

CULTURAL EVOLUTION IN NATURAL POPULATIONS:
A QUANTITATIVE BIOACOUSTIC ANALYSIS

By

Chenghui Ju

A dissertation submitted to the Graduate Faculty in Biology,

In partial fulfillment of the requirements for the degree of Doctor of Philosophy,

The City University of New York Graduate Center

2015

© 2015

CHENGHUI JU

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction
of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Chair of Examination Committee
Dr. David C. Lahti
Queens College, CUNY

Date

Executive Officer
Dr. Laurel A. Eckhardt

Dr. Jennifer A. Basil
Brooklyn College, CUNY

Dr. Ofer Tchernichovski
Hunter College, CUNY

Dr. Stefano Ghirlanda
Brooklyn College, CUNY
Stockholm University

Dr. Michael J. Hickerson
City College, CUNY

Dr. Kara L. Belinsky
State University of New York, New Paltz

Abstract

Cultural evolution in natural populations: a quantitative bioacoustic analysis

By Chenghui Ju

Advisor: Dr. David C. Lahti

Bird song is a powerful model system in behavioral biology, especially for learning and cultural evolution. Understanding the origins and maintenance of vocal diversity in nature is fundamentally important to acoustic biology. Here, we propose a large-scale, integrative population analysis of nearly 2000 songs of the house finch (*Haemorhous mexicanus*) (190 individuals) collected in western Long Island of New York during an interval of 37 years (between 1975 and 2012) to explore cultural change of bird songs. Generally, I have three aims in this study. In the chapter 1, to facilitate acoustic analysis for field recordings with ambient noise, of large sample sizes and with continuous variation, I introduce an algorithm (FinchCatcher), which extract signals from the background noise, summarizes songs as a series of spectral shapes and performs dynamic classification that optionally combine the advantages of hierarchical clustering and partitioning. I further test the algorithm on a geographical comparison of 240 songs of the house finch from eight localities, using previously published observations as ground truth for estimating transitions of song (or song element) sharing during the spread of the house finch in North America. In Chapter 2, I quantify spatial variation in the songs of the house finch in western Long Island in 2012. Previous studies have suggested large morphological and behavioural differences between house finches in the east and west coast, which may have profound impact on song evolution of eastern house finch. The result shows great acoustic variation across the sampling range, rather than finding discrete areas within which song and

syllable repertoires are highly similar between individuals. Nevertheless, spatial differentiation was neither simply clinal change with geographic distance, nor discrete dialects. To further explore the mechanisms underlying cultural change in this study region, in Chapter 3 I perform a temporal comparative analysis of house finch songs spanning nearly four decades. Substantial cultural change is observed in terms of local song and syllable sharing, song complexity and song type consistency. In addition, not a single song type persists over time and half of syllable types defined in the past data are not present in the recent songs. These results illustrate the potential interplay between multiple drivers of spatial and temporal variation. This body of research has provided a framework for understanding the spatiotemporal variation in house finch song in Western Long Island, incorporating partial migration, population growth, relaxed selection by females on male song, and developmental stress playing interacting roles.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. David C. Lahti for the continuous support and encouragement of my Ph.D study, as well as for his patience, motivation, passion and encyclopedic knowledge. His generous guidance made it possible for me to complete the research and writing of this thesis. I could not have imagined having a better advisor for my study.

I am hugely indebted to Prof. Jennifer Basil for leading me into the world of animal behavior and giving me great support in life and study during my first several years in U.S.

I am thankful to Prof. Stefano Ghirlanda for helping me shape up my mathematical skills.

I would like to express my gratitude to Prof. Ofer Tchernichovski for providing me with material and books that I could not possibly have discovered by my own, and for his precious comments to help me to develop my thesis.

I am grateful to Prof. Kara L. Belinsky for her immense interest in my topic of research, for her kind words and suggestion.

My sincere thanks also go to my fellow labmates Frances Geller, Elliot Aguilar, Khaleda Khan and Aaron Owen from Lab of Prof. David C. Lahti for their invaluable help and suggestion during my study, and for all the fun we have had in the last four years. In particular, I am grateful to Frances Geller for her great work in field sampling.

This dissertation is dedicated to the memories of Prof. Paul Mundinger who is one of the pioneer and leading researchers on the acoustic studies of House Finch songs.

