons could then be compared to genetic metrics of evolutionary history. These *in silico* experiments were the first systematic look at a large set of songbird vocalizations in the light of phylogeny.

### Methods

#### *Phylogeny construction*

To examine the genetic relationships between species in the Emberizidae family, I obtained DNA barcode sequences for available species and, using MEGA4 (Tamura et al. 2007), constructed a Minimum Evolution tree with 100 bootstrap replicates using the Maximum Composite Likelihood metric and pairwise deletion of gaps. I also calculated the genetic distance between each pair of species by dividing the number of nucleotide changes

between the two species' sequences by the number of nucleotide sites compared. This distance metric gives an average number of genetic changes per site and is useful when the unaligned sequences have different lengths.

#### *Data acquisition*

For each species in the phylogenetic analysis, I searched the Macaulay Library of the Cornell Lab of Ornithology database for recorded song. I chose at least ten song files from each species. When more than ten files were available, I chose files from different dates and places to minimize the chances of sampling the same bird more than once. The Macaulay Library generously donated the requested files for all subsequent work unless otherwise noted.

Each of the species selected for this analysis typically sings repeated bouts of one motif composed of a sequence of syllables, which can be defined as periods of sound separated by intervals of silence. The song files were manually screened and discarded if they contained no song bouts or if all instances of song were obscured by noise. The remaining files were separated into separate song files for each bout, and low frequency noise (below 1 kHz) was removed with Audacity software (<http://audacity.sourceforge.net/>).

#### *Sound spectrograms*

A sonogram, or sound spectrogram, is a three-dimensional representation of sound data with time on the x-axis, frequency on the y-axis, and the amplitude of each frequency

at a given time usually represented by the intensity of a color gradient (Figure 2-1A).

Spectrogram data can also be represented in a two-dimensional matrix for which the rows represent frequency intervals, the columns represent time intervals, and the magnitude of the number at  $\text{matrix}(x,y)$  corresponds to the amplitude of frequency  $x$  at time  $y$ .

Subsequent calculations performed on spectrograms should be assumed to refer to their numerical matrix representations, and the visual representation of these data will be used for illustration. Using MATLAB, a Gaussian-windowed sound spectrogram was generated for each file (Gardner and Magnasco 2006). This spectrogram was divided by its maximum value so that the values would range from zero to one. The spectrogram for each song bout was then segmented into syllables with a program designed to identify periods of sound greater than 16 milliseconds separated by a silence greater than 8 milliseconds (Figure 2-1C). For the purposes of segmentation, the spectrogram was made binary such that the top 1% of values were considered “sound” and were rounded up to one, whereas the bottom 99% of values were considered “silence” and were rounded down to zero (Figure 2.1-B) This helped filter noise and echoes so that the intervals between the syllables were silent. Occasionally syllables were made up of multiple notes, which were defined as discontinuous clusters of sound within a syllable (Figure 2-1D). Accordingly, each of these sound clusters was numbered sequentially using the MATLAB function `bwlabel`, and the properties of both the syllables and the component notes were analyzed.

### *Syllable properties*

After looking extensively at the songs for this group of species, I devised a list of syllable properties to use for statistical assessment of the song. For songs of the Emberizidae in general, each individual syllable was much less complex than a zebra finch syllable, so many of the advanced techniques developed for examining and comparing these syllables (Tchernichovski *et al.* 2000) were not used in this analysis. Within a song, I generated a numerical list of syllable “onsets” and “offsets” for the periods of sound and silence that were identified as syllables in the previous section, with the onset list representing the time each syllable began and the offset list enumerating the ending time of each syllable. I also prepared a copy of the binary spectrogram for which each note was numbered as described above. Clusters of sound that were too small to usually be considered song elements (those for which syllable length times height was less than 150 pixels, a threshold established by trial and error) were discarded. On the x-axis, each pixel represented a third of a millisecond, and on the y-axis, each pixel represented 43 Hz. With the spectrogram and these two metrics for separating the song into components, I recorded numerous variables describing each syllable and the song as a whole. Subtracting the syllable offsets from the syllable onsets gave a list of syllable durations for the song. The mean value of this list was recorded as the “mean syllable duration,” the maximum value gave the “length of the longest syllable,” and the minimum value was the “length of the shortest syllable.” Subtracting the syllable onsets from the previous syllable’s offset gave a list of silence durations between syllables, and the average value of this list was recorded as “mean silence duration.” The total number of syllable onsets gave the variable “number of syllables,” and

the time of the first syllable onset subtracted from the time of the last syllable offset equaled the “duration of bout.” Additionally, the variable “number of syllables divided by bout duration” was calculated from the two previous variables to represent the speed at which the song was delivered. Next, each note of the song was examined individually. I recorded the length of each note and calculated from this list the “mean note length,” “length of the shortest note,” and “length of the longest note.” The total number of notes in the song was used to calculate “number of notes,” and the “average number of notes per syllable” equaled the number of notes divided by the number of syllables. I then analyzed the frequency characteristics: for each note, the minimum and maximum frequency were recorded. The “song frequency peak” was the highest value in this list, and the “song frequency trough” was the lowest. The song frequency peak minus the song frequency trough equaled the “frequency range” for that song. The “mean frequency peak” was calculated by averaging the maximum frequency of each note, and the “mean frequency trough” was the average minimum frequency of each note. The minimum frequency of each note subtracted from the corresponding maximum frequency gave the frequency range for each note, and the average of these values was recorded as “mean note frequency range.”

After inspecting many notes, I hypothesized that the vast majority could be fitted to a second-order polynomial function; that is, most notes appeared to be a section of a parabola. With this in mind, I found the mean frequency value at each small time window, plotted these values, and then found the best fitting function (Figure 2-1E-F). I then recorded properties of this function to determine the note’s shape, namely the slope at



each point, the existence and location of a peak or trough in the function. If the function had a peak or trough near the center of the note, its shape was deemed a “parabola.” A predominantly negative slope was used to classify “downsweeps,” which included notes with a peak in the first 20% of the note, a trough in the last 20% of the note, or a negative slope over the whole note. Similarly, “upsweeps” included notes with a positive slope for the duration of the note, a trough in the first 20% of the note, or a peak in the last 20% of the note. “Pure tones” were syllables with a very flat function, regardless of the shape of the function. If the highest and lowest values of the fitted function spanned less than 150 Hz, it was classified as a pure tone.

#### *Syllable similarity and syntax*

Field guides often represent bird vocalizations with memorable phrases: “drink your tea,” “hey sweetie,” “fee-bee,” “who cooks for you?” These phrases say little about the mean syllable properties of a bird’s song as described in the previous section; rather, that a species’ song can be generalized in a short mnemonic implies the syntax of these syllables is somewhat conserved between conspecifics. To quantify the syllable pattern of a bird’s song, I first needed a metric to classify two syllables as a repetition of the same element. For this analysis, I separated the song into syllables using the syllable onsets and offsets as previously defined. I then performed a pairwise comparison of syllables, in which each syllable is compared to every other syllable. Since the spectrogram had been previously scaled so that the intensity values ranged from zero to one, multiplying the corresponding elements from different sections of the spectrogram could produce valuable data (Figure 2-2). If two

segments of the spectrogram had overlapping areas of signal, the product of two positive values would result in a positive number. However, if either segment of the spectrogram had no signal, multiplication of those elements would result in zero. Therefore, a sum of the product of two spectrogram segments should result in a number that correlates to the degree of overlap between the two segments. As a baseline, I multiplied by itself the section of the spectrogram corresponding to each syllable and summed the result to acquire a self-similarity score for each syllable. For the comparison of two different syllables, I determined which syllable was shorter. The shorter syllable was moved in a sliding window over the longer syllable, and the overlap was calculated at each point. The maximum overlap between the two syllables was then divided by the greater of the two self-similarity scores for the selected syllables, thus scaling the overlap between two syllables to a percentage of the maximum possible overlap for those syllables. This similarity was recorded in a matrix for each pair of syllables. If two syllables did not overlap at all, the score was very nearly zero. Through trial and error, I determined that a score of 50% was sufficient to determine that a syllable was the repetition of another syllable. The pairwise matrix of syllable similarities was thresholded such that scores less than 50 were set to zero. Then, for each syllable after the first one, I determined whether it was a repetition of a previous syllable by looking for a score greater than 50 and recording the number of its first syllable match. If it did not match any prior syllables, it retained its own syllable number. This algorithm also checked for previous repetitions; for example, if syllable five matched to syllable three, but syllable three had already matched to syllable one, all three were assigned the number one. This calculation produced a numerical representation of the syllable syntax of a song, such that

a bout that alternated between two syllables would be encoded as 1, 2, 1, 2, etc. These numerical representations were then converted to alphabetical ones by replacing the most common number with the letter A, the second most common with B, and so on. If two numbers were equally common, the one that occurred first was given the first letter.

Once the syllables were encoded in this way, I could add several more variables to the set of syllable properties. Using the numerical representation of the song syntax, I determined the “number of unique syllables” by counting the distinct elements in the list of syllable numbers. The repetitiveness of the song was then quantified in several ways. An overall average was obtained as the “number of syllables divided by number of unique syllables.” The fraction of times a syllable was followed by the same syllable was calculated by subtracting each element of the numerical song syntax from the previous element. If the same syllable was repeated, the result of this subtraction would be zero, so by counting the number of zeros and dividing by the number of syllables, I calculated the “degree of sequential repetition.”

### *Syllable pattern*

To compare the syllable syntax between the Emberizid species, I first determined a representative syntax for each species. I examined the syntax from ten members of each species and counted the number of repetitions of each element. Most birds in this group had very repetitive songs, so I calculated the average number of instances of the repeated element. For syllables that were not repeated, I used a majority rule algorithm for inclusion in the representative syntax. For example, for the three songs CBBBAAAAA,

CBBBAAAAAA, and BBBAAAAAAA, the majority of songs begin with a C, all have three B's, and there were an average of six A, so the representative syntax would be CBBBAAAAAA.

I then determined the distance between the syntax of each pair of species using a modified edit distance metric. Much like the comparison of two nucleotide sequences, this metric determined the distance between two strings of letters by assigning a penalty for each insertion, deletion, substitution, or conversion to a new repeated element. From the distance matrix that resulted from these comparisons, I could visualize the relationship between the syntax of each species using a neighbor-joining tree. This type of tree was formed using an algorithm that groups the two species with the smallest syntax distance, merges them into a cluster, and then recalculates the distances between every other species and the new cluster. This procedure was repeated until every element is in one cluster, creating a hierarchy of progressively larger clusters that represented the relationships between the syntax of every species analyzed.

### *Phonology*

The syllable property variables were generated for ten songs each for the *Zonotrichia*, *Spizella*, and *Pipilo* species. A distribution of frequency value in hertz is not directly comparable to counts of the number of syllables, however, so the data were standardized. For each variable, I calculated the mean and standard deviation across all birds sampled. Then, for each value of that variable, I subtracted the mean and divided by the standard deviation so that the variables would have overlapping distributions and could be

compared on the same scale. Finally, I took the species mean of each syllable property variable.

To calculate the interspecies “song distance,” I calculated for each pair of species the Euclidean distance between the average species values for all syllable properties. Sometimes referred to as “distance as the crow flies,” the Euclidean distance for two points in space ( $x$  and  $y$ ) is the length of the line segment connecting the two points, calculated by taking the square root of  $(x - y)^2$ . If  $x$  and  $y$  contain multiple measurements, as is the case for the syllable properties, the distance is calculated by taking the square root of  $[(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2]$ . In this way, the distance between the set of mean syllable properties for each pair of species was established. From this distance matrix, I followed the methods for constructing a neighbor-joining tree as described for syllable syntax.

Next, I compared this pairwise song distance to the pairwise genetic distance that I calculated by making the Emberizidae phylogeny. For corresponding pairs of species, I recorded the genetic distance in one column of a matrix and the song distance in another column. I then plotted genetic distance versus song distance and calculated Pearson’s correlation coefficient and Spearman’s rank correlation with corresponding p-values.

#### *Factor analysis*

When choosing the variables for the analysis of syllable properties, I sought to quantify the mean and extreme values of several aspects of syllable phonology. However, some of these variables were clearly interdependent. To reduce the dimensionality of the data set, I performed a factor analysis using SPSS (IBM). Related to a principal component

analysis, a factor analysis finds highly correlated variables, combines them into factors, and returns the amount of the variance in the variables that is explained by these factors. For example, field ornithologists usually collect data for numerous variables when they capture and color-band a bird: weight, wing length, tarsus length, beak length, etc. These variables could potentially be thought of as multiple observations of a more general “size” factor. However, some of the variance in these variables might be unrelated to the factor; for example, beak length might correlate to the general size of the bird but also vary based on food source. Performing a factor analysis on these data would demonstrate which correlated variables could be combined into factors and what percentage of the variance was explained by these factors. In contrast, a principal component analysis attempts to find underlying uncorrelated components that account for all of the variance in the variables, so this technique often explains more of the variance than a factor analysis but the components can be difficult to interpret in relation to the known variables.

I performed the factor analysis with the default parameters and ‘varimax’ rotation. One output of a factor analysis is the factor score for each of the input observations. Thus, for all songs in the analysis, the value of each factor was known, and I used these scores to generate a neighbor-joining tree by the procedure outlined above.

## Results

### *Phylogenetic construction*

Using data from the Barcode of Life Data Systems database, I constructed a phylogenetic tree for members of the family Emberizidae (Figure 2-3). From this phylogeny, I chose three branches to perform a song analysis. There were six branches of more than three species that had at least 75% bootstrap support. I chose the two largest of these groups, containing the genera *Spizella* and *Zonotrichia*, and one additional branch (*Pipilo*), for further analysis (Figure 2-4). One species in the *Spizella* genus, the American tree sparrow (*Spizella arborea*), was not in the same genetic group as the rest of the members of its genus, which was consistent with a previous molecular phylogenetic analysis of the Emberizid sparrows (Carson and Spicer 2003) but not with a tree constructed from morphological character state data (Patten and Fugate 1998). This species was included in the song analysis to determine whether its song was more associated with its closest genetic relatives or with its morphologically similar congeners.

### *Analysis of song data*

The Cornell Library of Ornithology graciously provided the data for these experiments. I compiled ten songs from each of the species identified in the previous phylogenetic analysis except for the clay-colored sparrow, which did not have sufficient recordings to be included. After the reduction of low frequency noise, an excerpt of one bout from each of these songs was processed in MATLAB, and the syllable properties and

syntax for each song were recorded and analyzed. A visual inspection of song spectrograms revealed qualitative similarities between songs within each of the three phylogenetic branches (Figure 2-5). I constructed a neighbor-joining tree from the syllable syntax data (Figure 2-6) and the syllable properties data (Figure 2-7).

Next, I compared the song distance, the Euclidean distance between syllable property variables of two species, to the genetic distance, the average number of genetic changes per site between the sequences of the same two species (Figure 2-8). The song distance and genetic distance were significantly correlated for this set of species (Pearson's correlation coefficient = 0.242,  $p = 0.033$ ), and the rank correlation test was nearly significant (Spearman's  $\rho = 0.219$ ,  $p = 0.054$ ).

#### *Factor analysis*

The factor analysis of Emberizid syllable properties returned eight factors, which accounted for the variance of all input variables (Table 1). Each of these factors explained the most variance in variables that seemed intuitively related. The first factor correlated best to variables related to syllable length, the second to variables related to the number of syllables, the third to the frequency properties of the syllables, the fourth to the change in frequency within a syllable, the fifth to the pattern of repetition, the sixth to syllable complexity, the seventh to the degree of repetition, and the eighth to syllable shape. These factors explain 85.1% of the variance in the data, but the remainder of the variance seems crucial to fully capturing the relationships between different species. The neighbor-joining



tree of factor scores for each species was not as accurate as the tree made from the standardized data (Figure 2-9).

## Discussion

A computational study of fourteen Emberizid sparrows in three genera revealed a robust and significant relationship between evolutionary history and song similarity. The syntax of a species' song accurately grouped it with other members of its genus (Figure 2-6). The finer structure of the genetic phylogeny was not as clearly resolved by this analysis: only one sister species relationship was maintained between the genetic and song trees, in the *Pipilo* species. For the other branches, the two most closely related song patterns were not sung by the most closely related species. In contrast to Payne's (1986) prediction that song makes the most useful predictions at the species level, a comparison of syllable syntax patterns from three monophyletic groups was most informative at the genus level.

In comparison, a neighbor-joining tree constructed from the normalized values of syllable properties did not as clearly delineate the coarse structure of the genetic phylogeny (Figure 2-7). The members of the *Spizella* genus (except the spuriously assigned *Spizella arborea*) were grouped together. The other two genera are paraphyletic, meaning that a monophyletic branch containing all of the members of either *Pipilo* or *Zonotrichia* would also contain species in other genera. However, the three most closely related branches combine members of one genus, and there is one such branch for each genus.

The neighbor-joining tree constructed from the results of the factor analysis was less accurate than the tree constructed from the syllable properties themselves (Figure 2-9).

However, the factor analysis demonstrated that, even though some variables are closely correlated with each other, all of the variables studied are contributing information to the analysis.

Trees generated from song data indicated that a relationship existed between the songs of these species and their genetic relatedness. To test the significance of this relationship most directly, we recorded the Euclidean distance of the syllable properties (referred to as song distance) and the genetic distance between each pair of organisms. These data points yield a noisy scatter plot, but the positive correlation is statistically significant: as genetic distance increases, so does song distance.

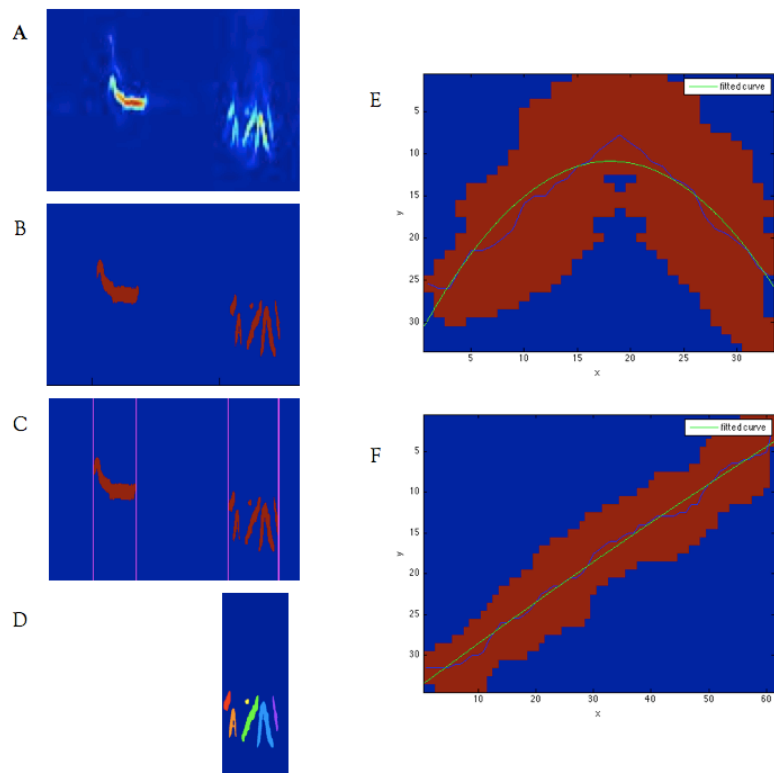


Figure 2-1. Song spectrogram analysis. In this example, the spectrogram of two syllables (A) was subjected to a threshold such that signals above the threshold were set to 1 and signals below were set to zero (B). The thresholded spectrogram was scanned for onsets and offsets of signal and segmented into syllables accordingly (C). Syllables with multiple parts without a full drop in amplitude between them were segmented into notes according to contiguous clusters of signal (D). These notes were then fit to a function to describe their shape (E, F). Horizontal axes are time and vertical axes are frequency.

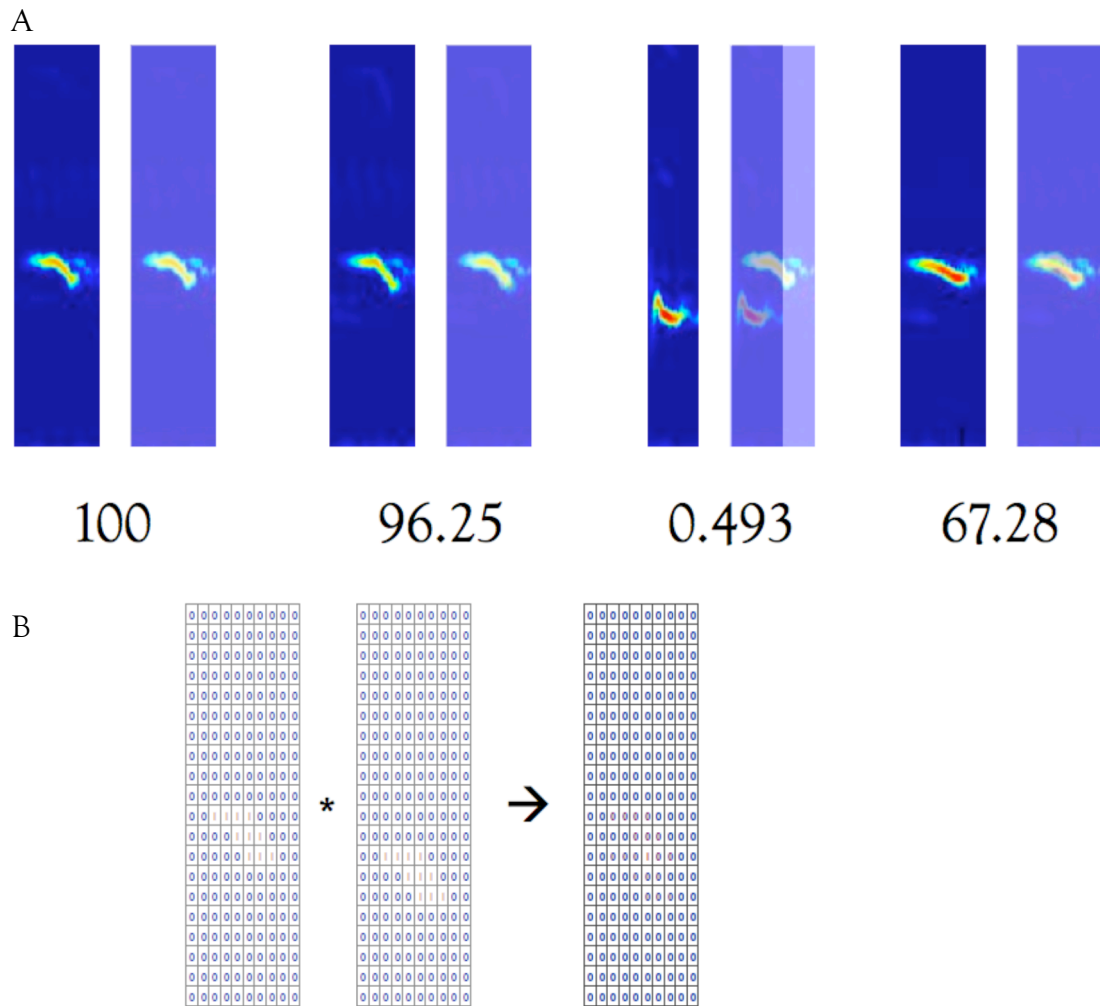
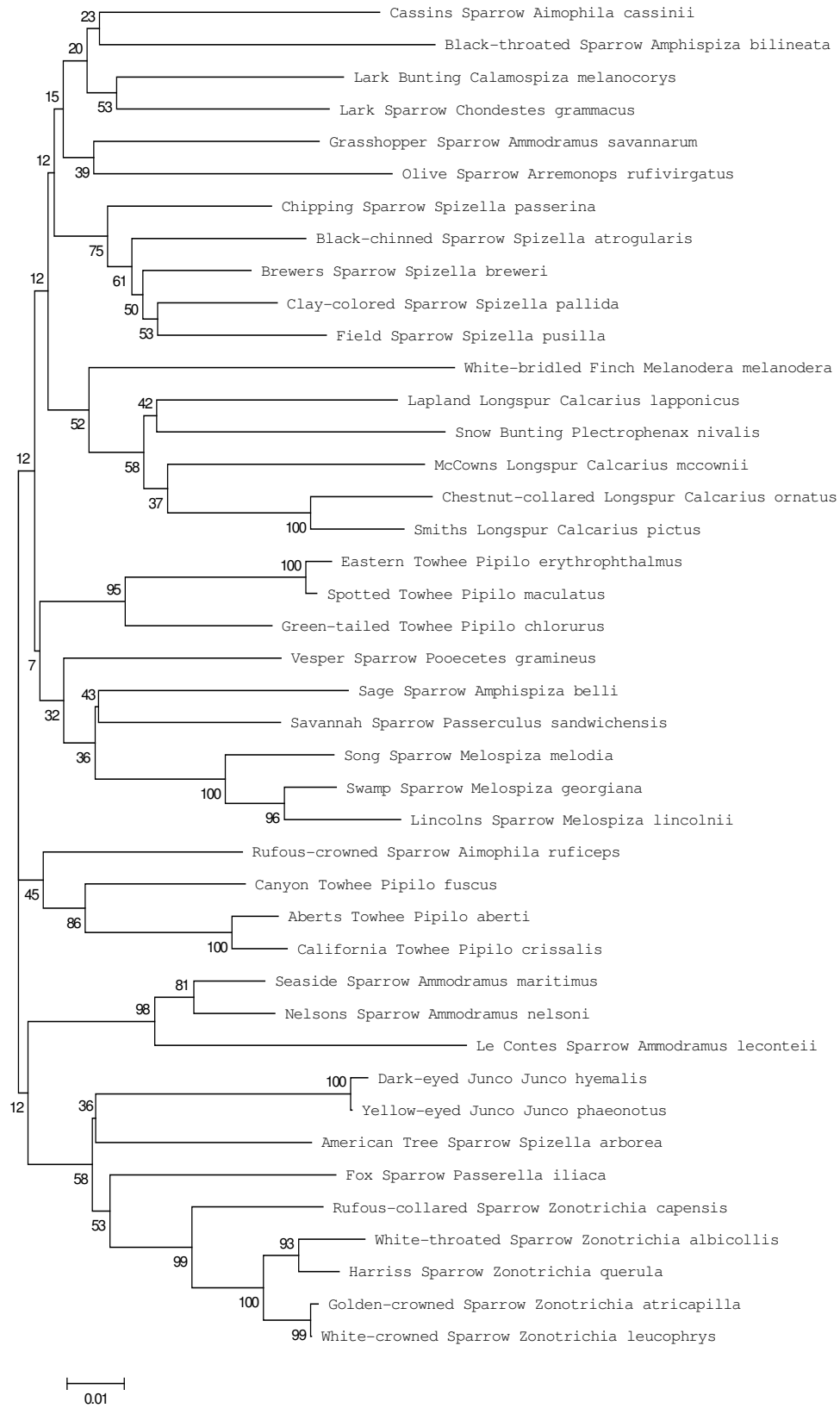


Figure 2-2. Calculating spectrogram overlap. When the elements of a spectrogram are scaled between zero and one, multiplying the spectrograms and summing the products can give a measure of overlap. This overlap is then scaled to a percentage of the maximum overlap between a syllable and itself. Four examples are given (A), in which each syllable is compared to the first one. A schematic overlap demonstrates that only for overlapping areas of signal will the product of the two matrices be greater than zero (B). Only one pixel has a “1” in both images, so the overlap between the two images is 1. All of the other pixels with a “1” in one image are multiplied by a “0” in the other image and do not add to the overlap. Horizontal axes are time and vertical axes are frequency.

Figure 2-3. DNA Barcode phylogeny of the Emberizidae. Overall, the phylogeny is not well-supported and the bootstrap values are low. Several genera appear in weakly supported branches in different branches of the tree (*Aimophila*, *Amphispiza*, *Ammodramus*). Well supported groups (*Spizella*, *Pipilo*, *Zonotrichia*) were used for subsequent song analysis. The scale bar represents average number of base pair changes per site.



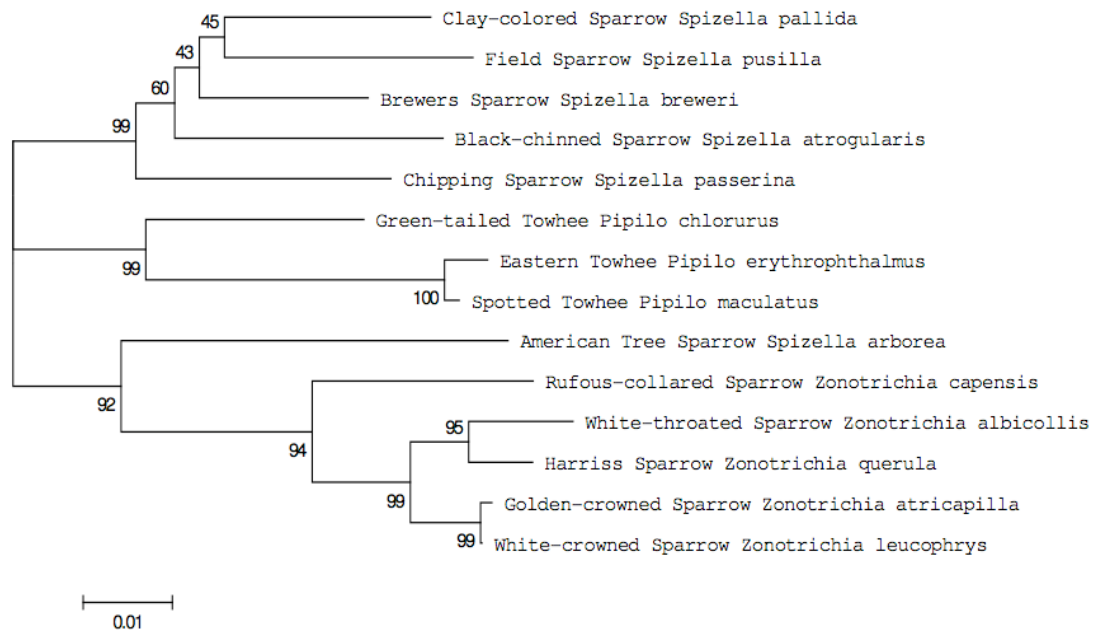


Figure 2-4. Phylogeny of the analyzed subset of Emberizidae. The relationship between the three groups is unresolved, but the groups themselves are strongly supported, with bootstrap values above 92. The scale bar represents average number of base pair changes per site.

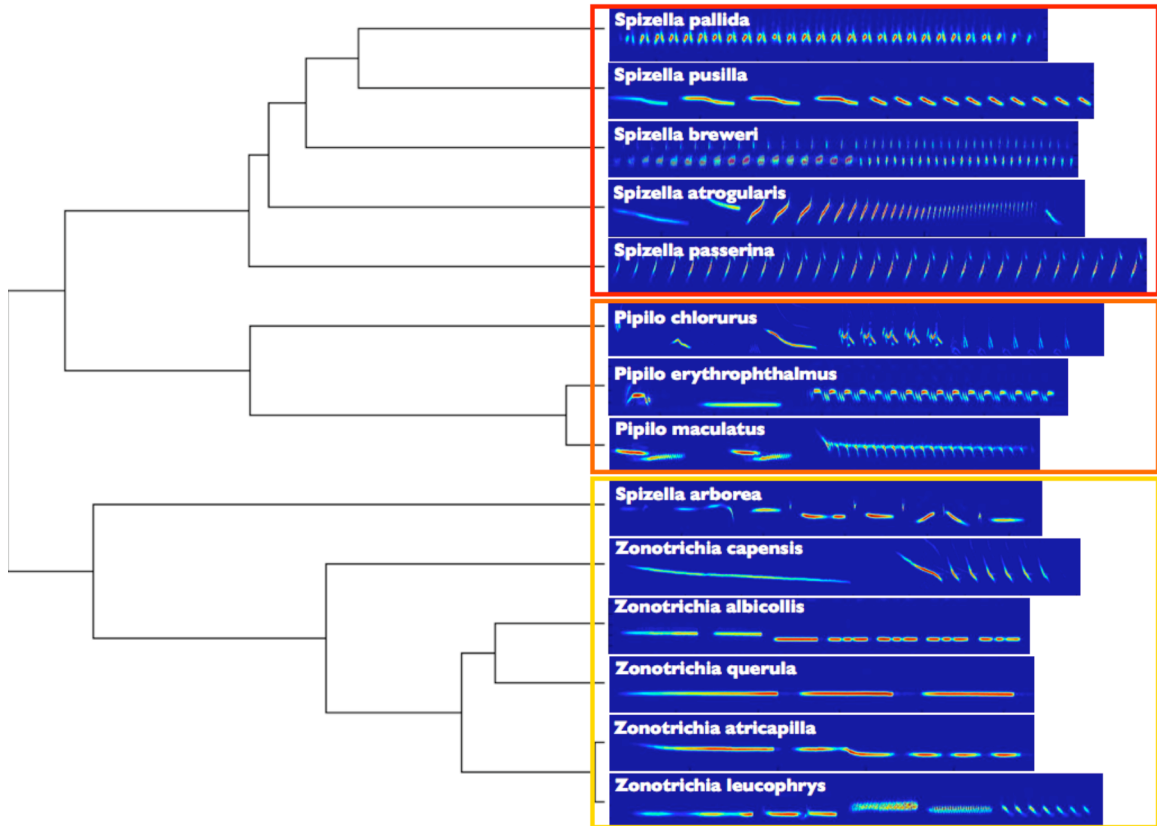


Figure 2-5. Emberizidae phylogeny with song illustrations. A phylogeny of the species used in the Emberizidae song analysis was overlaid with song spectrograms to illustrate the similarities between songs observable by eye. The top branch, the *Spizella*, have many repetitions of one or two syllables, sometimes increasing in rate. In the *Pipilo* branch, two syllables are followed by a short repeated note. In the *Zonotrichia*, most birds sing a series of long pure tones.



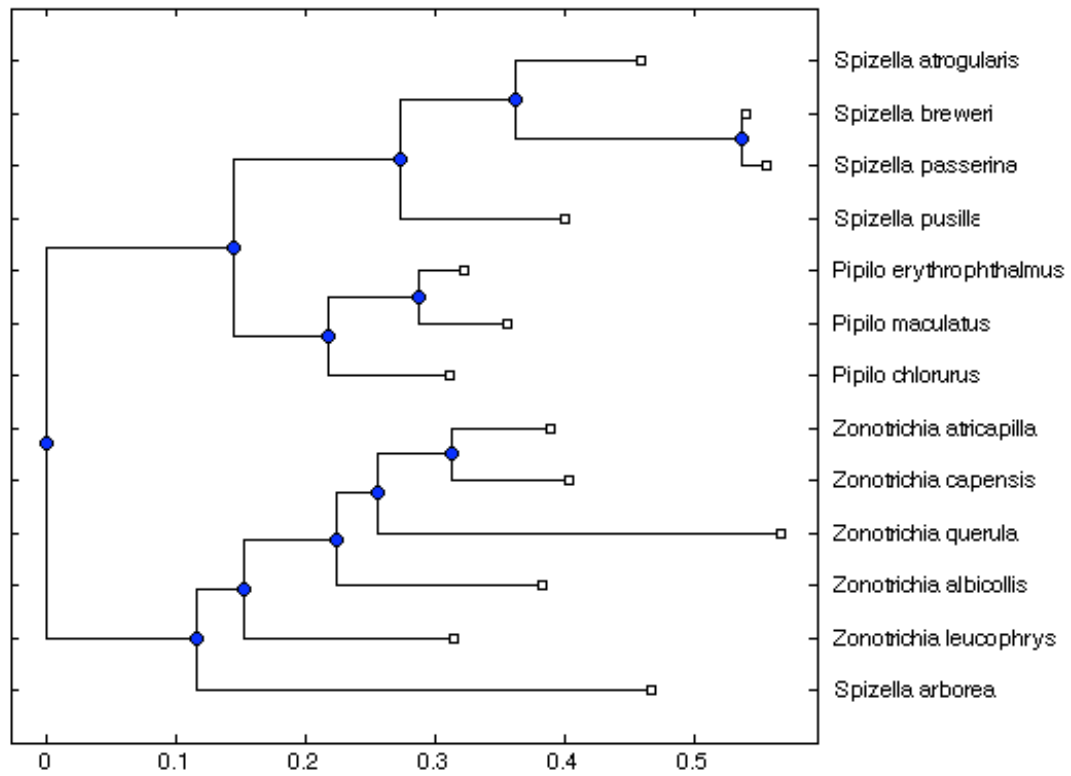


Figure 2-6. Neighbor-joining tree of syllable syntax distances. The edit distance between the syntax for each pair of species was calculated, and a neighbor joining tree was constructed from these distances. In this case, the data from song syntax clustered the species into their genetic genus groups, with *Spizella arborea* maintained as an outgroup to *Zonotrichia*. For this and all other neighbor-joining trees, the x-axis is a relative measure of distance and the y-axis is not meaningful.

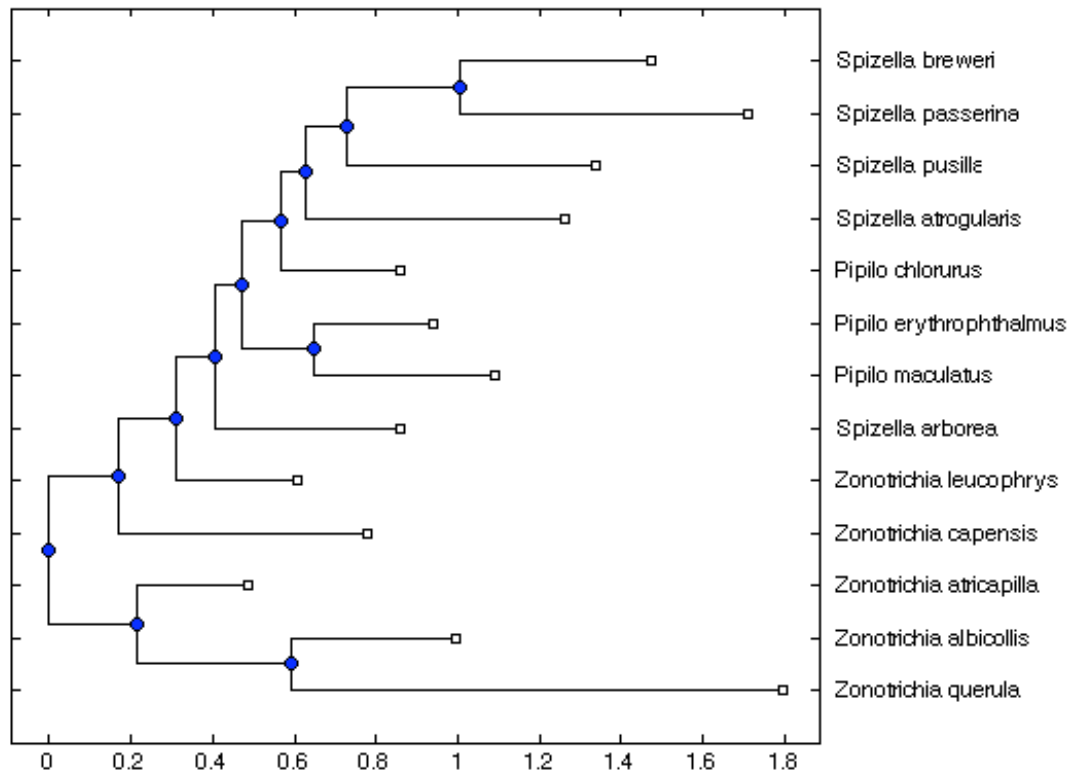


Figure 2-7. Neighbor-joining tree of normalized syllable property distances. The phonology characteristics of Emberizidae songs grouped closely related species but did not fully reconstruct the relationships between them. In this tree, *Spizella* is a monophyletic clade, and two *Pipilo* and three *Zonotrichia* grouped together, but the rest of the species were paraphyletic, i.e. they did not form a monophyletic group without the inclusion of *Spizella*.

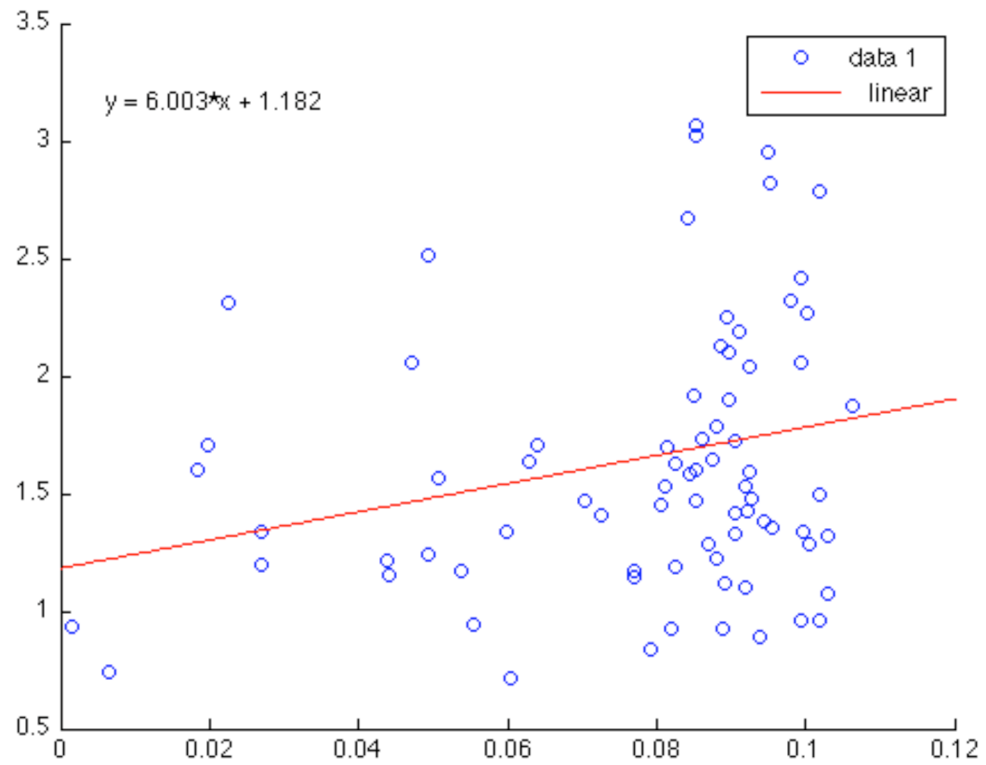


Figure 2-8. Correlation between genetic distance and song distance. The x-axis represents genetic distance, measured in the number of base pair differences between two sequences divided by the sequence length. The y-axis represents song distance, represented by the Euclidean distance between a normalized set of syllable property values for each species. There is a significant positive linear correlation between these two values.

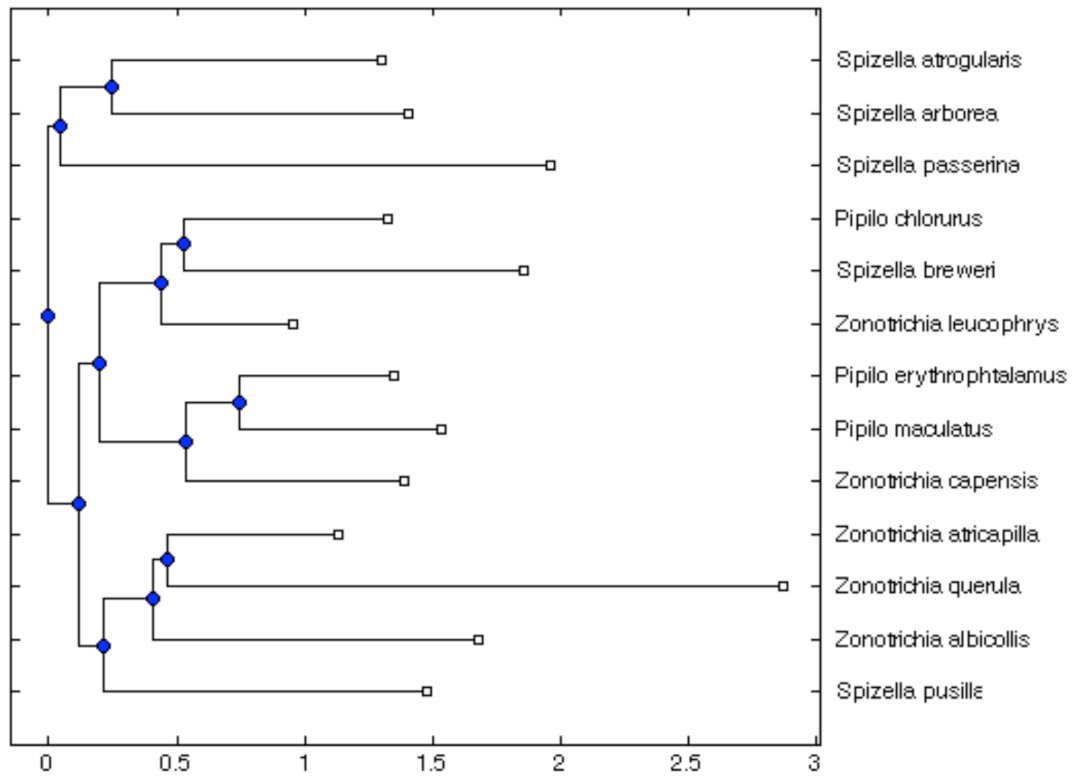


Figure 2-9. Neighbor-joining tree of factor analysis results. When the species values for each of the factors were used to make a distance matrix and then a neighbor-joining tree, the relationships between the species are less clear than when the data itself is used. For example, each branch of this tree contains species from different genetic clades: *Spizella arborea* is grouped with the true *Spizella*, and every other clade contains members of multiple genera.

|  | Component |       |       |       |       |       |       |       |
|--|-----------|-------|-------|-------|-------|-------|-------|-------|
|  | 1         | 2     | 3     | 4     | 5     | 6     | 7     | 8     |
| length of shortest syllable                      | .959      | .001  | -.065 | -.040 | .021  | .021  | -.072 | .041  |
| length of shortest note                          | .953      | .003  | -.074 | -.049 | .023  | -.011 | -.045 | .021  |
| average note length                              | .914      | -.181 | -.224 | -.155 | -.046 | -.137 | -.039 | -.006 |
| mean syllable durations                          | .910      | -.212 | -.218 | -.154 | -.064 | -.082 | -.066 | .004  |
| mean silence durations                           | .815      | -.283 | -.158 | -.181 | .050  | -.280 | -.014 | .044  |
| length of longest syllable                       | .686      | -.404 | -.392 | -.208 | -.052 | -.209 | -.051 | -.034 |
| length of longest note                           | .685      | -.403 | -.393 | -.210 | -.052 | -.220 | -.054 | -.031 |
| number of syllables divided by number of uniques | -.106     | .888  | .109  | -.043 | .089  | -.090 | -.120 | .000  |
| number of syllables                              | -.299     | .813  | .101  | .121  | -.041 | .073  | .375  | -.053 |
| number of syllables divided by bout duration     | -.374     | .718  | .225  | .215  | -.034 | .277  | -.069 | -.114 |
| number of notes                                  | -.266     | .700  | .156  | -.012 | -.009 | .524  | .259  | -.059 |
| minimum frequency                                | .307      | .501  | .266  | -.327 | .289  | -.362 | .161  | -.179 |
| average frequency peak                           | -.177     | .345  | .835  | .207  | -.029 | -.112 | .010  | .011  |
| maximum frequency                                | -.273     | .037  | .832  | .106  | -.050 | .224  | -.039 | .047  |
| average frequency trough                         | -.099     | .297  | .815  | -.144 | .029  | .013  | .074  | -.220 |
| frequency range                                  | -.362     | -.137 | .690  | .212  | -.146 | .335  | -.092 | .106  |
| downsweeps                                       | -.206     | -.050 | -.017 | .909  | .113  | .065  | .084  | -.071 |
| mean note frequency range                        | -.194     | .225  | .400  | .613  | -.099 | -.236 | -.090 | .350  |
| flat syllables                                   | .518      | -.232 | -.348 | -.612 | -.023 | .048  | .168  | -.114 |
| sequential repetition                            | .156      | .243  | -.179 | .033  | .817  | .068  | -.007 | .010  |
| notes per syllable                               | -.215     | .131  | .150  | -.059 | .084  | .844  | -.156 | .013  |
| duration of bout                                 | -.008     | .283  | -.234 | -.241 | .115  | -.325 | .759  | -.008 |
| number of unique syllables                       | -.413     | -.100 | .231  | .302  | -.187 | .074  | .660  | .135  |
| syllable stereotypy                              | .038      | -.244 | -.032 | .079  | .181  | .034  | .063  | .687  |
| upsweeps   | -.290     | .410  | .252  | -.191 | -.192 | -.110 | -.249 | .607  |
| parabolas  | -.306     | -.016 | .408  | -.075 | .127  | -.060 | -.147 | -.528 |

Table 2-1. Grouping of Emberizid syllable properties by factor analysis. The analysis reduces highly correlated variables to underlying factors. In this case, the factors were intuitively meaningful: syllable length, song length, frequency, frequency range, repetition, syllable complexity, song length, and syllable shapes and stereotypy.