Last but not the least, I would like to thank my family and my friends for supporting me spiritually throughout my Ph. D study and my life in general.

Table of Contents

Abstract.....	iv
Acknowledgements.....	vi
List of Tables.....	viii
List of Figures.....	ix
CHAPTER 1 : A method for characterizing bird song diversity between individuals and populations.....	1
References:.....	31
CHAPTER 2 : Quantifying spatial variation in house finch (<i>Haemorhous mexicanus</i>) song.....	36
References:.....	70
CHAPTER 3 : Temporal variability in house finch (<i>Haemorhous mexicanus</i>) song: Four decades of cultural evolution	74
References:.....	119

List of Tables

Table 1.1. Bird song recordings used for analysis	8
Table 1.2. Descriptions of ten features.....	15
Table 2.1. Measured features of the house finch vocal profile	50
Table 2.2. The linear discriminant function coefficient based on vocal profile.....	63
Table 3.1. Measured features of the house finch vocal profile	88
Table 3.2. Statistics of tests comparing means and variability	105
Table 3.3. Summary of song comparison between 1975 and 2012	118

List of Figures

Figure 1.1. Summary of FinchCatcher’s quantitative analysis procedure.	9
Figure 1.2. Illustration of signal localization for pure-tone sounds	14
Figure 1.3. Validation of comparison method on house finch songs	19
Figure 1.4. Representative songs of house finch from eight localities	21
Figure 1.5. Heatmap of UPGMA cluster analysis of house finch songs from eight localities based on similarity in syllable lexicon.....	22
Figure 1.6. Bioacoustical analysis of three swamp sparrow songs.....	24
Figure 1.7. Bioacoustical analysis of three Chaffinch songs	26
Figure 2.1. Sampling localities of 96 individuals in western Long Island	43
Figure 2.2. Bar plot showing frequencies of the 275 identified syllable types in the population.	51
Figure 2.3. Density plot exhibiting the distribution of maximum song similarity between individuals	52
Figure 2.4. Examples of limited sharing of songs between nearby individuals	53
Figure 2.5. Cumulative percentage of syllable types detected by nth song	54
Figure 2.6. Syllable sharing between individuals versus distance.....	55
Figure 2.7. Extent of syllable sharing between pairs of individual house finches in the map.....	57
Figure 2.8. Localized syllable composition based on UPGMA clustering.....	59
Figure 2.9. Scatterplot showing the relationship between $SS_j(\text{adj})$ values for phrase sharing and geographic distance	60
Figure 2.10. Sharing of phrases in the map	61
Figure 2.11. Geographic pattern of concavity.....	64

Figure 3.1. Sampling localities of 94 individuals (diamonds) from 1975 and 96 individuals (stars) from 2012 in western Long Island.....	81
Figure 3.2. Validity of similarity criteria to define song type.....	91
Figure 3.3. Violin plots showing distribution of song similarity scores between individuals	93
Figure 3.4. Pie charts illustrating song type sharing among house finches in Western Long Island in 1975.	95
Figure 3.5. Results of syllable classification and frequencies of identified syllable types in each year	97
Figure 3.6. Difference in sharing of house finch syllable repertoires between 1975 and 2012....	99
Figure 3.7. Geographic pattern of syllable repertoire sharing based on UPGMA clustering in 1975	100
Figure 3.8. Validation of sample size to study song type consistency	102
Figure 3.9. Within-individual song type consistency between years.....	103
Figure 3.10. Song repertoire size and syllable repertoire size between years	104
Figure 3.11. Boxplot showing tests comparing means and variability of ten vocal parameters between years	106
Figure 3.12. Distribution of frequency of syllable types shared by individuals in different regions and different years	117
Figure 3.13. Average proportion of common syllable types per song per individual in two different regions of 2012	118

CHAPTER 1:

A method for characterizing bird song diversity between individuals and populations

Abstract:

Acoustic analysis of animal communication in nature to assess geographic and temporal variation requires a standardized and effective way of identifying sound units, classifying them into distinct types, and measuring differences in any feature of interest. Each of these three steps, however, can be difficult and subject to inconsistency. Here we introduce an algorithm (FinchCatcher) optimized for field recordings of pure-tone bird signals with ambient noise, able to extract features informative for subsequent analysis and well suited to classification of a large number of songs or song elements. We incorporated the algorithm into an application program which automatically separates signals from the background noise, and summarizes songs as a series of spectral features. Finally, the software implements a dynamic branch cutting methods for syllable classification. We tested the algorithm on a geographical comparison of 240 songs of the house finch (*Haemorhous mexicanus*) from eight localities, using previously published observations as ground truth for estimating transitions of song sharing during the spread of the house finch in North America. We further test its performance on a few songs of two species (the swamp sparrow (*Melospiza georgiana*) and chaffinch (*Fringilla coelebs*) for its versatility. Our software provides an efficient way to quantify acoustic information of field data to facilitate subsequent analyses of geographic variation and cultural evolution of vocal signals.

Key Words: Spectral-shape features, acoustic analysis, FinchCatcher, automated recognition, bird songs, dynamic classification, field data, house finch

Introduction:

The study of acoustic communication has always been a productive and integrative field of research, especially with respect to bird song (Kroodsma, Miller, & Ouellet, 1982; Podos, Huber, & Taft, 2004; Thorpe, 1961). Comparing acoustic features of song has led to the testing of hypotheses relating to species recognition and sexual selection (Ballentine, 2009; Derryberry, 2010; Read & Weary, 1990). Description of spatiotemporal variation in bird song provides a basis for studies of cultural evolution (B. E. Byers, Belinsky, & Bentley, 2010; Payne, 1996). Attention to the acoustic features of song has also elucidated key principles underlying vertebrate learning and motor performance, such as the mechanism by which neural circuits generate sequential behaviors (Fee & Scharff, 2010), the relationship between song features and individual performance abilities (Podos, Lahti, & Moseley, 2009); and the developmental interaction between inherited and learned aspects of song (Lahti, Moseley, & Podos, 2011; Mundinger & Lahti, 2014; Podos et al., 2009). However, extensive variability in many aspects of acoustic signals, from song themes to individual syllables, undoubtedly leads to a methodological challenge: how to describe sound signals to facilitate comparison. The early way for researchers to describe a song was to pick a set of features at will, mainly relying on the eyes of the researchers or of volunteers to distinguish syllables and classify them into distinct categories, or a few measurements of acoustic features performed by hand (Fernández-Juricic et al., 2005; Mundinger, 1975). This traditional approach raises three major issues: (1) Human subjectivity is inevitable, resulting in selection of features without knowledge of which ones are objectively most variable or most biologically relevant. (2) Measurement of many features by hand would result in inconsistency, in particular when dealing with recordings from different acoustic backgrounds, in which properties of song can be masked or variously distorted (Brumm

& Slabbekoorn, 2005). (3) Quantitative analysis involving large sample sizes and multiple features would be prohibitively labor intensive. To address these issues, some researchers have adopted partly or fully automated procedures to identify song elements from a large number of recordings (S. E. Anderson, Dave, & Margoliash, 1996; Lachlan, Verhagen, Peters, & Cate, 2010; Nagy & Rockwell, 2012; Potamitis, Ntalampiras, Jahn, & Riede, 2014; Tachibana, Oosugi, & Okanoya, 2014; Wellock & Reeke, 2012). Some even provide a complete computational framework for analysis of animal communication from identification, to feature extraction, to classification, such as Sound Analysis Pro (SAP) (Tchernichovski, Nottebohm, Ho, Pesaran, & Mitra, 2000), Avisoft (<http://www.avisoft.com>), and Luscinia (<http://luscinia.sourceforge.net/index.html>) (Lachlan, 2007). These programs yield a sophisticated output of song features that facilitates downstream analysis. A few researchers have taken advantage of such datasets for advanced comparative analysis (Daou, Johnson, Wu, & Bertram, 2012; Laiolo, Vögeli, Serrano, & Tella, 2007; Leonardo, 2004; Wu, Thompson, Bertram, & Johnson, 2008).