### Part 3. Genetic phylogeny of Passeriformes

After developing the computational tools to compare bird songs between species and in relation to genetic similarity, I attempted to extend the analysis to more members of the Emberizidae family. Interpreting the relationship between genetics and behavior was very difficult when the genetic relationships between species were themselves indistinct. With this in mind, I sought to perform a large song analysis on a group of birds with a well-defined genetic phylogeny. To test the relationship between song distance and genetic distance on multiple levels of biological classification, I required a phylogeny with several species per genus, several genera per family, and several families in the Passeriformes order, with an analysis of bootstrap support for each branch and numerical values for genetic distance between pairs of species.

Putatively due to a large and ancient species radiation, the avian phylogeny remains contested even after numerous molecular genetics studies. Most studies to date focus on a small group of species (Moreno *et al.* 2003, Slack *et al.* 2006, Backström *et al.* 2008) or the two sequenced genomes (Nam *et al.* 2010), and a recent study of 169 species included most families in the Passeriformes order but only one species per family (Hackett *et al.* 2008). An avian “supertree” has also been constructed, which compiled many published phylogenies to generate a tree of 1723 oscine species (Jönsson and Fjeldskå 2006). Because this tree was generated as a summary of existing trees, it could provide neither measures of genetic distance nor indications of robustness for each branch. In order to continue my song

analysis based on these criteria, I set out to construct a phylogeny from existing avian genetic data.

For a set of species, there exists a “true” phylogeny; that is, the species have an evolutionary history with a specific lineage that can theoretically be traced back to the branching point of the most recent common ancestor. In practice, however, estimating the evolutionary history of a set of organisms can be complicated by numerous factors. Morphological data can include spurious features, genetic mutations can mutate back to the original form, and individual genes do not necessarily follow the same evolutionary trajectory as the organism itself. When constructing a phylogeny, similar organisms are grouped into branches of the phylogenetic tree, also known as clades. A clade typically refers to a phylogenetic grouping that contains an ancestor and all of its descendants, also known as a monophyletic group. To measure a level of confidence in the clade groupings of a phylogeny, a bootstrap test can be performed. This test resamples the data in the original phylogeny with replacement, so that some organisms might be sampled more than once and others not at all, and generates a new phylogeny with the same number of taxa. This procedure is replicated many times, and the percentage of the times that a given clade is reconstructed from the resampled data is called its bootstrap value. If the relationship between the members of the clade is strong, it should be robust to small perturbations such as duplicating or deleting a member.

## Methods

### *Data collection and processing*

To begin, I searched the National Center for Biotechnology Information (NCBI) GenBank database (Benson *et al.* 1994) for all genes tagged with the Taxon ID for Passeriformes (9126) and examined the resulting list for gene sequences that were sampled from multiple species. There were 19 genes with greater than 100 examples from the Passeriformes, and for each of these genes, I downloaded all available nucleotide sequences. The analysis included seven mitochondrial genes – ATPase (1671 sequences), COI (4774 sequences), COII (431 sequences), cytb (7714 sequences), ND2 (8104 sequences), ND3 (1895 sequences), ND5 (482 sequences) – and 12 genomic genes – AldoB (401 sequences), beta-fibrinogen (1918 sequences), c-Myc (174 sequences), CHDZ (197 sequences), GAPDH (1121 sequences), myoglobin (1758 sequences), ODC (949 sequences), RAG1 (1051 sequences), RAG2 (533 sequences), RHO (128 sequences), TGFB2 (624 sequences), and ZENK (181 sequences). In addition, I downloaded 5418 mitochondrial COI-5P sequences, which were the publicly available DNA barcode sequences for the Passeriformes order (Ratnasingham and Hebert 2007). For each of these genes, I aligned the sequences using the default values in MUSCLE 3.6 (Edgar 2004) and visually inspected the alignments in MEGA4. Often, sequences submitted to the NCBI database do not provide information for the entire gene, and these partial sequences may not overlap with one another. In this case, the two non-overlapping sections cannot be compared in an informative way. I identified the section of the gene with the most samples, and sequences that did not cover



this region of the gene were discarded. Sequences that contained information for more than one gene were cropped to include only the gene of interest. The sequences that passed these criteria for each gene were realigned and rechecked for sequences that could not be readily compared to the others. For the resulting set of sequences, I created a species consensus sequence for each gene using MATLAB. I imported the unaligned sequences and their corresponding species labels, aligned the sequences for each species, and retained the most common nucleotide at a given site. As a default, gaps were not counted in the calculation for most common nucleotide, so if only one sequence contained sequence information at a certain site, that information was retained in the consensus sequence. This setting was desirable to maintain as much information as possible since many of the sequences were incomplete. If multiple nucleotides were represented in equal numbers at a certain site, the consensus sequence at that site contained the nucleotide that was highest in the sequence C > G > A > T according to the convention of the `seqconsensus` function in MATLAB. Since the Passeriformes DNA barcode sequences were already aligned with all other barcodes by the Barcoding Life initiative, the consensus sequence was generated from the aligned sequences. However, since these sequences were taken from a larger alignment, some sites had a gap for every sequence and thus contained no information. For the `seqconsensus` function to properly recognize these all-gap sites, it was run with the additional parameters 'Alphabet', 'NT', 'Gaps', 'all'. This instructs the program use a nucleotide "alphabet" and to count all gaps, as opposed to the default value of counting no gaps as mentioned above.

### *Concatenation of species alignments and construction of phylogeny*

Consensus sequences for each gene were aligned using the default values in MUSCLE 3.6. Using MATLAB, I then created concatenated alignments for further study. For each species with a DNA barcode sequence, I searched every mitochondrial gene alignment for a sequence from that species. For each gene, sequences were concatenated to the end of the barcode sequence when they were present; if no sequence existed, an appropriate number of gaps (represented by dashes) were concatenated to the barcode sequence for that species. The same procedure was applied to the genomic genes, again using the DNA barcode sequences as an anchor so that the two sets of sequences would be easily comparable. I then constructed a Minimum Evolution phylogeny with 100 bootstrap replicates from the concatenated set of mitochondrial genes using the Maximum Composite Likelihood metric and pairwise deletion of gaps in MEGA4. Using the same parameters, I constructed a phylogeny from the concatenated set of genomic genes. For both the genomic and mitochondrial phylogenies, I identified monophyletic groups of three or more species with bootstrap support greater than 75% (Jönsson and Fjeldså 2006) and selected those species groups that were consistently identified by both phylogenies.

## **Results**

### *Data collection and processing*

All publicly available genetic sequences for Passeriformes species for genes with greater than 100 samples were compiled by gene and processed to obtain a single consensus

sequence for each species. These consensus sequences were concatenated for 1,212 species. For the mitochondrial genes, the concatenated aligned sequences of seven genes plus the DNA barcode sequences were 12,323 nucleotides in length; for the twelve genomic genes, the concatenated sequences plus the DNA barcode sequences contained 17,482 base pairs.

### *Phylogeny*

For the genomic and mitochondrial phylogenies, I selected monophyletic groups of three or more species with bootstrap support greater than 75%. Most of these groups contained members of the same genus; occasionally multiple genera were represented in one closely related branch. I compared the results from the two phylogenies and selected clades that were well supported in both the genomic and mitochondrial phylogenies. The vast majority of these branches had bootstrap support greater than 98%. There were 281 species in 41 clades meeting these criteria (Figure 3-1) out of the original 1,212 species. These were used in the subsequent song analysis when song data were available.

### **Discussion**

By compiling and analyzing existing avian genetic data, I constructed one phylogeny of mitochondrial genes and one of genomic genes and then identified well-resolved groups of species by choosing clades that were strongly supported in both phylogenies. For both the mitochondrial and the genomic data sets, I concatenated the gene sequences so that all of the separate samples were treated as one long gene. A concatenated gene phylogeny assumes that the phylogeny for each gene would accurately reconstruct the phylogeny for

the organisms themselves. In large phylogenies, this assumption can be invalid, because the genetic common ancestor of a gene sampled from two organisms was not necessarily present in the common ancestor of the organisms themselves. However, by concatenating numerous genes for this study and finding branches with strong bootstrap support, I placed a stringent constraint on the phylogeny such that closely related groups of species with genes evolving consistently would be most likely to show strong bootstrap support. It was possible for a closely related group of species to be only weakly supported in this tree if selection acted differently on different genes. Since my objective was to limit the song analysis to species with the most clearly defined genetic relationships, I used concatenated consensus sequences and chose only those groups that were highly supported by both genomic and mitochondrial data, thus placing a stringent requirement for unambiguous genetic relationships on the species selected for the subsequent song analysis.

In their pioneering study of avian molecular evolution, Sibley and Ahlquist (1990) identified three superfamilies within the Passeriformes, Passeroidea, Sylvioidea, and Muscicapoidea. These three monophyletic groups were supported by the avian “supertree,” which was an amalgamation of previously published phylogenies constructed in different ways. To date, there is no large-scale genetic analysis of the relationships between avian species. Recent phylogenies have attempted to clarify higher-order relationships by sampling one species from most families (Barker *et al.* 2004, Hackett *et al.* 2008). Even in light of these ambitious studies, there is still a lack of strong evidence supporting most of the major clades in the Aves order (Edwards *et al.* 2005).

Of the three superfamilies in the Passeriformes, the Passeroidea was best represented in this data set (Figure 3-1). Compared to the avian “supertree,” the overall structure of this phylogeny has several differences. The Passeroidea was the largest monophyletic group on the phylogeny, with only one aberrant genus present, the Polioptila, which is usually classified in the superfamily Muscicapoeida. The suboscine groups were outgroups of the phylogeny I generated, which is consistent with the avian supertree. The other two Passeriformes superfamilies, however, were not monophyletic in this tree; some members of the Sylvioidea were nested inside the Muscicapoeida branch, and one member of the Muscicapoeida (*Sitta*) was an outgroup of the rest. Surprisingly, one Corvida group (*Vireo*) consistently grouped inside the Passeriformes branch. However, it is not clear how much emphasis should be placed on these differences. With the advent of more precise molecular genetics, the Sibley and Ahlquist phylogeny, based on DNA-DNA hybridization, is no longer the accepted baseline for avian taxonomy. The avian supertree is controversial, as are all supertrees, because it was constructed from a large number (97) of smaller morphological and genetic trees without a full statistical regard for the uncertainties in each tree (Jönsson and Fjeldskå 2006, Ren *et al.* 2009). In addition, this supertree was constructed by hand instead of with an algorithm, and from the component trees themselves instead of the underlying data, so there is still a gap in the literature for a large and unbiased molecular genetic phylogeny. Finally, the method of phylogenetic construction I used here was most useful for identifying groups of species with very clear genetic relationships, which was better suited to the purpose of choosing species for a song analysis than to reconstructing the relationships between families. To verify these

differences, in future work I will perform a more detailed phylogenetic analysis without concatenation using Bayesian methods, which do not make the same assumption that all genes are evolving along the same tree as the species themselves.

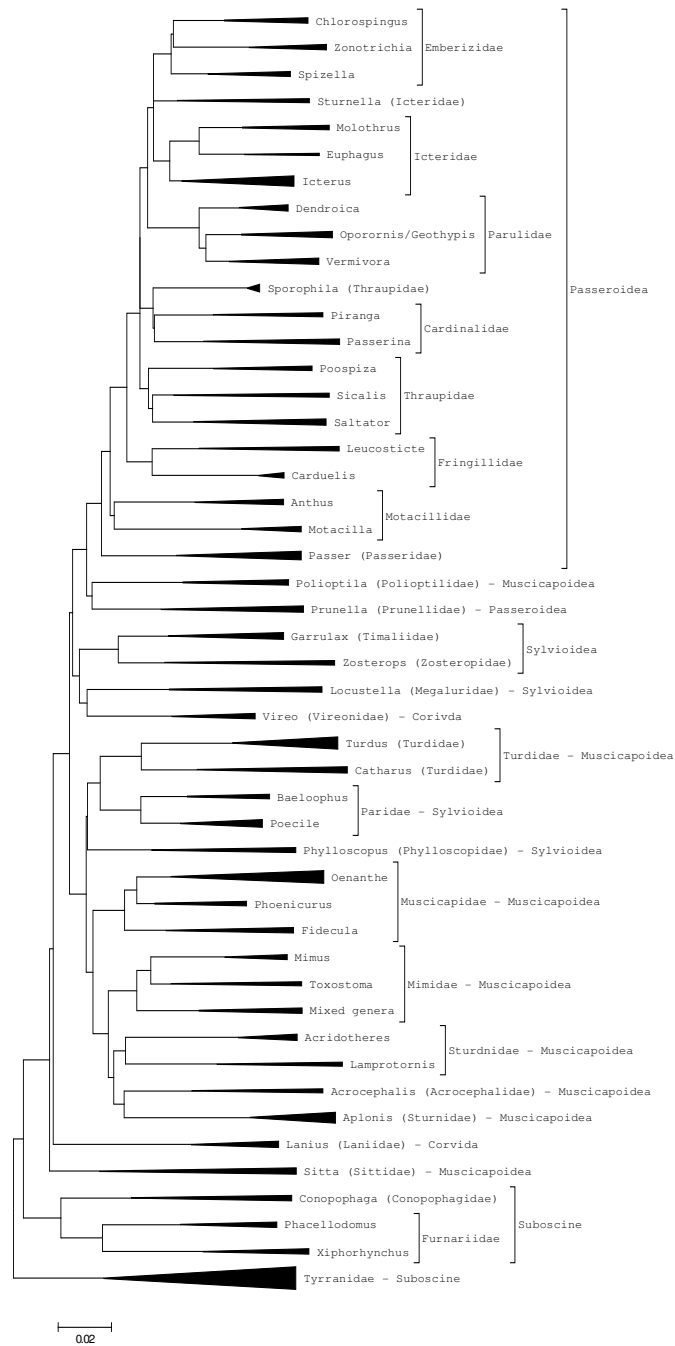


Figure 3-1. Phylogenetic tree topology of birds included in the song analysis. Species are consolidated by genus clades, which are represented by black triangles that correspond in height to the number of species in the group. Genera are labeled by family and superfamily. The scale bar represents average number of base pair changes per site.

## Part 4. Interfamily song analysis

By analyzing recordings of the learned songs of Emberizid sparrows, I demonstrated that both the phonology of song – the sound properties of the syllables – and the syntax of song – the pattern of the syllables – can reasonably predict the genetic relationships between a set of species belonging to three distinct genera. This finding challenged the historical assumption that song behavior can be useful for classifying birds into species but evolves too quickly to be practical when inferring higher-level relationships between species. However, my initial results applied only to one family of birds. Whether an analysis of songs can provide evolutionary information in other families, and whether similarities between closely related families could be observed, remained an open question.

To facilitate this inquiry, I compiled a large portion of the available genetic data for the Passeriformes and used it to find a set of species with clear and well-supported genetic relationships, as described in the previous section. This set of species would allow me to test on a wider set of birds the hypothesis that song can carry information about the evolutionary history of species. Furthermore, there were several additional questions I could ask given a data set of this size and diversity. I examined whether song carried any information about the higher-order species relationships, such as between families or superfamilies. I gathered data for three families of suboscines, which do not learn their song, and I observed the levels of variability within and between these species in comparison to the vocal learners. If song is innate, one might expect less variability overall,



but this prediction has not previously been tested. In addition, I investigated whether the song of the suboscines reflected their ancestral position in the Passeriformes phylogeny.

## Methods

### *Acquiring files*

I selected 48 well-supported genetic groups for song analysis, mostly corresponding to groups of species in the same genus, and then I chose corresponding files from the Cornell Lab of Ornithology. When a species identified by the phylogenetic analysis was present in Cornell's database, I chose as many files as were digitized unless there were more than 35, in which case I preferentially chose 35 of the available recordings based on the quality rating of the files. When possible, I selected files that were recorded in different years and locations to reduce the chance of including two recordings of the same bird. Once I obtained the set of recordings, I first listened to each one and noted the quality of the recording, the presence of song, and an initial assessment of whether successive song bouts were the same or variable. Next, I eliminated phylogenetic branches that contained fewer than three species with multiple song recordings (19 branches). In addition, for this analysis I eliminated species that did not have an apparent motif structure (12 branches). For the remaining 17 branches, I examined each recording and selected song bout excerpts that were free of extraneous bird and human noise. If the motif appeared consistent between bouts, I chose three bouts whenever possible, but if the motifs seemed variable, I selected up to ten bouts to obtain a better average of the syllable properties and syntax.

### *Syllable analysis*

The syllable analysis was conducted for these recordings as explained in Part 2. I filtered low frequency noise and segmented the song into syllables as previously described. However, since this database of recordings was from a diverse set of birds, there were several differences between this data set and the subset of Emberizidae analyzed previously. The Emberizids are predominantly grassland and forest edge species, so most recordings had low levels of background noise and echoes. Some species in this larger data set have deep forest or jungle habitats, where the amplitude of the background noise was consistently higher. To account for this, I checked the results of the syllable segmentation algorithm manually. Any changes to the syllable segmentation fell into three categories: eliminating noise or echoes that were segmented as a syllable, identifying low amplitude syllables that were fragmented by the amplitude threshold, and ensuring that repetitions of a syllable was segmented consistently with one another.

After the syllables were segmented, I followed the methods described in Part 2 to calculate the syllable phonology and syntax data for the species identified by the phylogenetic analysis. For this analysis, I calculated the average syllable property values per species slightly differently. Since there was not necessarily the same number of bouts analyzed for each bird, I first averaged the syllable property values for each bird and then averaged these values for each species. In this way, individual birds with different numbers of bouts recorded were weighted equally in the species-level analysis.

## Results

For this analysis, I analyzed 2,788 song bouts from 862 individual birds in 85 species from well-supported branches of an avian phylogeny (Figure 4-1). These species were members of 17 distinct families, three of which were suboscines (Figure 4-2). A neighbor-joining tree constructed from syllable properties did not completely recapitulate the genetic tree, but the species were by no means randomly distributed (Figure 4-3). Nearly all of the smaller branches, namely those containing four species or fewer, contained members of the same genus. As I observed in the analysis of Emberizid species, the phonological properties of a bird's song were conserved past the species level of the phylogeny. The larger structure of the phylogeny, for example the suboscines as an outgroup, was not preserved. The same observations held for the neighbor-joining tree built from syntax values: small branches often contained closely related species, but a higher-order structure among families and superfamilies was not readily apparent (Figure 4-4).

A factor analysis of the syllable properties demonstrated that many of the variables were correlated, and there were seven underlying factors in the data (Table 4-1). The first factor grouped seven variables that were related to syllable length and included "Pure tones," which, by personal observation, appeared to be longer than more frequency modulated syllables. The next factor clustered several variables related to the frequency range of the bird, as well as number of syllables divided by bout duration. The third group included notes per syllable, silence duration, mean note frequency range, and parabolic notes. The following factor joined variables related to the song length: duration of bouts, number of notes, number of syllables, and number of unique syllables. The fifth factor

combined variables that were related to the repetitiveness of the song and the frequency range. The sixth factor measured how stereotyped the songs were and the last factor grouped the remaining syllable shapes. A tree constructed from these factors also grouped closely related species together on small branches (Figure 4-5).

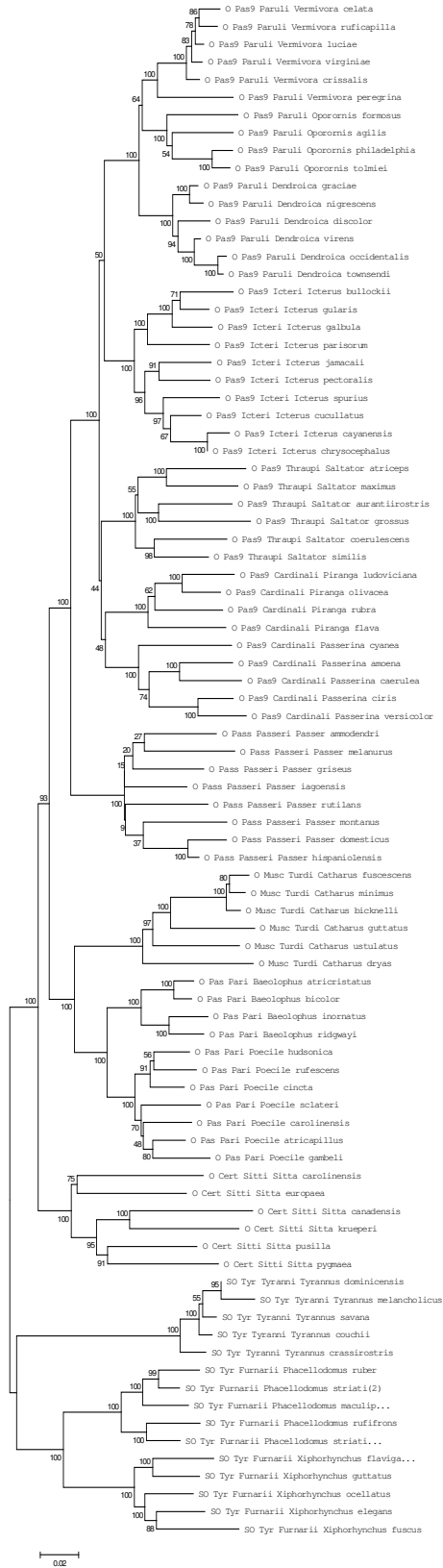
I generated phylogenies using exclusively song data from a variety of avian genera and families and observed that, for trees made from both phonology and syntax variables, genetically related species tended to group together on the most closely related branches of the tree. Correspondingly, I found a highly significant correlation between the genetic distance between two species and their “song distance,” calculated as the Euclidean distance between sets of normalized song variables for a pair of species. With a Pearson’s correlation coefficient ( $R=0.127$ ), the significance of the correlation was  $p=3.5 \times 10^{-13}$ . Using Spearman’s rank correlation test ( $Rho=0.218$ ), the significance of the correlation could not be distinguished from zero using MATLAB, indicating an exceedingly low p-value and thus a very significant relationship: across the Passeriformes phylogeny, as genetic distance increases, so does song distance.

## **Discussion**

By analyzing learned and unlearned songs from 85 species across 17 well-resolved genetic branches (families) of the avian phylogeny, I demonstrated a robust correlation between genetic distance and song distance. I reconstructed a neighbor-joining tree from both the syllable properties of the songs and the song syntax. Both of these trees grouped closely related species in the terminal branches of the phylogeny, but these groups did not

then cluster according to the family and superfamily structure of the genetic phylogeny. The syllable properties were reduced to seven principal factors, which accounted for 82% of the variance in the data. It might be possible to reconstruct the deeper branches of the phylogeny with a different set of characters, and I would like to add more variables to the analysis. Additionally, the variability of syllable property values for individual birds within a species could provide valuable information that I lost by averaging these values among members of a species. Likewise, with a species set this large, it could be beneficial to conduct a song syntax analysis in which the elements of the syntax pattern have a relationship to the syllable shapes that they represent. Many of the birds sampled sang variations on a few patterns, ABC, AAAAA, BAAAAC, etc., but some characteristics of “A” were preserved among closely related species but differed between genera. If I used syllable syntax letters that carried information about syllable properties, I could potentially better discriminate between divergent groups. On the other hand, it is possible that families that diverged 50 million years ago cannot be reasonably reconstructed by a learned behavior. With this analysis, I confirmed a significant correlation between genetic distance and song distance, which stands in contrast to previous hypotheses that learned song carries little evolutionary information above the species level.

Figure 4-1. Species selected for song analysis. The phylogeny of 1,212 bird species was culled to the subset of 85 species shown on this tree based on 75% bootstrap support for the clades, consistency between mitochondrial and genomic data, and available song recordings. The scale bar represents average number of base pair changes per site.



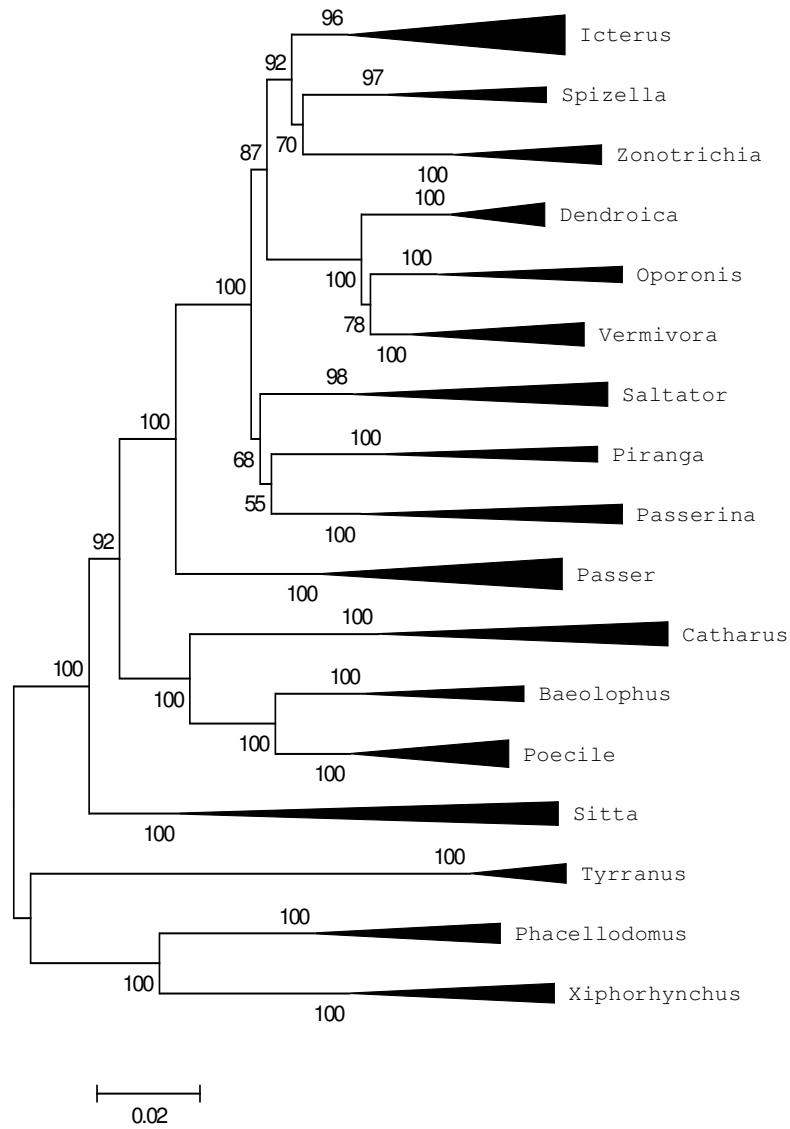


Figure 4-2. Genera included in the song analysis. The phylogeny in Figure 4-1 was compressed to illustrate the relationships between the 17 genera. Each genus is represented by a black triangle proportional in height to the number of species. Branches are well-resolved, and there are several levels of relatedness represented: species, genus, family, superfamily, and suborder. The scale bar represents average number of base pair changes per site.



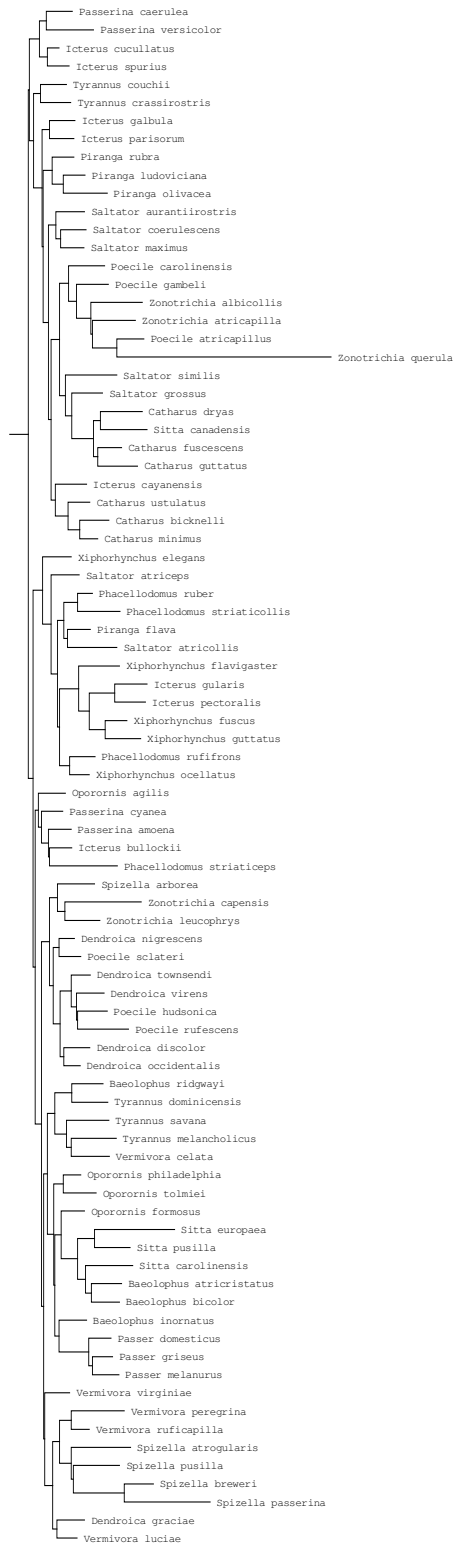


Figure 4-3. Song phonology tree. The most closely related song branches are very often genetically close as well, but the larger structure of the tree does not match well with the Passeriformes phylogeny.

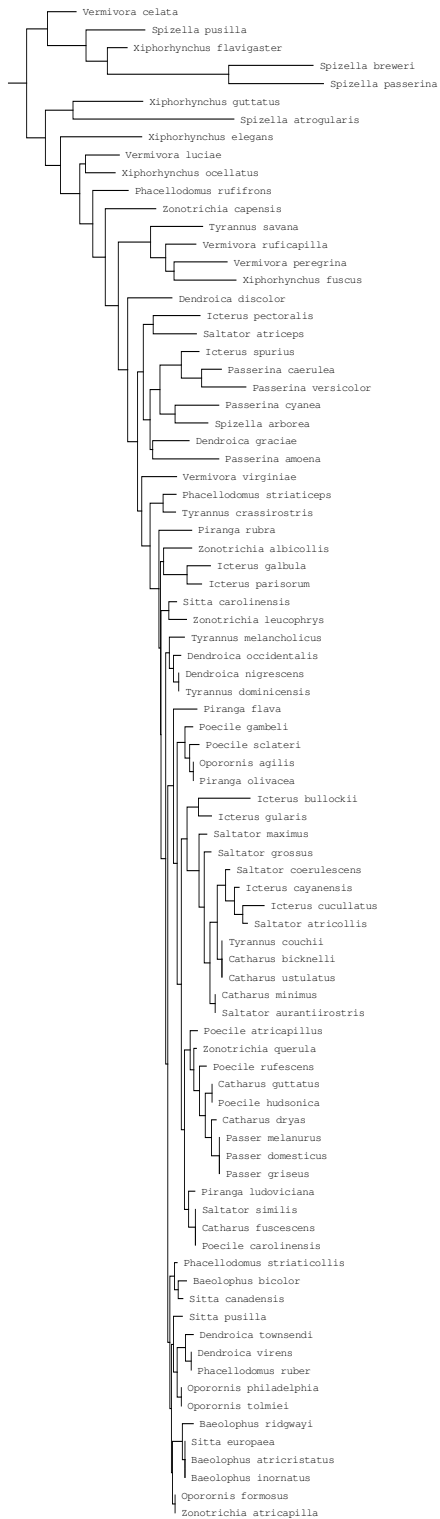


Figure 4-4. Song syntax tree. As in Figure 4-3, the very closely related song branches are very often genetically close as well, but the larger structure of the tree does not match well with the Passeriformes phylogeny.

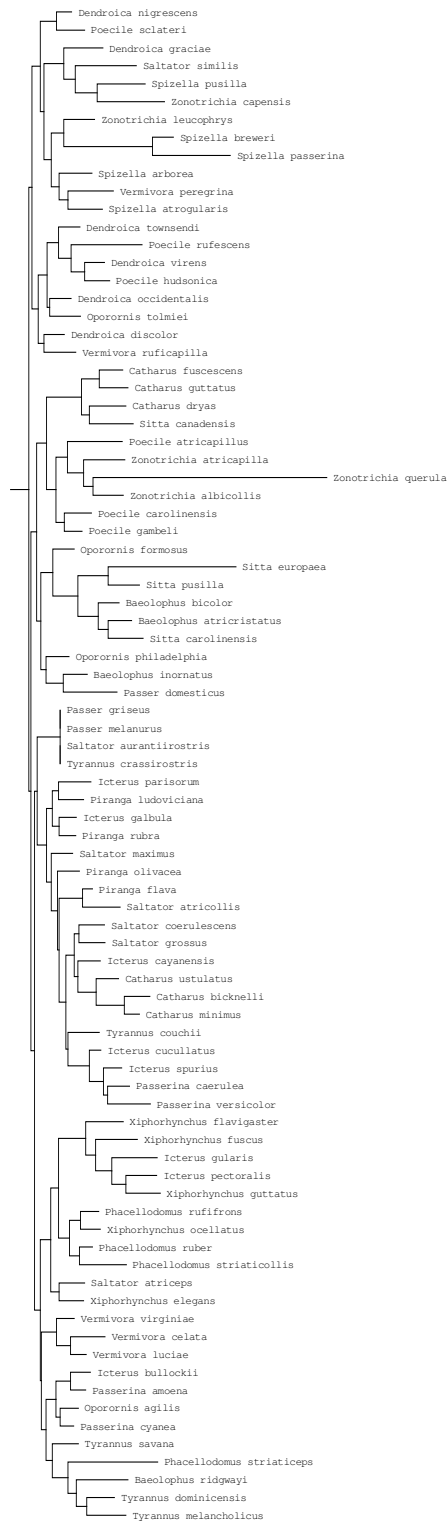


Figure 4-5. Neighbor-joining tree reconstructed from seven song factors identified by factor analysis, which explain 82% of the variance in the data. Once again, the smallest monophyletic song branches are very often closely related species.

|   | Component |       |       |       |       |       |       |
|---|-----------|-------|-------|-------|-------|-------|-------|
|   | 1         | 2     | 3     | 4     | 5     | 6     | 7     |
| Mean note length  | .948      | .018  | .160  | -.137 | .112  | -.057 | .038  |
| Length of longest note                                    | .932      | .008  | -.124 | -.064 | -.044 | .080  | -.154 |
| Length of shortest note                                   | .860      | .036  | .177  | -.085 | .180  | -.088 | .211  |
| Length of longest syllable                                | .826      | .045  | -.378 | -.114 | -.162 | .150  | -.111 |
| Mean syllable durations                                   | .778      | -.091 | -.440 | -.299 | -.111 | .025  | .024  |
| Length of shortest syllable                               | .613      | -.195 | -.371 | -.435 | .080  | -.158 | .126  |
| Pure tones  | .601      | -.300 | -.536 | .142  | -.043 | -.270 | .112  |
| Average frequency trough                                  | -.001     | .946  | .114  | -.092 | -.107 | -.009 | -.016 |
| Average frequency peak                                    | -.094     | .925  | .266  | -.080 | -.090 | .020  | -.047 |
| Minimum frequency   | .222      | .825  | .132  | -.176 | .287  | .069  | -.133 |
| Number of syllables divided by bout duration              | -.423     | .439  | .402  | .254  | .412  | -.134 | .108  |
| Notes per syllable  | -.159     | -.229 | -.821 | .029  | -.268 | -.094 | .142  |
| Mean silence duration                                     | .162      | -.388 | -.646 | -.322 | -.260 | .037  | .052  |
| Mean note frequency range                                 | -.336     | .431  | .604  | -.006 | .003  | .101  | -.121 |
| Parabolas   | -.433     | -.003 | .512  | -.226 | -.229 | .101  | .262  |
| Duration of bout  | .073      | -.312 | -.057 | .804  | .015  | .052  | -.063 |
| Number of notes   | -.265     | -.237 | -.299 | .784  | .059  | -.130 | .080  |
| Number of syllables                                       | -.284     | .269  | .247  | .722  | .395  | -.061 | -.019 |
| Number of unique syllables                                | -.223     | .050  | .359  | .704  | -.348 | .227  | -.060 |
| Number of syllables divided by number of unique syllables | -.188     | .155  | .090  | .171  | .843  | .025  | .131  |
| Sequential repetition                                     | -.027     | -.105 | .316  | .084  | .741  | -.087 | .126  |
| Frequency range   | -.427     | -.030 | .083  | .271  | -.707 | -.082 | .252  |
| Maximum frequency   | -.347     | .502  | .181  | .197  | -.628 | -.050 | .203  |
| Stereotypy  | .395      | .290  | -.020 | -.232 | .066  | .566  | -.168 |
| Downsweeps  | -.193     | .235  | .250  | .036  | .012  | .113  | -.824 |
| Upsweeps  | -.431     | .263  | .171  | -.077 | .274  | .249  | .534  |

Table 4-1. Grouping of the syllable properties by factor analysis for the phylogenetic song analysis. The syllable property values grouped into seven distinct underlying factors of correlated variables. These factors correspond to variables related to syllable length, syllable frequency, etc.

## Part 5. Cultural evolution of song in chipping sparrows

The computational methods for analyzing and comparing the songs of diverse species of birds, as described in the previous sections, can be readily applied to an in-depth analysis of one species. In the oscine songbirds, which usually learn their song from a conspecific tutor, song differences within a species have been demonstrated in several cases, corresponding to various factors depending on the species studied, including location (Marler 1962), altitude (Nottebohm 1969), time of day (Liu and Kroodsma 2007), ambient noise (Luther and Baptista 2009), and number of years since recording (Derryberry 2007). Marler and Tamura (1962, 1964) demonstrated that male white-crowned sparrows (*Zonotrichia leucophrys*) displayed local “dialects,” with birds of one locality producing songs that are much more similar to one another than to conspecific birds a few miles away. In white-crowned sparrows sampled within one of these dialect regions, small frequency and timing differences were recently noticed between current and historical recordings; birds in a natural setting respond most aggressively to a playback of the current dialect and less aggressively to both historical recordings and current recordings from a nearby population, demonstrating that birds are sensitive to the song differences that can accumulate over time and between dialects (Derryberry 2007, Derryberry 2010). Another study of the same species demonstrated that, over the course of thirty years of field recordings in the San Francisco Bay area, the minimum frequency of songs has increased, corresponding to an increase in low-frequency urban noise over the same period that could potentially interfere with the portions of the song at lower frequencies (Luther and Baptista 2009). A member

of the same genus, the rufous-collared sparrow (*Zonotrichia capensis*), exhibits a uniform dialect over a large area of the pampas region of Argentina, but numerous dialects are found within relatively short distances when the area contains altitude shifts, with altitudinal transects delineating the boundaries between dialect regions (Nottebohm 1969, Nottebohm and Selander 1972). Another sparrow in the Emberizidae family, the chipping sparrow (*Spizella passerina*), showed variation in its song depending on the time of day. Chipping sparrows have a simple repertoire of one syllable repeated many times, and the number of repetitions per song bout of an individual bird was found to fluctuate depending on whether the bird was singing at dawn or during the day (Liu and Kroodsma 2007).

The chipping sparrow has a wide range across North America, and it has been recorded extensively for more than six decades. Chipping sparrows can learn a song from their father as juveniles, but they continue to show plasticity after their first spring migration, when they must establish a territory for the breeding season. Male chipping sparrows in their first breeding season can initially produce several song types before settling on one syllable to make up their repertoire, and the birds tend to sing the syllable type that is closest to the syllable of a bird in a neighboring territory, which can be modified to achieve a very close imitation of the song of the neighbor (Liu and Nottebohm 2007). The syllables of 157 chipping sparrows from the eastern United States were analyzed and placed into around thirty distinct syllable type categories by eye (Liu 2001). Since the birds can be influenced by their neighbors, syllable types of successful birds, those with large or long-held territories, could gain prominence in one region; this pattern of song

transfer or random drift of song types could both lead to regional and chronological differences in syllable types. By examining the syllables of chipping sparrows with the time and location of the recordings in mind, it is possible to search for cultural evolution of song within this species. The aforementioned examples of cultural evolution of song within sparrow species relied on extensive and prolonged fieldwork within a subset of the species' geographical range. The analysis reported here relies on existing databases of recordings, which can provide a large number of samples and an extensive range of dates and places with a smaller initial time investment than a long-term field study. However, a computational assessment of recordings taken out of their natural context might lack the resolution to detect nuances and patterns that could stem from prolonged observations of known individuals.

## **Methods**

### *Gathering chipping sparrow data*

To perform the most comprehensive analysis possible, chipping sparrow recordings were collected from all known publicly available databases: Cornell Lab of Ornithology (72 song files from unique individuals), Borror Laboratory of Bioacoustics (350 unique song files), and Xeno-Canto (11 unique song files). In addition, Wan-chun Liu generously donated the audio data from an intensive field study conducted between 1997 and 2008 (233 unique song files). The recordings were manually examined for multiple instances of the same individual: if samples from the same location and roughly the same date

contained songs that appeared identical, all but one sample was eliminated. This method potentially eliminated neighboring pairs of individuals in which one bird had accurately imitated another; however, in the absence of observable song nuances to discriminate between individuals or field notes dictating that two individuals were recorded, seemingly identical song samples recorded in a single location within one field season were considered duplicate recordings of one bird and eliminated for stringency. Songs were also eliminated from the analysis if the background noise was abnormally high relative to the song amplitude.

For each song used in this analysis, we compiled a matrix of attributes from the information provided in the recording databases: day of recording, year of recording, time of recording, latitude of recording location, and longitude of recording location.

#### *Analysis of full songs*

To analyze chipping sparrow song properties, we collected recordings for which an entire song bout was free of interfering birdsong and excessive noise. We used a highpass filter to remove low frequency noise (below 1 kHz) and exported the syllable selection in wav format sampled at 44,100 Hz, as described in Part 2. These recordings were segmented into syllables using software I developed for this purpose in MATLAB. To ensure that the syllables in each excerpt were segmented consistently with one another, the automated segmentation for each song was manually checked for accuracy by two independent researchers. The recordings, along with the corresponding values for syllable onsets and offsets, were analyzed in MATLAB, and the mean values of the following variables were



calculated for each song: mean syllable duration, length of the longest syllable, length of the shortest syllable, mean silence duration, duration of song bout, number of syllables divided by bout duration, total number of syllables, number of unique syllables, number of syllables divided by unique syllables, degree of sequential repetition, mean stereotypy of repeated syllables, mean note length, number of notes, length of shortest note, length of longest note, number of notes per syllable, mean syllable frequency peak, mean syllable frequency trough, song frequency peak, song frequency trough, frequency range, mean frequency modulation, pure tones, upsweeps, downsweeps, parabolas. A thorough explanation of each of these variables can be found in Part 2.

#### *Bout analysis*

After the analysis of chipping sparrow song bouts, we probed the resulting syllable properties for significant relationships to attributes of the recordings, specifically latitude, longitude, calendar day, time of day, and year. To look for song changes over time, we divided the recording years into time periods – early (1948 to 1969), middle (1970 to 1989), and late (1990 to 2010) – and compared each unique pair of time periods for each of the syllable properties. For these comparisons, all t-tests performed were two-tailed Student's t-tests for independent samples with unequal variance. For a given syllable property, we tested the null hypothesis that the mean value for that syllable property was equal for a pair of time periods; we compared the values of recordings from the early time period versus the middle time period, early versus late, and middle versus late. To determine whether the songs varied by location, we divided the longitude into region categories – western ( $-129^{\circ}$  to

-110°), central (-109° to -90°), and eastern (-89° to -69°) – and performed a Student’s t-test between each pair of region categories for each of the syllable properties. We divided latitude into two categories, northern (40° to 49°) and southern (30° to 39°). To find seasonal effects on chipping sparrow song independent of the year of recording, we separated the data into four seasons of recording – March 22<sup>nd</sup> to April 30<sup>th</sup>, May 1<sup>st</sup> to May 31<sup>st</sup>, June 1<sup>st</sup> to June 30<sup>th</sup>, and July 1<sup>st</sup> to August 8<sup>th</sup>. To explore the relationship between time of day and chipping sparrow song, we divided the songs into two categories. First, we divided the song into dawn song and day song, defined as the period beginning one hour before sunrise and lasting approximately 30 minutes (Liu, personal communication). The time of sunrise was approximated for each month, and any songs recorded more than 30 minutes before that time were considered dawn song. In addition, since the vast majority of the data were recorded before noon, we split the morning into four-hour increments: 4:00 to 7:59AM, and 8:00 to 11:59AM. A Student’s t-test was performed between pairs of season categories and pairs of time categories. For all t-tests in this experiment, we used a Bonferroni correction to stringently compensate for testing multiple hypotheses on the same set of data by dividing the 0.01 p-value threshold for significance by the number of tests. We performed 15 t-tests on each of 27 syllable properties, so the threshold for significance was lowered to  $2.5 \times 10^{-5}$ . Using SPSS, we looked for significant linear correlations between the recording attributes (calendar day, time of day, year, longitude, and latitude) and the syllable properties.

### *Heat maps*

For each song, the latitude and longitude of the recording location were rounded to the nearest degree. A plot was created with longitude on the x-axis and latitude on the y-axis. To visualize the variability in syllable property variables (syllable duration, frequency range, etc.) by location, the value of a variable was placed on the plot according to its latitude and longitude. When a given latitude/longitude point had more than one entry, the corresponding syllable property values were averaged at that point. The number of samples at each point was recorded in a separate matrix. The values for each variable were then visualized as a heat map by location, with a color scale representing the numerical values and the x-y position representing the recording location to the nearest degree of latitude and longitude.

### *Song analysis of chipping sparrow excerpts*

A chipping sparrow's song consists of one syllable repeated numerous times in rapid succession, often increasing in amplitude at the beginning of the song and decreasing at the end. This shift in amplitude can make comparisons of the syllables at the edges of the song bout to those at the center of the bout less informative than comparisons of adjacent syllables. To determine whether the amplitude fluctuations were significantly influencing the chipping sparrow syllable properties, we performed an analysis on three-syllable excerpts of each song in addition to the previous analysis on entire song bouts. We limited our analysis to excerpts in which a single chipping sparrow was singing, without simultaneous interference from other birds or human voices; any competing signals

complicate syllable segmentation and comparison. By performing an analysis on several syllables, we were able to include in this analysis fifty additional songs for which no complete bouts were suitable for the previous motif analysis. In addition, we could compare the results of this analysis to the syllable properties of the full bouts to determine whether significant associations between syllable properties and recording attributes were affected by the inclusion of the lower amplitude syllables at the beginning and the end of the bout. Each syllable comprises the entirety of the individual's adult repertoire, so an analysis of a subset of syllables should provide valuable information about individual syllable properties for each chipping sparrow, but bout-level information (number of syllables per bout, bout duration, etc.) must be obtained from the previous analysis.

For this analysis, we took an excerpt of three syllables from the middle of a song bout from each recording with low levels of background song and noise. As above, we selected an excerpt from each unique individual and filtered out the low frequency noise (below 1 kHz). Each file was then segmented into syllables using MATLAB software developed for this purpose and checked for errors by two individuals, as described in previous sections.

Once an excerpt from each recording was compiled, we determined the syllable property variables for each recording as described previously. Some of these variables became redundant because we chose three-syllable excerpts, so the number of syllables was artificially held constant, but these values were checked for consistency with expectation.

We examined the results of the analysis of chipping sparrow song excerpts for relationships between syllable properties and attributes of the recordings, specifically

latitude, longitude, and year. Since chipping sparrows, with rare exceptions, sing one syllable for their adult lives, we did not examine the syllable properties of these excerpts for relationships to the time or calendar day of recording. We divided the recording years, longitude, and latitude into categories as described for the chipping sparrow bout analysis and performed a Student's t-test between each pair of categories for each of the syllable properties. Using SPSS, we looked for significant linear correlations between the year, longitude, and latitude attributes and the syllable properties.

### *Syllable types*

To detect patterns of transmission in chipping sparrow syllables, we first determined which birds were singing the same songs. Using the data collected for the analysis of 667 chipping sparrow song excerpts, we constructed a database of one representative syllable for each bird. Using MATLAB, we compared each possible pair of syllables using the syllable comparison algorithm described in Part 2 and recorded the similarity score in a matrix. To group the syllables in a chipping sparrow population into unique syllable clusters, we needed to reconsider the threshold for syllable similarity that should be used to classify two syllables as the same. In the syllable syntax analysis described previously, the algorithm was used to determine whether a given syllable could be considered a repetition of another syllable within the same bird's song; the syllables being compared were recorded under identical or near-identical conditions and performed by the same individual, but some modulation of the syllable was tolerated. For the chipping sparrow population analysis, we compared recordings varying widely in song amplitude and background noise, so we used a

thresholded version of the spectrogram for each syllable such that the background was set to zero and the syllable itself set to one. In addition, we sought a similarity threshold that was optimized for chipping sparrow syllables: lenient enough to include all occurrences of the same syllable in a single cluster but stringent enough to restrict the inclusion of dissimilar syllables. After testing, we used a 62% similarity threshold to balance these two concerns. Elements of the 667 by 667 matrix of pairwise similarity scores were set at zero if they were less than 62%, and diagonal elements representing comparisons of a given bird with itself were also set to zero so that the matrix only contained passing scores for syllable comparisons from two different chipping sparrows. In addition, the difference in syllable length was calculated for each pair of syllables. To prevent a short syllable from receiving a passing score by overlapping with a portion of a larger syllable, comparisons between syllables that differed in length by more than 25% of the length of the shorter syllable were set to zero. Each column of the resulting matrix represented the comparisons between one bird's syllable and every other bird's; if all elements in a column were zero, the syllable was placed in a group by itself. Otherwise, the syllable was grouped with the syllable corresponding to the largest of the similarity values in the column. This procedure of grouping a syllable with its best match was continued for all columns in the matrix, and if one member of the syllable pair in question was already a member of a group, the other syllable in the pair was added to that group. It was possible by this method for a subset of a larger group of apparently identical syllables to form a separate cluster if the elements of the subset matched best to each other. To fuse groups that contained the same syllable, we calculated the mean similarity score between elements of each group to every other group

(with similarity scores below the threshold and self-similarity scores again set to zero) and tested for a threshold that joined groups that were similar but kept dissimilar groups separate. We used a 23% overlap to fuse the groups, which required more than a third of between-group syllable comparisons to exceed the 62% threshold for syllable similarity.

#### *Cultural analysis by year and place*

For each of the 667 chipping sparrow recordings used in the previous analysis, the syllable group categories were annotated with the year, longitude, and latitude of the recording. We divided the year category into decades, and for each syllable group with more than five members we calculated the fraction of the syllables in each decade made up by members of that group. Similarly, we divided the longitude values into four categories, two western ( $-135^{\circ}$  to  $-116^{\circ}$  and  $-115^{\circ}$  to  $-96^{\circ}$ ) and two eastern ( $-95^{\circ}$  to  $-76^{\circ}$  and  $-75^{\circ}$  to  $56^{\circ}$ ), and calculated the fraction of syllables in each longitude category made up by members of a given group.

## **Results**

#### *Gathering chipping sparrow data*

In total, 667 song files from unique chipping sparrow individuals were compiled for this analysis; 328 additional recordings were screened and excluded due to noise levels or repeated recording of the same bird. All of the analyzed song files were annotated with the

year and location of recording. Information about the exact day of recording and the time of recording was available for 68.6% and 65.4% of recordings, respectively.

#### *Analysis of chipping sparrow bouts*

We examined 617 chipping sparrow bouts for significant associations between syllable properties and attributes of the recordings. The mean of many syllable properties were significantly different between longitudinal categories: mean syllable duration, longest syllable duration, shortest syllable duration, mean silence duration, number of syllables divided by bout duration, number of syllables, mean note length, number of notes, shortest note length, longest note length, mean frequency peak, mean frequency trough, maximum frequency, frequency range, and downsweeps (Table 5-1). Most syllable properties were significantly different for comparisons between both eastern vs. western and eastern vs. central longitudinal categories (see Table 5-1), and no significant differences were found between the central and western categories. For example, the mean syllable duration was significantly higher in the east (Figure 5-1), but the number of syllables was significantly higher in the west (Figure 5-2). Several recording sites are highly represented, though, especially in the east (Millbrook, NY; Amherst, MA; and Columbus, OH). To determine whether the relationship between longitude and syllable properties was still evident when all the samples from a given location were averaged, we generated “heat maps,” on which darker colors represent higher values of the syllable property in question. The longitude and latitude points represented in the song data are illustrated in Figure 5-3. For each of these points, the average syllable property value for that location is displayed on the map.



Plotting the mean syllable duration values in this way clearly suggested that syllables are longer in the east (Figure 5-4), and the opposite pattern was observed with number of syllables (Figure 5-5).

No syllable properties were significantly different between latitude categories. Only duration of bout was significantly different between year categories for the early vs. late comparison. The maximum frequency values were significantly different when the earliest and latest seasonal categories were compared (March and April vs. July and August), but this difference could potentially be attributed to a significant bias in these months in the longitude distribution, which was also correlated to maximum frequency. No variables were significantly different based on the time of day when it was divided into four-hour increments. When the time of recording was divided into “dawn” and “day” periods based on the median sunrise time per month, both the bout duration and the number of syllables were significantly lower in the dawn period.

The syllable properties that were significantly different between recording attribute categories would not necessarily be significantly linearly correlated if the relationship between values in different attribute categories was nonlinear. However, we found most of the significant differences between attribute categories were supported by significant linear correlation. When we considered only relationships significant ( $p < 7.4 \times 10^{-5}$ , a  $p = 0.01$  significance level corrected for  $27 * 5$  tests) with a Pearson correlation coefficient greater than 0.2, we found consistent results for time of year (a correlation with maximum frequency), year, time of day (both significantly correlated with bout duration), and latitude (no significant correlations). The longitude of recording was significantly correlated with all

of the same values from the previous analysis except shortest note length and downsweeps (Table 5-1).

#### *Analysis of excerpts*

The analysis of three-syllable excerpts from chipping sparrow songs was completely consistent with the analysis of full bouts. Since the number of syllables was held constant, variables that did not apply to song excerpts were excluded: bout duration, number of syllables, number of syllables divided by bout duration, and number of notes. All other variables were applicable on a syllable level and were included. For these variables, the results of the analysis of syllable excerpts were identical to the correlation results for the chipping sparrow bouts. Year and latitude did not significantly correlate to any syllable properties and the t-test did not reject the null hypothesis that the mean of any variable was different between year or latitude categories. The longitude of recording was significantly correlated with several variables, all of which were predicted by the bout analysis: mean syllable duration, longest syllable duration, shortest syllable duration, mean silence duration, mean note length, longest note length, mean frequency peak, mean frequency trough, maximum frequency, and frequency range. The Student's t-test results showed the same set of variables to be significantly different between the eastern and western longitude categories. Between the eastern and the central longitude categories, all of the same variables were significantly different except mean note length and longest note length.

#### *Syllable types and cultural analysis*

By comparing one representative syllable from each chipping sparrow song, we classified the 667 syllables into 124 chipping sparrow syllable types. Each of these 667 syllables was labeled with its syllable type, year, and longitude. We then plotted the fraction of each longitude category composed of different syllable types (Figure 5-6) and the same for time categories (Figure 5-7). From these plots, in which higher fractions of a given syllable are represented by darker colors, we identified certain syllable types that were biased toward early versus late recordings and eastern versus western recordings (Figure 5-6 and 5-7, B and C).

## **Discussion**

Although an adult chipping sparrow generally has a single syllable in its repertoire, the number of repetitions of this syllable in a given bout can vary and fluctuates in different contexts; for example, an individual chipping sparrow's dawn song tends to be shorter in duration and delivered more often than its song during the day (Liu 2001). With this analysis, we found that this observation, made by observing banded birds at different times of day and comparing song characteristics within individuals, generalized to a large set of recordings in which each bird was recorded only once. Thus, it seems that each individual chipping sparrow's propensity for shortening its song bout during the dawn chorus cumulatively lowers the average bout duration and number of syllables in a bout during the period ending 30 minutes before sunrise.

In addition to supporting this field observation, we saw many differences when chipping sparrows recorded in the eastern part of the United States and Canada were

compared to those recorded in the western region. For eastern chipping sparrows, the mean syllable duration was significantly higher and the number of syllables significantly lower than in western birds, but the bout duration was indistinguishable between regions. Similar values to syllable duration, such as silence duration and note length, followed the same pattern of higher values in the east, and similar values to number of syllables, such as number of notes and the speed of syllable delivery (number of syllables divided by bout duration), had higher values in the west. This result implied that, on average, western birds are singing bouts of the same length as eastern birds and that the song of the western birds contained more syllables sung at a faster pace. Some frequency properties of chipping sparrow songs also correlated with longitude: most frequency-related variables were higher for the eastern recordings. The majority of tested syllable properties correlated very strongly with longitude, even with a very stringent p-value correction for testing multiple hypotheses on the same data set, whereas between zero and two variables correlated with the other recording attributes.

Our findings suggested that longitude is a salient factor when studying the cultural changes in chipping sparrow songs, but this analysis does not have the resolution to predict whether the east and west populations were evolving independently, whether habitat and body size factors were different in these two populations, or whether sexual selection favored fast-paced and lower frequency songs in the west and slower, higher frequency songs in the east. A set of playback experiments in the field would be well suited to examining this phenomenon further: the aggressiveness of a chipping sparrow's response to playback of typical eastern and western recordings, as well as eastern syllables modified to the speed

or frequency of a western song and vice versa, could test the relative importance to the bird of these significantly different variables. Separately, in the cultural analysis we classified the set of chipping sparrow songs into unique syllable types and found several syllable types that had biased distributions; some were more common in the east versus the west, and others were more common in recent recordings versus older recordings. Field tests using songs of these different syllable types could help determine whether this bias in distribution corresponds to a difference in behavioral response.

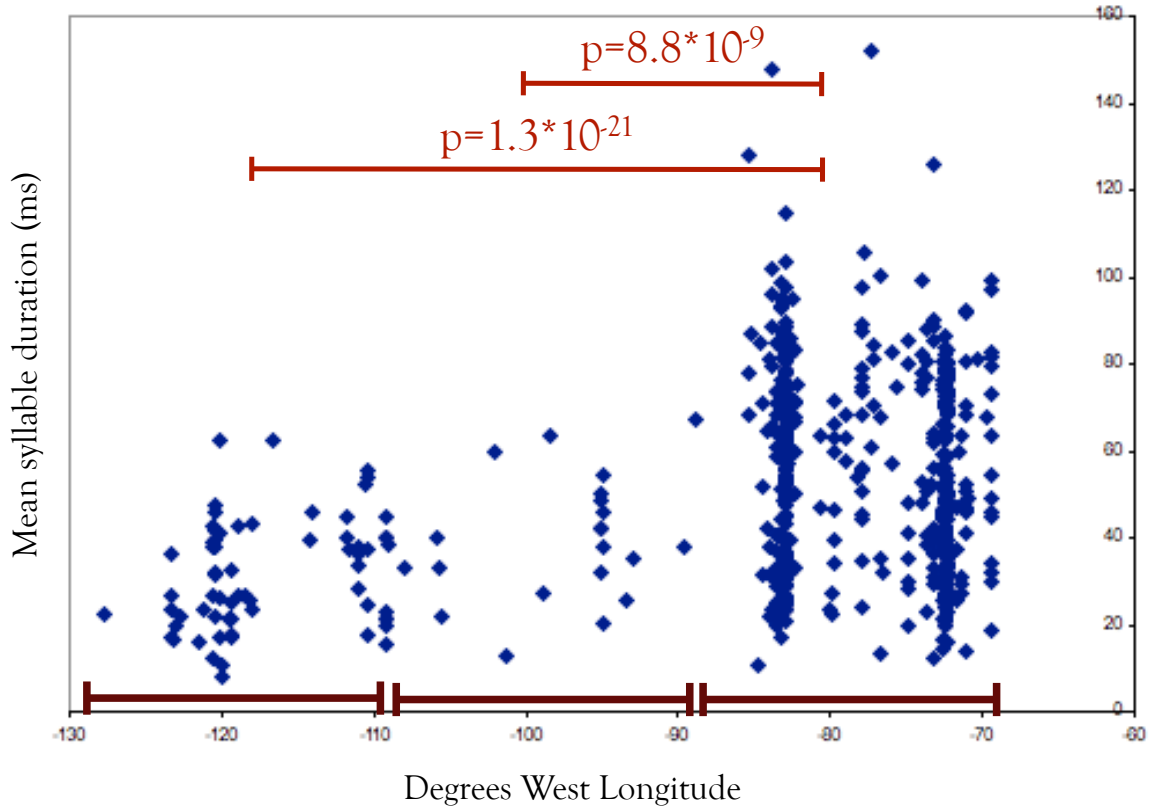


Figure 5-1. Distribution of chipping sparrow syllable duration by longitude. Each point represents the mean syllable duration of one bird's song. Chipping sparrows in the eastern part of North America had significantly higher syllable durations than birds in the west or central regions.

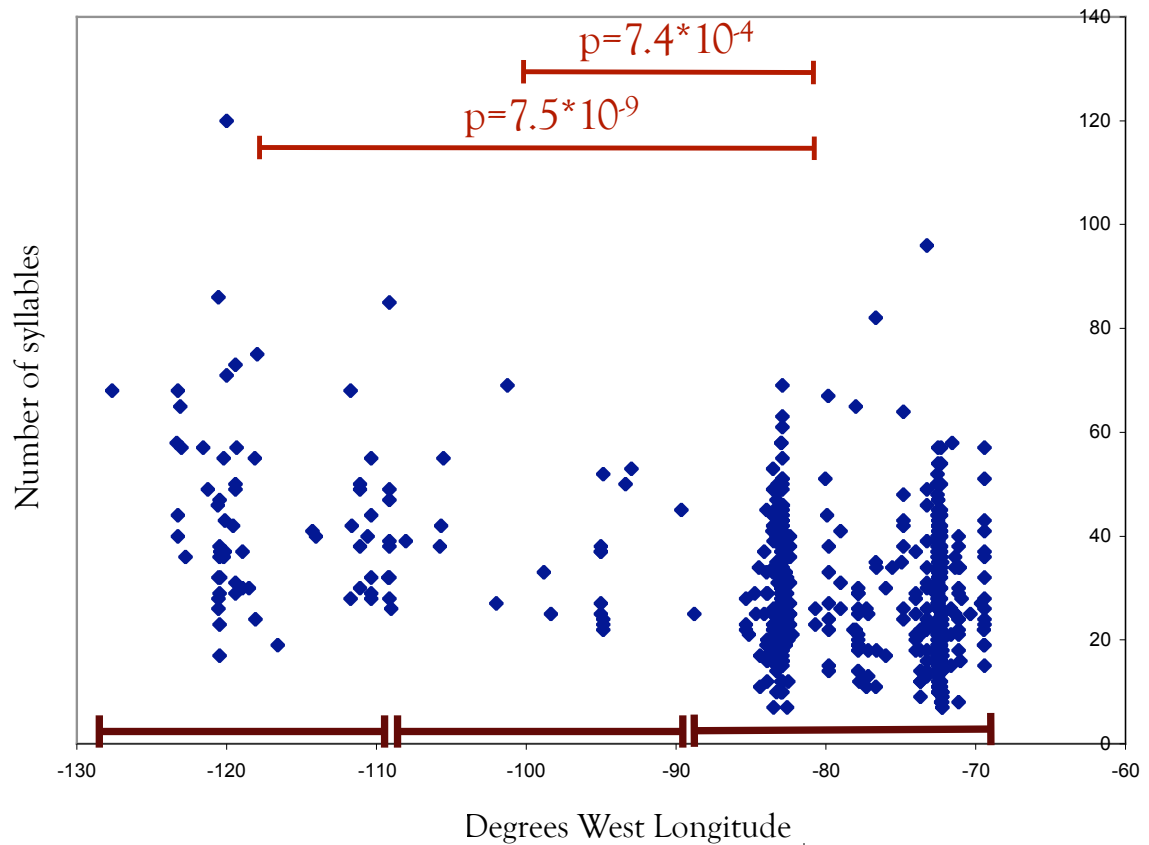


Figure 5-2. Distribution of chipping sparrow number of syllables by longitude. Each point represents the number of syllables in one bird's song. Songs in the western part of North America had significantly more syllables than songs recorded in the central or eastern regions.

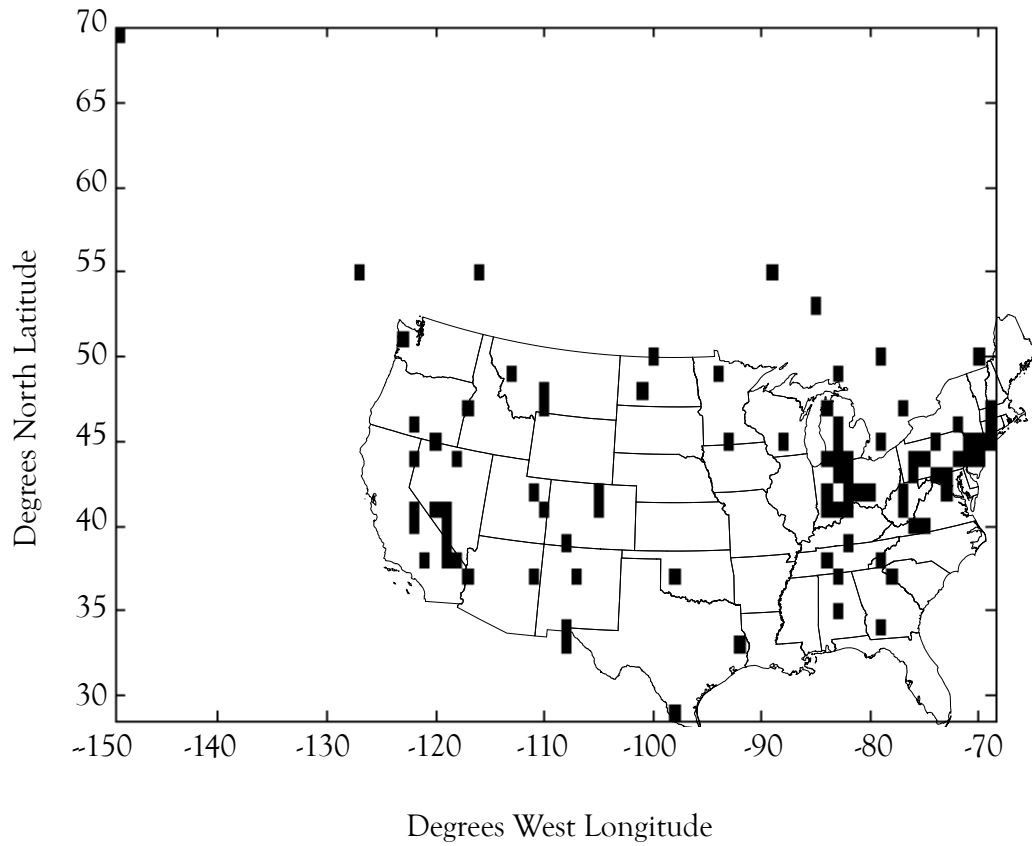


Figure 5-3. Distribution of chipping sparrow data by longitude and latitude. Longitude and latitude values were rounded to the nearest degree and plotted on the map to show the density and spread of sampling



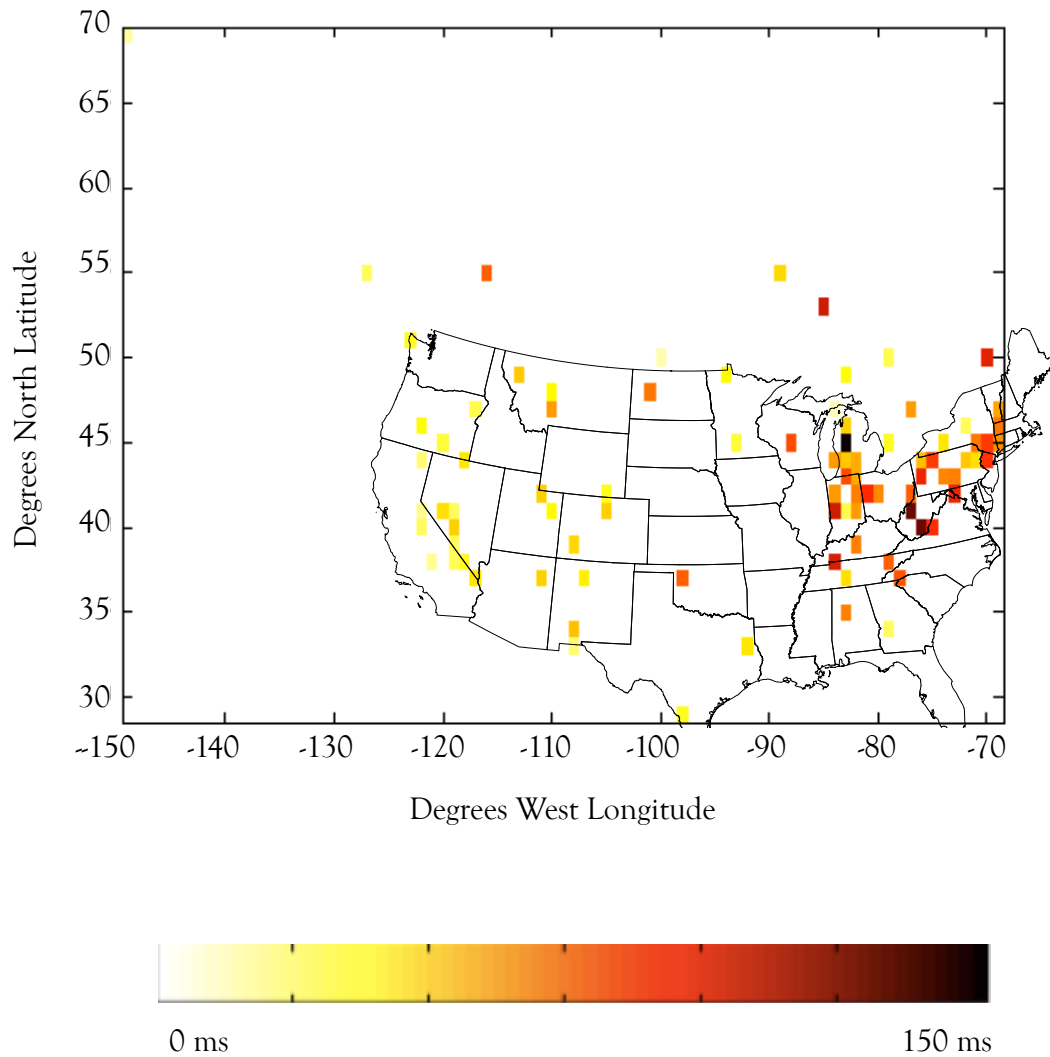


Figure 5-4. Heat map of chipping sparrow syllable durations by latitude and longitude. The location of each recording was rounded to the nearest degree, and the syllable duration values within that area were averaged and plotted on the map. The syllable duration is longer as the colors of the points go from yellow to red to black. The syllables in the western region are significantly shorter than those in the east.

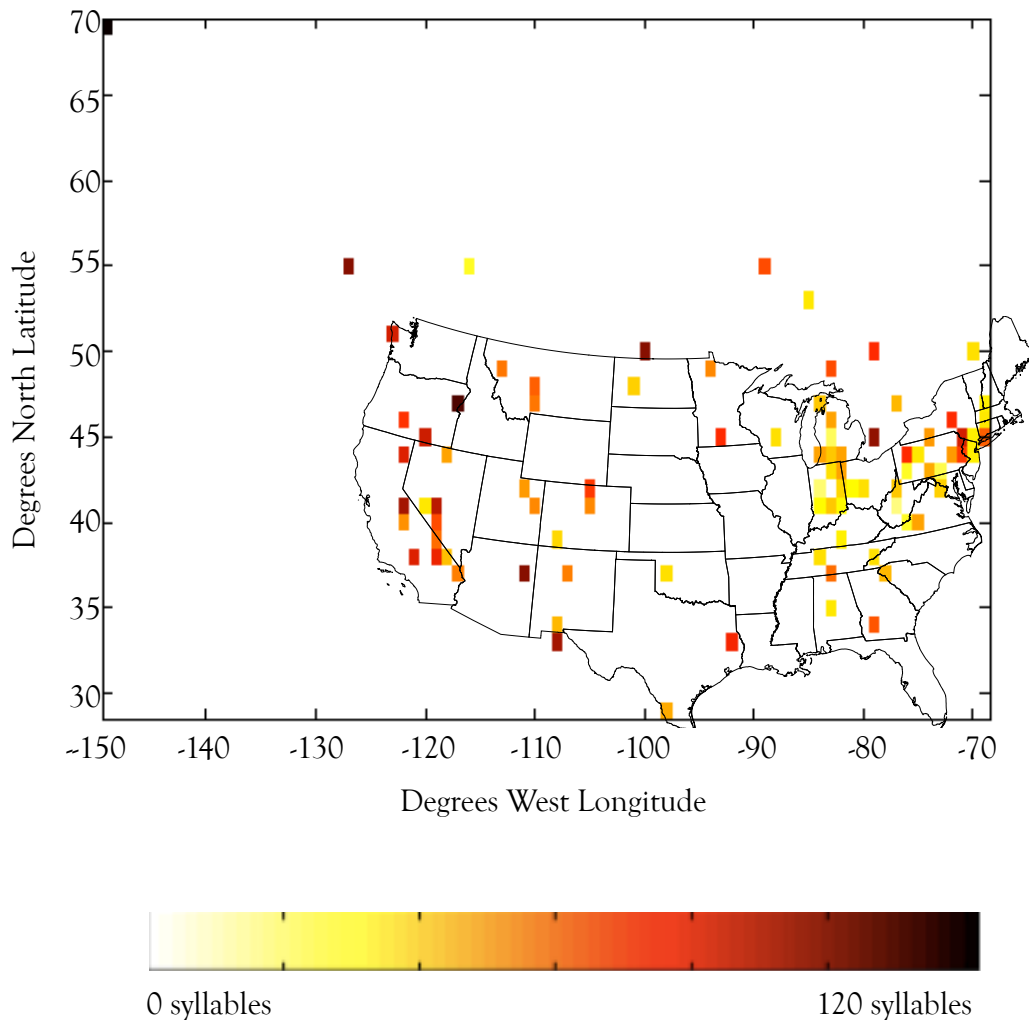


Figure 5-5. Heat map of chipping sparrow number of syllables by latitude and longitude. The location of each recording was rounded to the nearest degree, and the values for number of syllables for recordings within that area were averaged and plotted on the map. The number of syllables is greater as the colors of the points go from yellow to red to black. The number of syllables in a chipping sparrow song was significantly lower in the eastern region of North America.

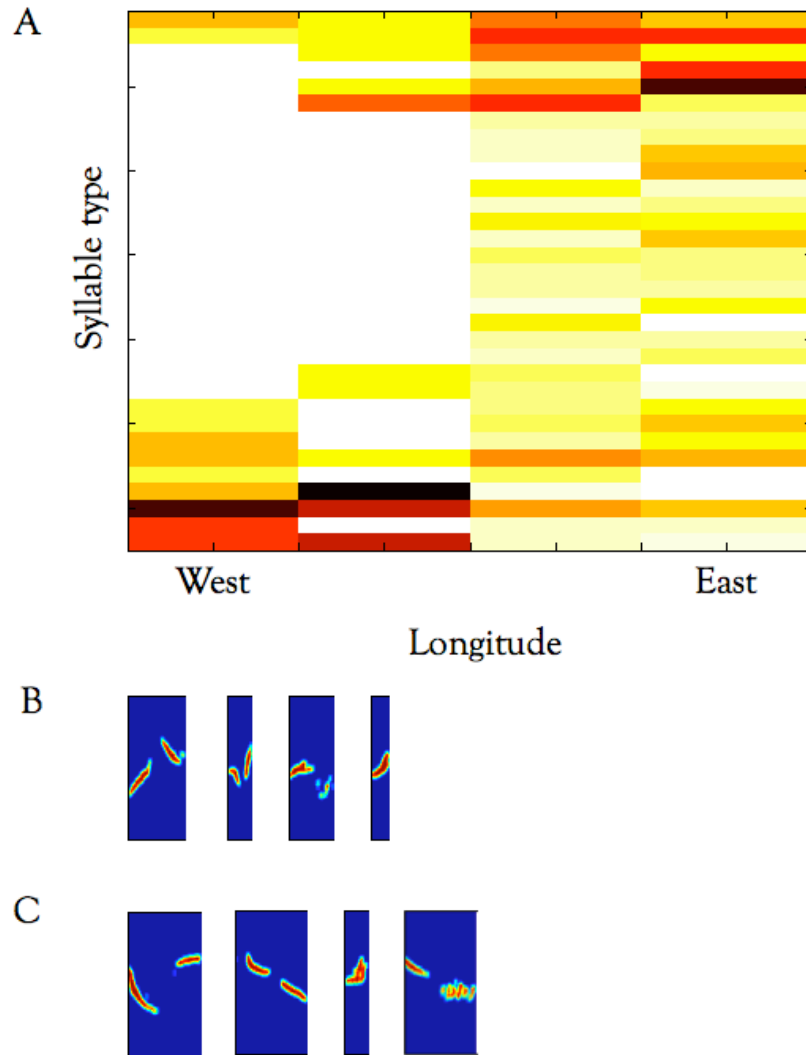


Figure 5-6. Cultural analysis of chipping sparrow syllables by longitude. Each horizontal line in the graph in (A) represents a single syllable type, and the colors for each region correspond to the fraction of the syllables in that region made up by that syllable type. The occurrence of the syllable increases as the colors progress from yellow to red to black. Syllables at the bottom of the graph occur mostly in the west, and syllables at the top occur mostly in the east. Several syllables were biased in their distribution, with those shown in (B) appearing mostly in the west and those in (C) appearing predominantly in the east.

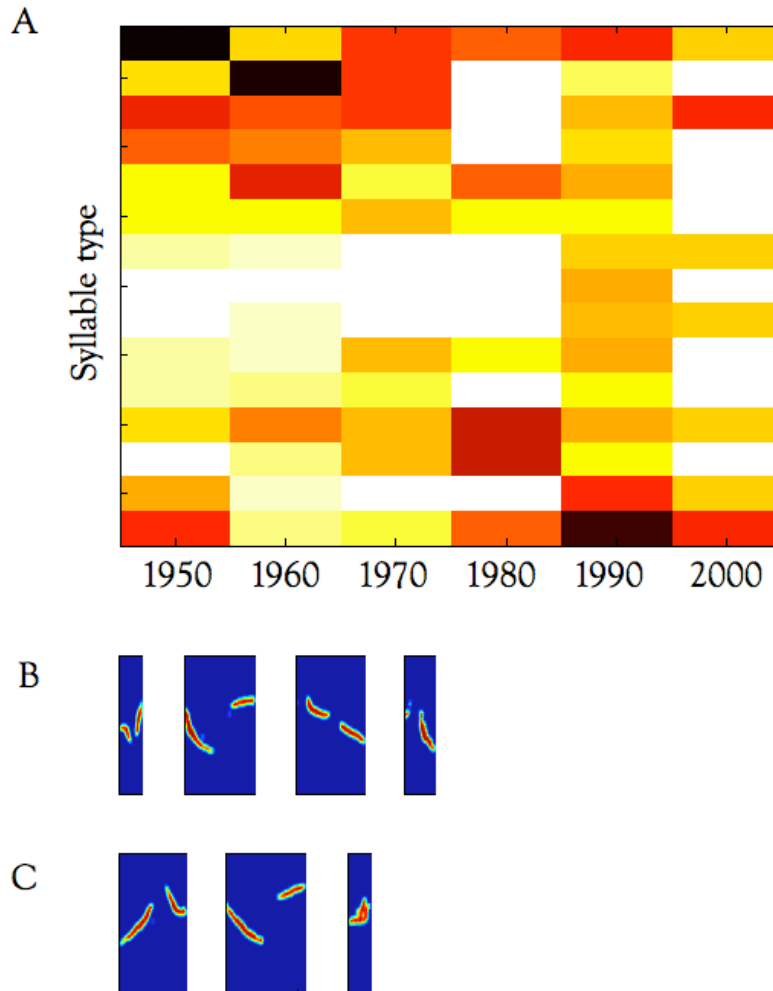


Figure 5-7. Cultural analysis of chipping sparrow syllables by year. Each horizontal line in the graph in (A) represents a single syllable type, and the colors for each decade correspond to the fraction of the syllables in that decade made up by that syllable type. The occurrence of the syllable increases as the colors progress from yellow to red to black. Syllables that are darkest at the top left of the graph occur mostly in the older recordings, and syllables that are darkest at the bottom right occur mostly in the newer recordings. Several syllables were biased in their distribution, with those shown in (B) appearing mostly in the newer recordings and those in (C) appearing predominantly in the older recordings. The second syllable in both (B) and (C) seemed very closely related and warrants further investigation.

Table 5-1. Significant associations between syllable properties and the date and location of recording. The time of day correlated significantly with duration of the song bout, which was consistent with previous research. Many of the song property variables were significantly different when grouped by longitude. P-values are reported for t-tests, and the specific t-test that was significant is noted. Longitude values with a “C” were also significantly different between the eastern and central regions. Cells highlighted in pink showed a significant positive correlation, and cells highlighted in blue showed a significant negative correlation.

|                                  | Longitude                                   | Year   | Time of day                            |
|----------------------------------|---|--|--|
| Mean syllable durations          | p=1.3*10 <sup>-21</sup><br>East vs. West, C |  |  |
| Length of longest syllable       | p=5.3*10 <sup>-23</sup><br>East vs. West, C |  |  |
| Length of shortest syllable      | p=3.4*10 <sup>-19</sup><br>East vs. West, C |  |  |
| Mean silence duration            | p=1.3*10 <sup>-23</sup><br>East vs. West, C |  |  |
| Duration of bout                 |   | p=2.6*10 <sup>-10</sup><br>1950-1969 vs. 1990-2009 | p=6.6*10 <sup>-7</sup><br>Dawn vs. day |
| # of syllables/<br>bout duration | p=2.0*10 <sup>-12</sup><br>East vs. West, C |  |  |
| Number of syllables              | p=7.6*10 <sup>-12</sup><br>East vs. West    |  | p=4.1*10 <sup>-6</sup><br>Dawn vs. day |
| Mean note length                 | p=3.9*10 <sup>-17</sup><br>East vs. West    |  |  |
| Number of notes                  | p=8.4*10 <sup>-6</sup><br>East vs. West     |  |  |
| Length of shortest note          | p=2.3*10 <sup>-7</sup><br>East vs. Central  |  |  |
| Length of longest note           | p=1.1*10 <sup>-13</sup><br>East vs. West    |  |  |
| Mean note frequency peak         | p=1.4*10 <sup>-8</sup><br>East vs. West, C  |  |  |
| Mean note frequency trough       | p=2.2*10 <sup>-8</sup><br>East vs. West, C  |  |  |
| Maximum frequency                | p=3.1*10 <sup>-17</sup><br>East vs. West, C |  |  |
| Frequency range                  | p=3.3*10 <sup>-23</sup><br>East vs. West, C |  |  |
| Downsweeps                       | p=5.6*10 <sup>-7</sup><br>East vs. West     |  |  |

## Part 6. The effects of hybridization on learned song

Oscine songbirds must learn their songs from an adult tutor, but several studies suggest that there is an innate component to learned song that can affect the song a bird produces. When the sound environment of a bird is experimentally manipulated during development, these innate predispositions can be elucidated. For example, when a synthetic song containing both song sparrow and swamp sparrow elements was used to tutor swamp sparrows, they consistently imitate the swamp sparrow elements. When a swamp sparrow note was properly embedded in a song sparrow song, however, the song sparrow could successfully imitate it (Marler and Peters 1989). In another example of tutoring with synthetic songs, canaries were exposed to very unnatural sounds during the sensitive period for vocal learning. They were capable of reproducing these sounds very accurately, but at maturation began to use elements from the synthetic song to produce typical canary phrases of repeated syllables (Gardner, Naef, and Nottebohm 2005). These examples illustrate that birds with learned vocalizations can have an innate song learning program, and the basis of learning preferences can differ by species.

A more naturalistic test of a similar principle can be performed by replacing the synthetic tutor song with a live tutor of another species. In these cross-fostering experiments, the wild type songs of males from two different species can be compared to the songs of each species when raised by a tutor of the other species. For example, when zebra finches are raised by Bengalese finch (*Lonchura striata* var. *domestica*) tutors, they can produce elements of the tutor's song but they tend to structure these elements in phrases

more typical in length of zebra finch songs. Correspondingly, Bengalese finches raised with zebra finches sang zebra finch song elements with increased periods of silence between syllables and a longer song length characteristic of Bengalese finches (Clayton 1989). Given exposure to his own species' song, however, a zebra finch seems to prefer it: when housed with Bengalese finches during the sensitive period for song learning (35 to 70 days post-hatching), zebra finches usually imitated a conspecific song that they heard prior to day 35 (Clayton 1987). When zebra finches were cross-fostered with Bengalese finches and subjected to song preference tests, they demonstrated a reduced discriminatory preference for their own species' song compared to normally raised zebra finches, which strongly preferred conspecific song to other species' songs (Campbell and Hauber 2009).

One recent study compared the songs of two zebra finches raised with a Bengalese finch to the song of hybrid offspring of a Bengalese finch and a zebra finch (Takahasi *et al.* 2006). The hybrid finch produced a better imitation of the Bengalese finch song than the cross-fostered zebra finches. In a study of two canary strains, offspring preferentially learn the song of their own strain when exposed to tutors of both song types, but hybrid offspring of the two strains learned from both tutor songs (Mundinger 1995).

Field research on hybridized birds can shed light on the role of song in both the hybrids' mate choice and the interspecific mate choice necessary to the production of hybrids. On Daphne Major in the Galapagos, Darwin's finches occasionally hybridized, and the hybrids mated according to the paternal song (Grant and Grant 1997). Thus, the genetic makeup of the F2 hybrids was determined by a learned behavior. In indigo and lazuli buntings, females were shown to choose mates based on plumage and song traits, but



the males are capable of learning song traits from the other species. Even though the fitness of hybrids is lower than that of either purebred species, hybridizations still occur in zones of overlap, putatively because females cannot (or do not) discriminate against song elements of their own species when sung by another species (Baker and Boylan 1999). In this way, the genetically determined mating signals of the plumage are superseded by the learned, and possibly error-prone, mating signals in song, resulting in offspring that are at a fitness disadvantage.

Two features of avian genetics make birds a noteworthy system for the study of hybridization and its effects on learned behavior. The time to the most recent common ancestor between two species is an estimate, usually based on molecular clock theory, of the amount of time that two species have been diverging. For most mammals, if the time to the most recent common ancestor is greater than three million years, the two species will not be capable of hybridizing. In contrast, the average time to the most recent common ancestor of hybridizable species of birds is 22 million years (Prager and Wilson 1975). It remains unknown precisely why mammals lose the ability to hybridize an order of magnitude more quickly than birds. Additionally, according to Haldane's rule, "[w]hen in the F1 offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterozygous [heterogametic] sex" (Haldane 1922). In birds, the heterogametic sex is the female (ZW) and the homogametic sex is male (ZZ), whereas mammalian males have heterogametic sex chromosomes (XY in males vs. XX in females). Therefore, if the hybridization of two divergent species of birds is biased toward one sex, the resulting offspring will be male, as opposed to a female bias in hybrid mammals, such as when

donkeys and horses hybridize to produce female mules. As a result of these two features, species of birds with diverse songs can potentially be hybridized and their song examined under different tutoring circumstances, and the hybrid offspring will be predominantly or exclusively males, which is beneficial when studying a group of birds in which all or nearly all singing behavior is performed by males.

## Methods

### *Selection of species*

To begin the interspecies breeding, we first chose an appropriate pair of species. The zebra finch (*Taeniopygia guttata*) was a logical choice for one species, since a robust year-round breeding effort was already in place at the Rockefeller Field Research Center in Millbrook, NY. The second species needed to have the ability to produce offspring with the zebra finch, but we also sought to maximize the song differences between the two species. If hybrids were indeed produced, it was preferable to be able to classify as many characteristics as possible of a hybrid's song as more similar to a zebra finch's song or to the other species' song. The Indian silverbill (*Lonchura malabarica*, also known as the white-throated munia and occasionally classified as *Euodice malabarica*) had the ability to hybridize with the zebra finch, and its songs were readily distinguishable from those of the zebra finch by visual inspection of the spectrograms (see Figure 6-3). These two birds would never have the opportunity to hybridize in nature, since the zebra finch inhabits Australia and the silverbill's range is in the Middle East and South Asia. A precedent for hybridization of

these two species had already been established in the laboratory (Lombardino, unpublished data). Several silverbills from these original experiments were still present at the Field Research Center and were included in the breeding groups for subsequent experiments, but none of the original hybrids

#### *Animal housing*

All birds were housed at the Rockefeller University Field Research Center in Millbrook, NY. The light cycle was kept constant at 12 hours of light and 12 hours of dark throughout the study, and all birds were given seed and water *ad libitum* and provided with oyster shell grit and cuttlebone. Silverbills and hybrids were provided with millet sprays in addition to finch seed mix. Breeding pairs were housed in individual cages with a nest and nest material, and their food was supplemented with a mixture of hard-boiled eggs with shells, softened seeds, and vitamins. Non-breeding adults were housed in same-sex group cages. A protocol for this work was approved by the Rockefeller University Institutional Animal Care and Use Committee.

#### *Same-species breeding*

Juvenile males raised by parents of their own species were placed in one of two learning paradigms: raised one-on-one with a same-species tutor (transferred to an individual tutor immediately after independence, approximately post-hatch day 30) or raised with a group of siblings by a same-species tutor (either raised from hatching with

their father or separated from their father before post-hatch day 12, raised by their mother, and transferred to a tutor on approximately post-hatch day 30).

#### *Description of singing behavior*

Songs from zebra finches and silverbills tutored by a male of their own species were analyzed for similarity to the tutor's song using Sound Analysis Pro 2 (Tchernichovski et al. 2000). I compared a representative motif from recordings of the tutor to five motif renditions from the adult pupil. The amplitude and entropy thresholds were adjusted in Sound Analysis Pro if necessary to segment the songs into syllables accurately. To score the similarity between two songs, I used the "Asymmetric" scoring method, suitable for comparing imitations to a model, and the "Time courses" setting for calculating distance. For each pupil, I recorded the percent similarity from each comparison and calculated the mean value to describe the mean similarity to the tutor. Usually similarity calculations for species other than the zebra finch require adjustment of feature scales in Sound Analysis Pro; however, changing these settings to suit silverbills would complicate analyses to follow, in which zebra finch, silverbill, and hybrid songs are compared to one another. For consistency, the default zebra finch settings were used throughout this analysis.

In addition to quantifying a baseline level of same-species tutor imitation for both zebra finches and silverbills, I calculated the mean values of the syllable properties and the song syntax for each species using MATLAB, as described in previous sections. Recordings of seven adult zebra finch songs and seven adult silverbill songs were processed through the analysis program, using one motif from each individual. Since these zebra finch and

silverbill syllable properties would later be compared to those of cross-fostered and hybrid individuals, I used an independent set of zebra finch and silverbill recordings obtained from birds that were not used to tutor cross-fostered or hybrid birds, such that imitation of song elements from a specific tutor would not be conflated with the extent of similarity that normally occurs between the song of these two species.

#### *Cross-fostering*

For cross-fostering experiments, clutches of eggs with similar hatch dates were switched between zebra finch and silverbill nests. Zebra finches were reared with silverbill foster parents, and juveniles were separated into two groups: one subset of birds was moved just after independence (approximately 30 days post-hatching) to individual sound isolation boxes for one-on-one tutoring by silverbills until 120 days post-hatching, and another subset was kept together with nest-mates in the silverbill foster parent's cage until 120 days post-hatching. The equivalent experiments were also performed with silverbill juveniles and zebra finch foster parents.

#### *Hybridization*

For hybridization experiments, silverbill males were paired with zebra finch females and vice versa. Pairs were given several weeks to build a nest, and the female was replaced if the pair showed no mating activity. When the interspecific mating was successful, the hybrids were divided into several categories: tutored one-on-one by an adult silverbill male, tutored one-on-one by an adult zebra finch male, raised in a group with an adult silverbill

male, raised in a group with an adult zebra finch male, and raised in isolation. Each of these categories included hybrids with a zebra finch father and hybrids with a silverbill father, and all birds remained with the tutor until 120 days post-hatching. Clutches of birds to be raised in a group remained in their hatching cage until adulthood: birds fathered by a silverbill but raised with a zebra finch tutor were relocated as eggs. Hybrids to be raised in isolation or one-on-one with a tutor of the other species were removed from their hatching cages before fledging, most often between day 10 and day 13, and hand-reared until independent. These birds were then placed in sound isolation chambers, either alone or with a tutor, until they reached 120 days post-hatching.

#### *Song recording*

Birds were recorded in sound attenuating chambers equipped with lights, airflow, and microphones. These isolation chambers were constructed from 100-quart coolers (Coleman) lined with Soundfoam-M (Soundcoat). Lavalier microphones (Audio-Technica AT803B) were attached to computers via a DirectPro preamplifier and soundcard (Aardvark), and the bird's vocalizations were recorded using Sound Analysis Pro 2 (Tchernichovski et al. 2000), which constantly monitors the sound amplitude in the chamber and begins saving temporary data once an amplitude threshold is exceeded. These temporary recordings are then processed by the program, which retains only those files that meet user-modified amplitude and entropy thresholds characteristic of song. Birds were recorded on or soon after 120 days post-hatching. In addition, hybrids were recorded periodically as adults to determine whether their song remained stable. When a bird was

raised in an isolation chamber, either with a tutor or as an isolate, its song ontogeny was also recorded continuously, barring computer crashes and power outages.

#### *Song imitation analysis*

To analyze the vocalizations of silverbills, zebra finches, and hybrids and compare between tutoring protocols, I accumulated songs from each bird recorded in sound isolation chambers as described above. As described for same-species tutoring, I chose at least five recordings from each pupil, compared each to the tutor's song using Sound Analysis Pro, and then averaged the percent similarity values for the five comparisons.

In addition, I used software developed in MATLAB to extract the syllable properties and the syntax of each bird's song, as described in previous sections. Song properties were grouped for individuals reared in the same conditions (silverbill tutored by silverbill, silverbill tutored by zebra finch, zebra finch tutored by zebra finch, zebra finch tutored by silverbill, hybrid tutored by zebra finch, hybrid tutored by silverbill) and compared between conditions with two-tailed Student's *t*-test for independent samples with unequal variance.

#### *Hybrid breeding test*

To determine whether the zebra finch-silverbill hybrids were fertile, each hybrid male was paired with a zebra finch female and given a nest. To avoid confusion due to sperm storage, care was taken to use females that had resided in an all-female aviary for an extended time. Many of these pairs successfully built a nest and laid and incubated eggs,

but none successfully developed or hatched. If these pairs seemed to be sitting on the eggs inconsistently, the eggs were moved to an incubator.

## Results

### *Breeding experiments*

Mating between zebra finches and silverbills was successful, but I anecdotally observed that interspecific pairs were less likely to build a nest and lay eggs than purebred pairs. Until the first group of cross-fostered birds reached adulthood, we were unsuccessful at crossing male zebra finches with female silverbills, perhaps because the sexually dimorphic plumage of the male zebra finch was not attractive to the sexually monomorphic silverbill. Female silverbills raised with zebra finch foster parents were much less resistant to mating with male zebra finches, and the males that succeeded in pairing with them were often cross-fostered as well.

Hybrid offspring were more likely to die just after hatching, be underfed, develop slowly, and reach independence late. It is unclear whether these animals were less robust because of their hybrid genetic material or because their mouth markings and begging calls were not strong signals to both parents to feed them. The mouth markings of hybrids were intermediate between those of silverbills and zebra finches (Figure 6-1). The mouth markings of hybrids with a zebra finch father were indistinguishable from those with a silverbill father, as were their begging calls, which were similar to those of a zebra finch. Their adult plumage was more similar to a silverbill's, but occasionally showed some zebra



finch elements, including chestnut and white markings on the flanks, dull orange cheek patches, and black striping on the chest, to varying degrees (Figure 6-2). The beaks and legs of adult hybrids were a pale peach color, unlike that of either parent. The hybrids with a zebra finch father were more likely to display zebra finch plumage elements, including darker orange beaks. Haldane's rule seemed to apply to the hybridization of silverbills and zebra finches: out of 39 adult hybrids, all were male and none was observed to be fertile. The most general learning properties of the hybrids seemed comparable to each of the parent species: by approximately 100 days post-hatching, each hybrid had a crystallized song that remained stable for the duration of the experiment.

#### *Song analysis*

I compared the similarity between tutors and male pupils for birds in six groups: zebra finches tutored by zebra finches, zebra finches tutored by silverbills, silverbills tutored by zebra finches, silverbills tutored by silverbills, hybrids tutored by zebra finches, and hybrids tutored by silverbills. All of these birds were tutored one-on-one. Each species was best at learning its own species song: zebra finches tutored by zebra finches imitated with 86% similarity (n=8) and silverbills copied silverbills with 64% similarity (n=5). In contrast, zebra finches imitated silverbill tutors with 38% similarity (n=4) and silverbills imitated zebra finch tutors with 31% similarity (n=10; Figure 6-3). Qualitatively, the hybrids copied each tutor species better than cross-fostered birds but not quite as well as purebred birds raised with their own species (Figure 6-3). Both silverbill-tutored hybrids (41.5% similarity to tutor, n=8) and zebra finch-tutored hybrids (80% similarity to tutor, n=6) were

statistically indistinguishable from purebred birds in their percent similarity to the tutor's motif. However, hybrids were significantly better at imitating zebra finch songs than silverbill songs ( $p=6.7 \times 10^{-3}$ ). Zebra finches tutored by a silverbill did not have significantly lower similarity scores than silverbills tutored by their own species. Cross-fostered silverbills, however, had significantly lower similarity scores to zebra finch tutors than both zebra finches ( $p=4.6 \times 10^{-6}$ ) and hybrids ( $p=7.3 \times 10^{-4}$ ). There was no significant difference in the similarity of imitation between hybrids with a silverbill father and hybrids with a zebra finch father (Student's t-test results:  $p=0.41$  for similarity of imitation to a zebra finch tutor by silverbill-fathered versus zebra finch-fathered hybrids,  $p=0.53$  for similarity of imitation to a silverbill tutor by silverbill-fathered versus zebra finch-fathered hybrids), so these groups were pooled according to tutor species.

In addition to the testing the fidelity of imitation between tutors and pupils, I examined the phonology and syntax of the songs of silverbills, zebra finches, and hybrids. I averaged each phonology variable based on the parentage and tutor of the bird and generated a neighbor-joining tree from the distances between the values. On average, the hybrids tutored by zebra finches grouped with purebred zebra finches, and hybrids tutored by silverbills grouped with purebred silverbills. Hybrid isolates clustered with the zebra finches and zebra finch-tutored hybrids. In terms of syllable phonology, the cross-fostered birds grouped with their genetic species and not with their tutor species (Figure 6-4). When inspected on an individual level, the syllable phonology features still grouped the birds into two major clades, one clade containing silverbills and hybrids raised by silverbills (in particular, the purebred silverbills, the hybrids tutored by silverbills, the silverbills cross-

fostered with zebra finches, one zebra finch cross-fostered with silverbills, and one hybrid isolate), and one clade containing zebra finches and hybrids tutored by zebra finches or isolated (in particular, purebred zebra finches, hybrids tutored by zebra finches, hybrid isolates, zebra finches cross-fostered with silverbills, and one silverbill cross-fostered with zebra finches; Figure 6-5). In terms of syntax, the individual birds grouped quite differently. The zebra finches, hybrids tutored by zebra finches, and hybrid isolates clustered into one closely related clade (Figure 6-6). The hybrids tutored by silverbills are in outgroups to this clade, and the purebred silverbills are in further outgroups. These groups are paraphyletic, however, meaning that they do not form their own monophyletic group without including the zebra finch clade. In a separate group are all of the cross-fostered birds, which grouped more closely to each other than to the groups of either their genetic species or tutor species.

When multiple juvenile males are reared together, the relative importance of vertical learning (from parents) and horizontal learning (from peers) can be observed. Often, one bird in the clutch will sing an accurate rendition of the tutor song, but other males in the clutch imitate to varying degrees. As a result, a set of zebra finches reared together by the same tutor can each sing a unique song. I replicated this finding in zebra finches, but groups of silverbills reared together display a different learning strategy. Each male in a silverbill clutch sings a nearly identical song, even when that song differs from the tutor's song substantially (Figure 6-7). In addition to rearing clutches of purebred silverbills and zebra finches, I cross-fostered groups of silverbills with a zebra finch foster father and groups of zebra finches with a silverbill foster father. Similarly, either a zebra finch tutor or a silverbill tutor raised clutches of hybrids. Zebra finches raised in a group with a silverbill

tutor demonstrated the same learning strategy as with a tutor of their own species: each bird in the group sang a different song, each with certain elements copied from the silverbill tutor. Silverbills cross-fostered in a group, in contrast, each sang a strikingly similar song, which bore little resemblance to the zebra finch tutor's song. Hybrids seemed to demonstrate the same learning strategy as silverbills, with each male singing a similar song that often differed from the tutor's song, whether they were raised with a zebra finch tutor or a silverbill tutor.

## **Discussion**

With one exception in each group, cross-fostered birds showed syllable properties more closely related to their genetic species than to their tutor species. The zebra finch tutored by a silverbill that succeeded in reproducing syllable properties of the silverbill, as well as the silverbill tutored by a zebra finch that accurately imitated zebra finch phonology, demonstrated that there was no species-wide physiological barrier to producing the song of the other species. However, most cross-fostered birds did not accurately imitate the phonology features of their tutor species even though they were exposed only to that species throughout development. When I examined the syntax for the same set of birds, I found the hybrids to predominantly group with their tutor species and the hybrid isolates to group with zebra finches, but both the cross-fostered silverbills and cross-fostered zebra finches formed their own clade in a monophyletic outgroup to the other birds. Thus, birds raised by a tutor of another species in this study tended to maintain phonology characters similar to their genetic species and to exhibit a syntactical structure that was different from

either species. Taken together, these results supported a hypothesis of innate predispositions that guide learned song (Marler 1990). Both species were physically capable of producing sounds that closely resembled that of the other species, but most individuals did not.

From this sizable comparison of normal breeding, cross-fostering, and hybridization, I observed a difference in the apparent inheritance patterns of song features and learning strategies. Hybrid offspring of zebra finches mated to Indian silverbills were capable of imitating a tutor of either species, and the similarity to each type of tutor was not significantly different from normal imitation in either of the purebred species. However, the songs of hybrids tutored by zebra finches were significantly more similar to their tutors than those of hybrids tutored by silverbills. In addition, the begging calls of all hybrids recorded before fledging resembled a zebra finch begging call, and the songs of isolated hybrids resembled zebra finch songs in both phonology and syntax. These results suggested that the hybrid offspring of zebra finches and Indian silverbills are predisposed to zebra finch vocalizations, although they more readily learn and produce silverbill vocalizations than zebra finches themselves. Conversely, zebra finches and silverbills demonstrated disparate learning strategies when reared in a group, and hybrids raised in this setting consistently followed a silverbill learning trajectory, in which all members of a clutch sang very similar songs, as opposed to a zebra finch learning trajectory, in which each member of a clutch often sang a unique song. Previous research has demonstrated that zebra finches have a curious learning strategy that manifests itself when multiple males are raised in the same clutch (Tchernichovski and Nottebohm 1998). A male zebra finch raised one-on-one

with a tutor will usually sing an accurate rendition of the tutor's song (vertical transfer of information from adult to juvenile), but males in a larger clutch will each sing a unique song. Two juvenile zebra finches isolated from any vertical input will usually sing a song that is very similar to one another's (horizontal transfer of information between juveniles) (Morrison 1991). Thus, zebra finches are capable of both horizontal and vertical learning. Silverbills demonstrated the same capabilities in slightly different contexts: birds tutored one-on-one could imitate a tutor well, but juveniles raised together with a single tutor tended to converge on a single song that was not always consistent with the tutor's song. Taken together, these findings in hybridized finches illustrate a species bias in two aspects of song learning, which can be inherited independently: the propensity for production and imitation of certain syllable types and syntax patterns, and the strategy for learning in an environment with both vertical and horizontal input.

Zebra finch

Hybrid

Indian silverbill

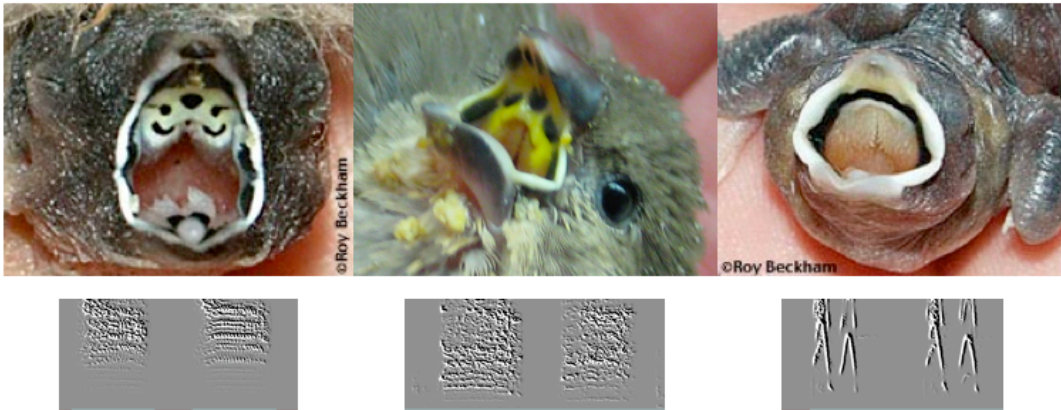


Figure 6-1. Mouth markings and begging calls of juvenile zebra finches, hybrids, and Indian silverbills. The mouth marking of the hybrid is intermediate between the zebra finch and the silverbill, but the hybrid begging call is much closer to that of the zebra finch.

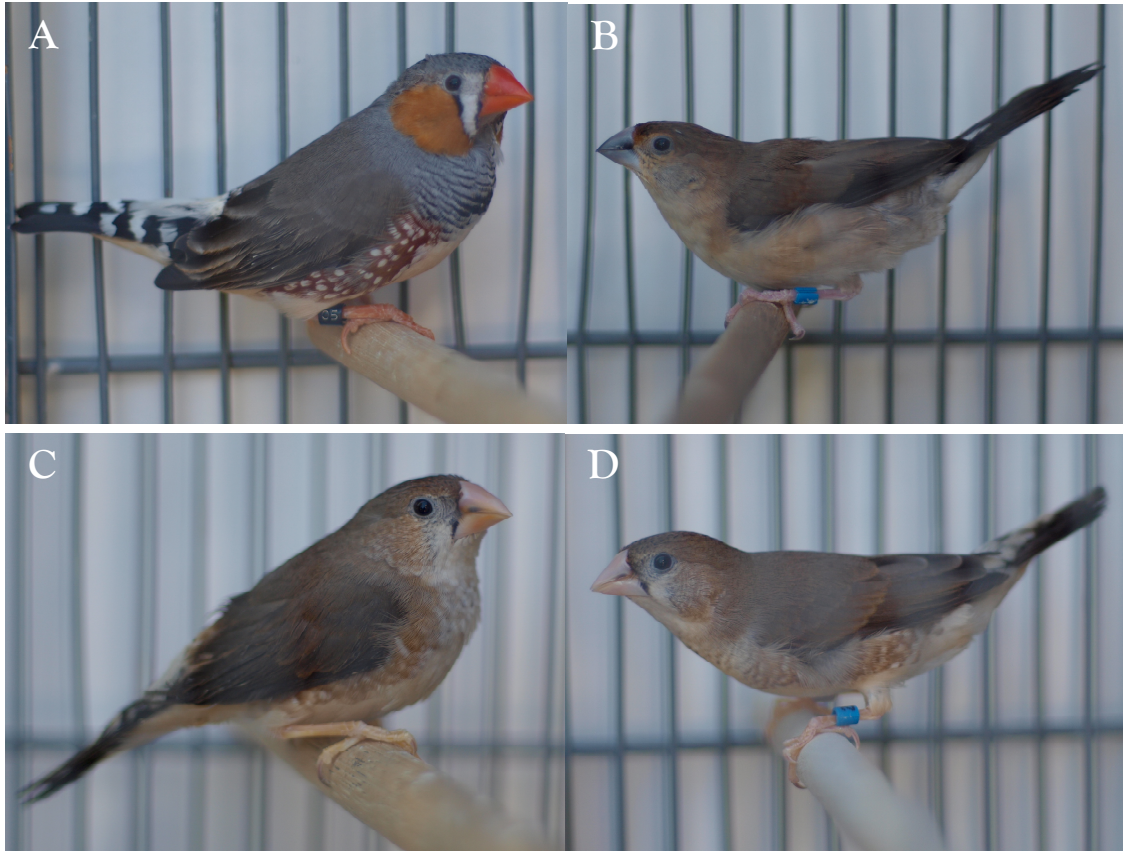


Figure 6-2. Plumage of zebra finch, silverbill, and hybrid males. Adult plumage of the (A) zebra finch, (B) Indian silverbill, (C) hybrid offspring of a zebra finch male and a silverbill female, and (D) hybrid offspring of a silverbill male and a zebra finch female. The hybrid (C) with a zebra finch father has slightly more orange coloration on the cheeks and beak.



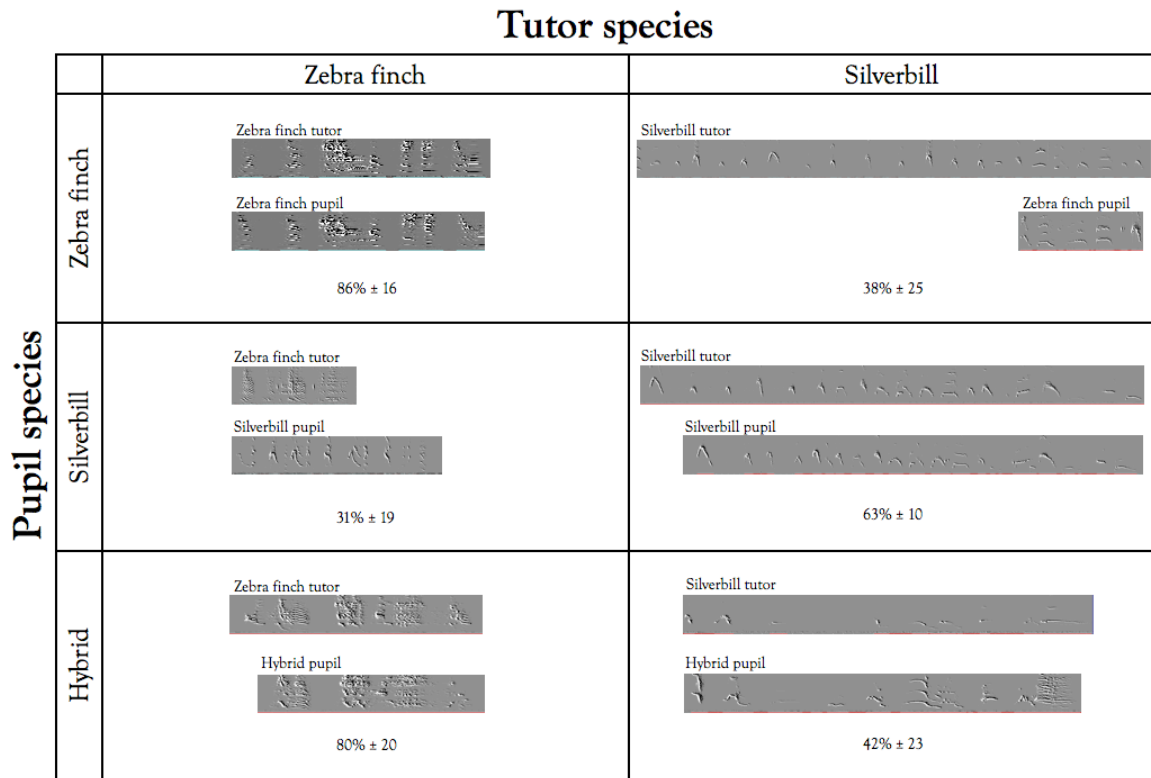


Figure 6-3. One-on-one tutoring in zebra finches, silverbills, and hybrids. Both purebred species are capable of accurate imitation of a conspecific tutor. Neither species is particularly adept at learning the other species' song, as illustrated by the similarity scores (displayed as average percent similarity, plus or minus the standard deviation). The hybrids are capable of learning either song, and the similarity to the tutor is not significantly different from that of same-species learning. Hybrids were significantly better at imitating zebra finch song than silverbill song, however.

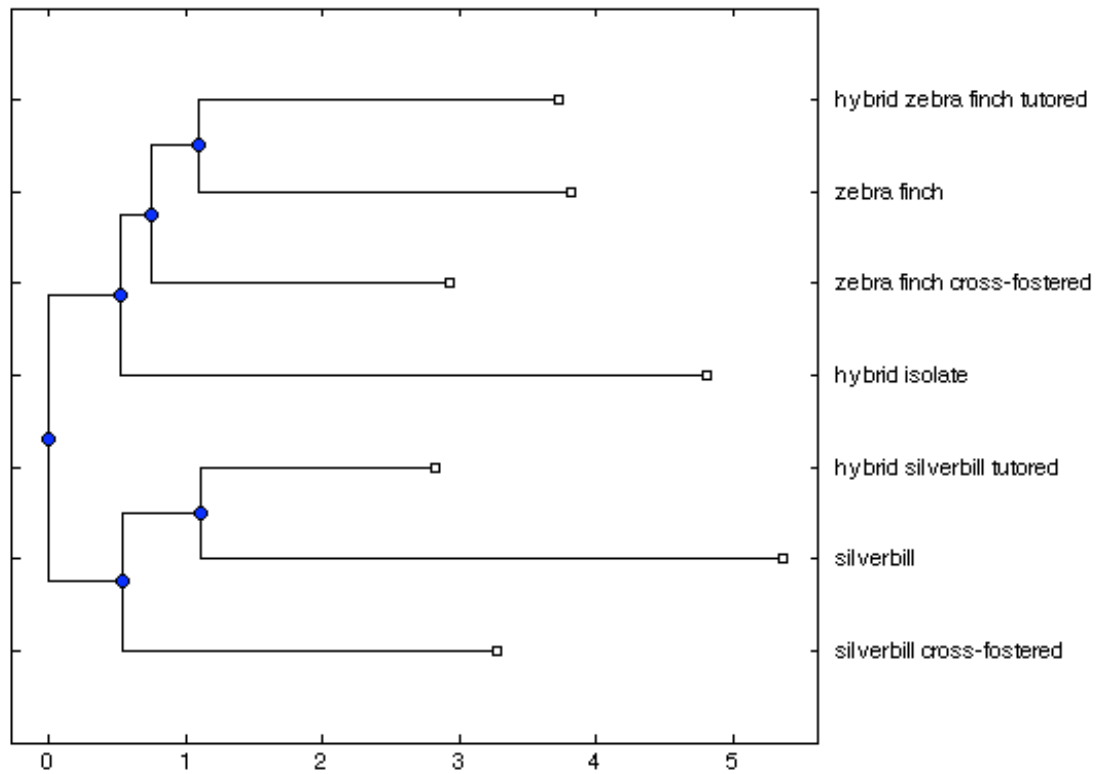


Figure 6-4. Neighbor-joining tree of average syllable properties for zebra finches, silverbills, and hybrids. The syllable properties of zebra finches and silverbills formed separate clades. Hybrids tutored by either species were capable of imitating the sound properties closely. However, cross-fostered birds maintained the syllable features of their genetic species instead of acquiring those of the tutor species. Hybrid isolates grouped in the zebra finch clade. As above, the horizontal axis measures arbitrary units of distance and the vertical axis separates the taxa.

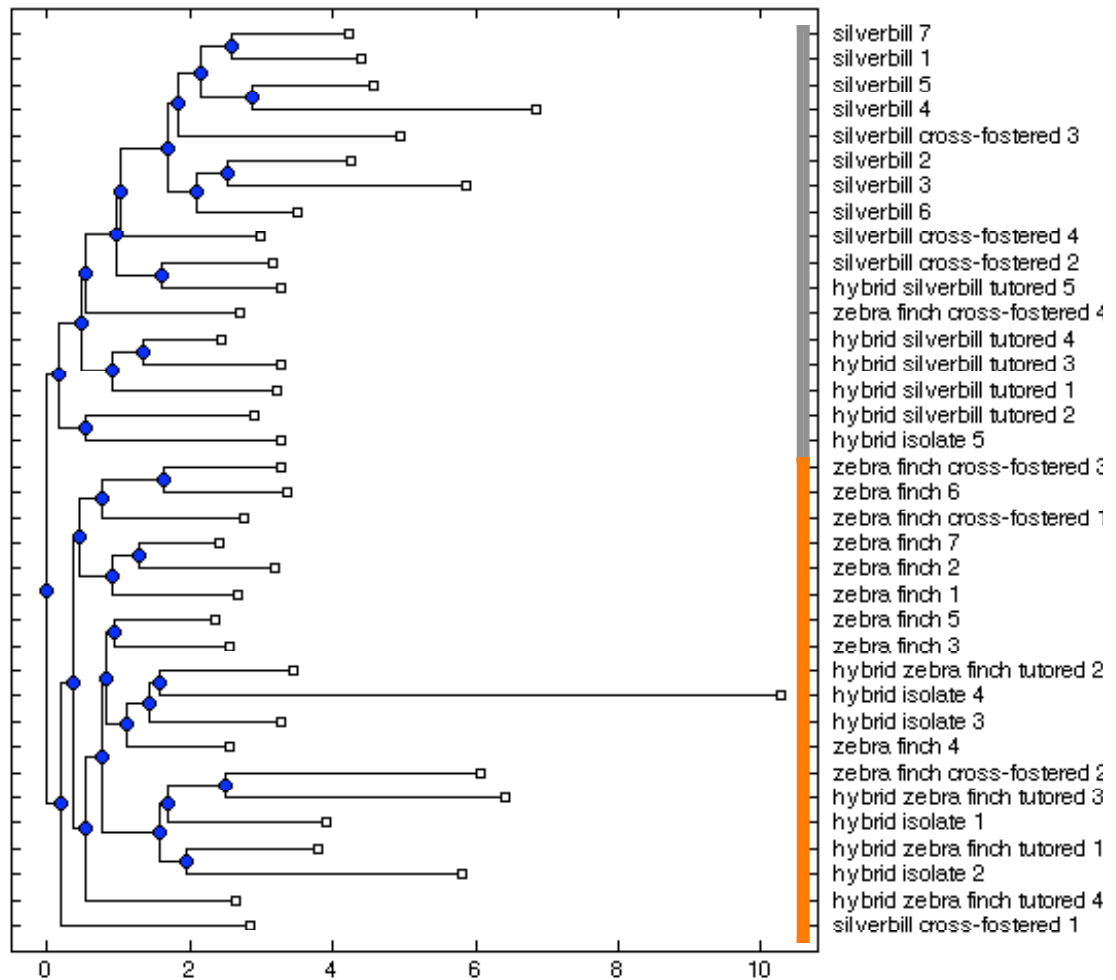


Figure 6-5. Neighbor-joining tree of syllable properties for individual zebra finches, silverbills, and hybrids. As in Figure 6-4, the birds grouped into two major clades, the gray one of mostly silverbills and silverbill-tutored hybrids, and the orange one of mostly zebra finches and zebra finch-tutored hybrids. One of each type of cross-fostered bird grouped with its tutor species instead of its genetic species, indicating that each species is capable of imitating the sounds of the other species, but most cross-fostered individuals had syllable properties more closely related to their genetic species.

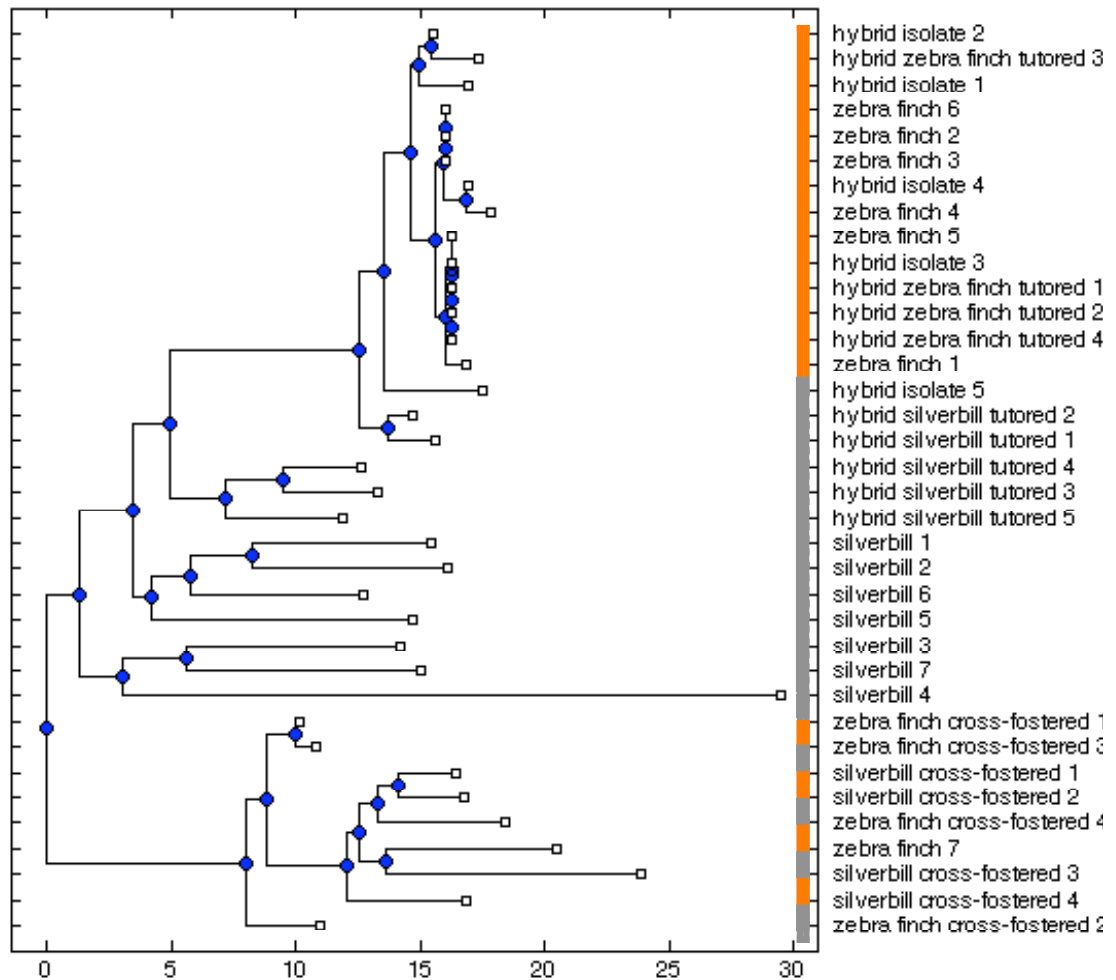


Figure 6-6. Neighbor-joining tree of song syntax for zebra finches, silverbills, and hybrids. The song syntax grouped in three parts. The orange clade was made up of zebra finches, zebra finch-tutored hybrids, and hybrid isolates. The gray clade is not a monophyletic group, but it puts all of the silverbill-tutored hybrids and silverbills in successive outgroups from the zebra finch clade. The grey and orange clade is a separate group predominantly made up of cross-fostered individuals, for which the song syntax does not group with either the genetic species or the tutor species.

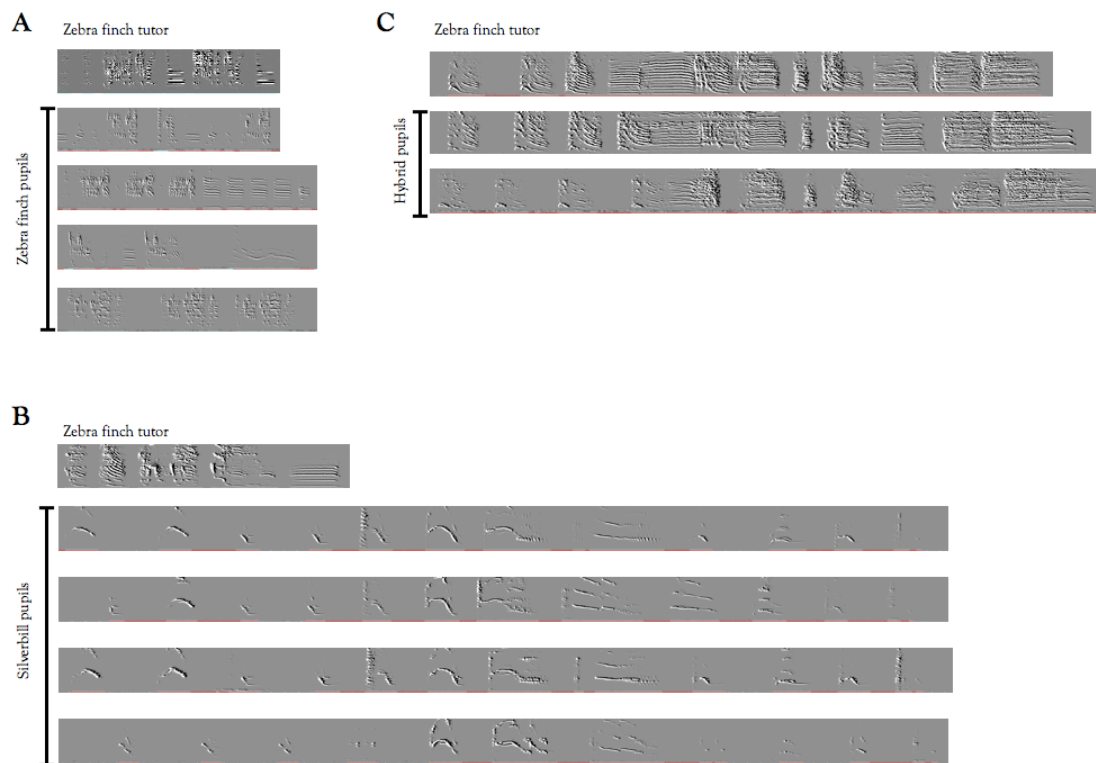


Figure 6-7. Group tutoring in zebra finches, silverbills and hybrids. All pupils were tutored by zebra finches. When multiple male zebra finches are tutored together, they often sang very different songs from one another (A). When multiple silverbills are reared together, on the other hand, their songs are very similar to one another and not necessarily similar to the tutor (B). Hybrids reared together seem to follow the same strategy as silverbills: the two birds raised together sing a very similar song, and there are a few song elements that are present in both pupils' songs and not the tutor's song, such as the extension of the final syllable (C).

## Part 7. Discussion

Using a computational analysis of song databases, I have demonstrated that birdsong, a learned behavior with evolutionary relevance, can also be a repository of evolutionary information. This analysis used both syllable properties and song syntax to examine the song relatedness of oscine songbirds and their suboscine outgroup. With these analyses, I have shown that genetic relatedness and song similarity can be closely correlated. In a small subset of species, I could accurately reconstruct the major clades of a genetic phylogeny using only song information. In a larger set, a song phylogeny demonstrated that closely related birds often have similar song properties, but the larger structure of the phylogeny was not resolved with only the song data. It is possible that more or different song properties might add additional evolutionary information, and this investigation is already underway. I have examined the song primarily at the level of individual song motifs, but higher-level comparisons could yield interesting results. For example, the amount of time between song motifs, the degree of stereotypy between motifs, and the percent of syllables that are maintained between motifs could potentially add more information at the family level of the phylogeny.

In the oscine songbirds, there is a potential for a huge diversity of sounds, as exemplified by outstanding mimics such as the lyrebird. Why, then, were these learned songs so highly conserved over millions of years? As seen in the canaries that could imitate a random set of syllables but then modified this output to conform to typical canary syntax, learned song is an interplay between the freedom to learn many sounds and the innate

predispositions that dictate which sounds are most relevant. The extent of the influence of these predispositions could be evolutionarily relevant at the more ancient nodes of the phylogeny.

The suboscine outgroup to the oscine songbirds are thought to have an innate song; that is, they do not need auditory input from a tutor during development. If singing behavior is innately programmed, then song changes must have corresponding genetic or epigenetic changes, and relatively small changes in song could potentially be sufficient for accurate species recognition in the context of speciation. In these birds, sexual selection might not act strongly on song, as it can in the vocal learners, because the content of an innately determined behavior pattern is unlikely to be a reliable indicator of male fitness. Accordingly, I hypothesized that suboscine song might have changed more slowly than oscine song over evolutionary history and thus would be more closely related between species. Upon testing, this does not seem to be the case: members of the suboscine group in this study (*Tyrannus*, *Phacellodomus*, and *Xiphorhynchus*) were found spread throughout both the syllable property tree and the song syntax tree. The song dynamics in this group certainly warrants further investigation.

In addition to demonstrating a mathematical relationship between a learned behavior and evolutionary history, during this work I have developed a set of tools that are useful for examining song on several levels. Studying zebra finches and Indian silverbills, I explored the song relationships between two species and their hybrids and demonstrated that the hybridized birds were capable of learning the phonology and syntax of either species' song. However, each purebred species was not capable of the same task: when

fostered by a heterospecific tutor, only a small subset of birds matched the phonology and syntax of the other species. Using the same analysis techniques, I closely examined a group of chipping sparrow songs to find cultural and historical patterns. I was able to find features of song that varied across times and places, and I also found individual syllables that were biased toward certain regions and years. There were strong differences in chipping sparrow songs based on longitude, but birds from the western and eastern regions are putatively members of the same subspecies.

### **Future research**

In this work, I have presented an experimental design in which groups of species were chosen based on a genetic assessment of common ancestry, and the relationships between the songs of these species were assessed and compared to the corresponding genetic relationships. Another approach to the comparison of evolutionary history and song similarity would be to begin from the behavioral data: instead of choosing species to study based on phylogeny, I could examine a large set of songs and find the species that are the most similar. I could then determine how often species with extremely similar songs are closely related.

Another area that remains to be thoroughly examined is the relationship between song and natural history. With a set of species this size, I could catalog the known features of habitat, feeding, migration, territoriality, breeding, size, foliage density, surrounding biodiversity, and more, and then determine which song features correlated with these natural history factors. With this technique, I could discover new relationships in an



unbiased way as well as test many of the hypotheses in the literature. For instance, sexual selection on a character usually drives accelerated change, and song is a powerful sexual stimulus. However, if sexual selection acted predominantly on another character, it is not clear how song would change in the process. For example, Darwin noted that the most prolific singers were often species with drab plumage; “with birds the power of song and brilliant colours have rarely been both acquired by males of the same species” (*Descent* p. 408). Examining my database, I could explore Darwin’s observations and determine which song characteristics are enriched in drab or brightly colored birds. Additionally, if a female is choosing her mate primarily based on plumage, the song divergence between species could follow a different pattern than that of birds using song as a primary attractor. Similarly, if species are radiating quickly to fill environmental niches, as in the case of Darwin’s finches on the Galapagos Islands, it is unclear how extensively their songs will diverge. On one hand, song differences could function like regional dialects to help females recognize the mates that are best suited to a certain niche. By choosing a mate with a song similar to her father’s, the female might try to ensure that her offspring will be suited to the same niche that she occupies. In response, the song changes in birds of differing niches might have become exaggerated to emphasize the ecological variation. On the other hand, the female might be better equipped to decide on a mate by observing his beak size or his technique for cracking seeds, and song will not be under strong selection and thus might change more gradually. I may be able to approach this question from another perspective, searching for clades with abnormal levels of variability and predicting which selection pressures drove speciation.

The methods for song evaluation presented here have many potential applications. I have tested these methods on several other species-level projects not mentioned here, such as determining the repertoire size of vocal mimics and finding the magnitude of song changes between pairs of closely related species, but there is a huge potential for analysis. For example, vocal learning has evolved multiple times in birds, so comparing the relationship between song and evolutionary history in three separate sets of species (songbirds, hummingbirds, and parrots) could shed light on which types of vocalization changes are consistent in the evolution of all vocal learners and which are particular to one group. In addition, adding these other avian orders to the analysis enables relatively independent tests of hypotheses regarding the relationship between song and natural history. Further, a set of field experiments could give valuable insights into whether any or all of the significantly different song features are salient to the birds in nature. Birdsong is uniquely well-suited to the study of the evolution of learning, the effects of natural selection on behavior, and the impact of the environment on behavioral characteristics. Further analysis is limited only by human creativity.

## References

- Atkinson, Q.D. and Gray, R.D. (2005). Curious parallels and curious connections—phylogenetic thinking in biology and historical linguistics. *Systematic Biology* 54:513-526
- Atz, J.W. (1970). The application of the idea of homology to behavior. In *Development and Evolution of Behavior*, ed. L.R. Aronson, E. Tolbach, D.S. Lehrman, J.S. Rosenblatt, San Francisco: Freeman.
- Backström, N., Fagerberg, S., Ellegren, H. (2008). Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Molecular Ecology* 17:964-980.
- Baker, M.C. and Boylan, J.T. (1999). Singing behavior, mating associations, and reproductive success in a population of hybridizing lazuli and indigo buntings. *The Condor* 101:493-504
- Barker, F. K., Cibois, A., Schikler P., Feinstein, J., and Cracraft, J. (2004). Phylogeny and diversification of the largest avian radiation. *PNAS* 101:11040-11045.
- Benson D.A., Boguski M.S., Lipman D.J., and Ostell J. (1994). GenBank. *Nucleic Acids Research.*, 22, 3441-3444.
- Bougey, M. J. and Thompson, N. S. (1976). Species specificity and individual variation in the songs of the brown thrasher (*Toxostoma rufum*) and catbird (*Dumetella carolinensis*). *Behavior* 57:64-90.
- Campbell, D.L.M. and Hauber, M.E. (2009). Cross-fostering diminishes song discrimination in zebra finches (*Taeniopygia guttata*). *Animal Cognition* 12:481-490.
- Carson, R.J. and Spicer, G.S. (2003). A phylogenetic analysis of the Emberizid sparrows based on three mitochondrial genes. *Molecular Phylogenetics and Evolution* 29:43-57.
- Cavalli-Sforza, L.L. and Feldman, M.W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. New Jersey: Princeton University Press.
- Chappuis, C. (1971). Un exemple de l'influence du milieu sur les émissions vocales des oiseaux: L'évolution des chants en forêt équatoriale. *La Terre et la Vie* 25:183-22.
- Clayton, N.S. (1987). Song learning in cross-fostered zebra finches: a re-examination of the sensitive phase. *Behaviour* 102:67-68

- Clayton, N.S. (1989). The effects on cross-fostering on selective song learning in Estrildid finches. *Behaviour* 109:163-175.
- Darwin, C. (1838). Notebook B: [Transmutation of species]
- Darwin, C. (1859). On the origin of species. London: John Murray.
- Darwin, C. (1874). *The Descent of Man and Selection in Relation to Sex*. New York: D. Appleton and Company.
- Derryberry, E.P. (2007). Evolution of bird song affects signal efficacy: an experimental test using historical and current signals. *Evolution* 61:1938-1945.
- Derryberry, E.P. (2010). Male response to historical and geographical variation in bird song. *Biological Letters*. epub August 4, 2010.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5):1792-1797.
- Edwards, S.V., Bryan Jennings, W., and Shedlock, A.M. (2005). Phylogenetics of modern birds in the era of genomics. *Proceedings of the Royal Society B: Biological Sciences* 272:979-992
- García-Morena, J., Sorensen, M.D., and Mindell, D.P. (2003). Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. *Journal of Molecular Evolution* 57:27-37.
- Gardner, T.J., Naef, F., and Nottebohm, F. (2005). Freedom and rules: reprogramming of a bird's learned song. *Science* 308:1046-1049.
- Gardner, T.J. and Magnasco, M.O. (2006). Sparse time-frequency representations. *PNAS* 103:6094-6099.
- Gould, S.J. (1971). *Ontogeny and Phylogeny*. Massachusetts: Belknap Press, Harvard University
- Grant, P.R. and Grant, B.R. (1997). Mating patterns of Darwin's finch hybrids determined by song and morphology. *Biological Journal of the Linnean Society* 60:317-343
- Haldane, J.B.S. (1922). Sex ratio and unisexual fertility in hybrid animals. *Journal of Genetics* 12:101-109.

- Handford, P. (1981). Vegetational correlates of variation in the song of *Zonotrichia capensis*. *Behavioral Ecology and Sociobiology* 8:203-206.
- Hackett, S.J., Kimball, R.T., Reddy, S. Bowie, R.C., Braun, E.L., Chojnowski, J.L., Cox, W.A., Han, K.L., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., and Yuri, T. (2008). A phylogenomic study of birds reveals their evolutionary history. *Science* 320:1763-1768.
- Jønsson, K. and Fjeldså, J. (2006) A phylogenetic supertree of oscine passerine birds (Aves: Passeri). *Zoologica Scripta* 35:149-186.
- Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie* 22:770-783.
- Kroodsma, D. E. (1984). Songs of the Alder Flycatcher (*Empidonax alnorum*) and Willow Flycatcher (*Empidonax traillii*) are innate. *The Auk* 101:13-24.
- Kroodsma, D. E. and R. A. Canady (1985). Differences in repertoire size, singing behavior, and associated neuroanatomy among marsh wren populations have a genetic basis. *The Auk* 102: 439-446.
- Kroodsma, D. E. and M. Konishi (1991). A suboscine bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Animal Behaviour* 42:477-487.
- Lanyon W. E. (1969). Vocal characters and avian systematics. In *Bird Vocalizations*. R. A. Hinde (ed.) Cambridge: Cambridge University Press.
- Liu, W. C. (2001). Singing behaviors of chipping sparrows. Doctoral thesis, University of Massachusetts, Amherst.
- Liu, W. C. and Kroodsma, D. (2007). Dawn and daytime behavior of chipping sparrows (*Spizella passerina*). *The Auk* 124(1):44-52.
- Liu, W. C. and Nottebohm, F. (2007). A learning program that ensures prompt and versatile vocal imitation. *PNAS* 104:20398-20403.
- Lorenz, K. (1941) Vergleichende Bewegungsstudien an Anatinen. *Jour. f. Ornith.*, 89:194-294.
- Luther, D. and Baptista, L. (2009). Urban noise and the cultural evolution of bird songs. *Proceedings of the Royal Society B: Biological Sciences* 277: 469-473.

- Marler P, Tamura M. (1962). Song “dialects” in three populations of white-crowned sparrow. *The Condor* 64:368-77
- Marler, P. and Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science* 146:1483-1486.
- Marler, P. and Peters, S. (1989). Species differences in auditory responsiveness in early vocal learning. In *The Comparative Psychology of Audition: Perceiving Complex Sounds*. R.J. Dooling and S. Hulse. New Jersey: Erlbaum.
- McCracken, K. G. and F. H. Sheldon (1997). Avian vocalizations and phylogenetic signal. *PNAS* 94:3833-3836.
- Morrison, R.G. (1991). Neural correlates of sensitive periods for song learning in zebra finches. Doctoral thesis, The Rockefeller University.
- Morton, E.S. (1975): Ecological sources of selection on avian sounds. *American Naturalist* 109:17-34.
- Mundinger, P.C. (1995). Behaviour-genetic analysis of canary song: inter-strain differences in sensory learning and epigenetic rules. *Animal Behavior* 50:1491-1511
- Nam, K., Mugal, C., Nabholz, B., Schielzeth, H., Wolf, J.B.W., Backström, N., Künstner, A., Balakrishnan, C.N., Heger, A., Ponting, C.P., Clayton, D.F., Ellegren, H. (2010). Molecular evolution of genes in avian genomes. *Genome Biology* 11:R68.
- Nottebohm, F. (1969). The song of the chingolo, *Zonotrichia capensis*, in Argentina: description and evaluation of a system of dialects. *Condor* 71: 299-315.
- Nottebohm, F. and Selander, R.K. (1972). Vocal dialects and gene frequencies in the Chingolo Sparrow (*Zonotrichia capensis*). *The Condor* 74:137-143.
- Nottebohm, F., Stokes, T.M., and Leonard, C.M. (1976). Central control of song in the canary, *Serinus canarius*. *Journal of Comparative Neuroscience* 165:457-486.
- Patten, M.G. and Fugate, M. (1998). Systematic relationships among the Emberizid sparrows. *The Auk* 115:412-424
- Payne, R. B. (1986). Bird songs and avian systematics. In *Current Ornithology*, Vol. 3 Johnston, R.F. (ed.). New York and London: Plenum Press.

- Prager, E.M. and Wilson, A.C. (1975). Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. *PNAS* 72:200-204.
- Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative Physiology and Psychology* 93:260-277.
- Prum, R. (1994) Phylogenetic analysis of the evolution of alternative social behavior in the manakins (Aves: Pipridae). *Evolution* 48:1657-1675
- de Quieroz, A. and Wimberger, P. (1993). The usefulness of behavior for phylogeny estimation: Levels of homoplasy in behavioral and morphological characters. *Evolution* 47:46-60.
- Ratnasingham, S and Hebert, P.D. (2007). BOLD: The Barcode of Life Data System. *Molecular Ecology Notes* 7(3):355-364.
- Ren, F., Tanaka, H., and Yang, Z. (2009). A likelihood look at the supermatrix - supertree controversy. *Gene* 441:119-125.
- Renfrew, C. (1999). Time depth, convergence theory, and innovation in proto-Indo-European: 'Old Europe' as a PIE linguistic area. *Journal of Indo-European Studies* 27:257-293.
- Schleicher A. (1863). *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: Hermann Bohlau.
- Slack, K., Delsuc, F., Mclenachan, P., Arnason, U. and Penny, D. (2006). Resolving the root of the avian mitogenomic tree by breaking up long branches. *Molecular Phylogenetics and Evolution* 42:1-13.
- Sibley, C. and Ahlquist, G. (1990). *Phylogeny and classification of birds*. Yale University Press, New Haven, CT.
- Takahashi, M., Kagawa, H., Ikebuchi, M., and Okanoya, K. (2006). Case studies of song and call learning by a hybrid Bengalese-zebra finch and Bengalese-fostered zebra finches: assessing innate factors in vocal learning. *Ornithological science* 5:85-93.
- Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24:1596-1599.

Tchernichovski, O. and Nottebohm, F. (1998) Social inhibition of song imitation among sibling male zebra finches. *PNAS* 95:8951-8956.

Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B. and Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. *Animal Behavior* 59:1167-1176.

Thielcke, G. A. (1976). *Bird Sounds*. Ann Arbor: The University of Michigan Press.

Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the Chaffinch (*Fringilla coelebs*). *Ibis* 100:535-570.

Wasserman F. E. (1979). The relationship between habitat and song in the White-throated Sparrow. *Condor* 81:424-426.