Nevertheless, many existing software are optimized for recordings under laboratory conditions, and their performance can be compromised for field recordings with ambient noise, potentially resulting in inaccurate and inconsistent quantitative measurements of acoustic features, and inefficient separation of individual syllables under different acoustic backgrounds. In addition, most available software either lacks a hierarchical clustering method, or provides a standard one that can be inappropriate for acoustic data of large sample sizes and with continuous variation. SAP, one of the exceptions, uses nearest neighbor hierarchical clustering (NNC), a nonparametric, density-based method to define clusters regardless of cluster shapes (SAP online manuals:

<http://soundanalysispro.com/manual-1/chapter-12-clustering-syllables/introduction-to-clustering>).

Existing application programs provide varying methods for spectrographic representation of acoustic signals. For example, Avisoft uses spectrographic cross-correlation to produce a similarity matrix in which each similarity value is measured as the peak of the correlation between the spectrograms of the two signals (Clark, Marler, & Beeman, 1987), while SAP and Luscinia extract features to represent standard and mathematically well-defined properties of audio signals. Cross-correlation values can only provide reliable measures of similarity between simple and continuous signals. When dealing with more complex signal structures, cross-correlation often returns misleading coefficients because of the complementarity of frequency and time resolution (Cortopassi & Bradbury, 2000; Khanna, Gaunt, & McCallum, 1997). Feature extraction is limited by the scope of the feature sets used and, as noted by the authors themselves, determining the “right features” to characterize animal sounds *a priori* can be difficult (Giret et al., 2011; Tchernichovski et al., 2000).

In this study, we present a computational method that is well suited to field recordings of large sample sizes, and provides alternative analytical techniques to complement the existing range of programs used by behavioral biologists. Our application program, FinchCatcher, separates acoustic signals from background noise, minimizing the influence of varied acoustic backgrounds on measurements. FinchCatcher introduces dynamic tree-cut procedure for classification. This method first takes advantage of hierarchical clustering to perform an

unsupervised classification, defining a series of core clusters with small within-cluster variances; then it relies on a partitioning algorithm to extend these core clusters and exclude outliers.

FinchCatcher uses feature extraction to describe signals, as many researchers have performed automated procedures to extract a number of acoustic features (temporal or spectral) in the signal for modeling purposes (Angle & Coskun, 2014; Kakishita, Sasahara, Nishino, Takahasi, & Okanoya, 2007; Saar & Mitra, 2008; Somervuo, Harma, & Fagerlund, 2006). FinchCatcher emphasizes spectral shapes (see below), as classification based on such features generally agrees well a human expert classifying based on spectrographic representation (Giret et al., 2011). In this way, FinchCatcher can be considered to be a quantification or formalization of a method long used in studies of bird song that have involved categorization of syllables according to overall shapes by eye (e.g., Marler & Pickert, 1984). In addition, several of these features might relate to motor performance based on the concept of motor limits to birdsong production (Podos and Nowicki, 2004) and therefore provided information valuable for subsequent analysis.

Spectral-shape features have also been popular in more quantitative studies. For instance, geometric morphometrics (Bookstein, 1996) and SoundPoints (Taft, BN: <http://www.landmarkacoustics.com>) use landmarks, defined as points in time-frequency-amplitude space, to represent shapes quantitatively. FinchCatcher breaks each song into segments or syllables, each of which is converted to a set of measurements of spectral-shape features. In addition to commonly included features pertaining to single and simple aspects of shapes such as syllable duration, FinchCatcher introduces more complex derivations like the number of concavity and frequency-time excursion length. The current version of FinchCatcher is better suited to acoustic analysis of bird songs of pure tones, as many spectral-shape features can be clearly defined in these songs.

To test our approach, we tested the performance of our program on songs of three bird species, choosing different sets of features introduced above (simple and complex spectral-shape features). We first perform a full analysis on a sample of house finch (*Haemorrhous mexicanus*) songs to test several previously published observations. The house finch, native to the western United States and Mexico, was introduced from California into the New York City area around 1940 (BirdLife International, 2013). The introduced population gradually spread westward, reaching Iowa in 1982 and first reaching the eastern edge of the western population around 1990 in central Oklahoma (Belthoff & Gauthreaux Jr, 1991; Elliott & Arbib, 1953). We predicted that song patterns of the house finch throughout its geographic distribution would corroborate this historical information. We then briefly tested the versatility of acoustic analyses on songs of the swamp sparrow (*Melospiza georgiana*) and chaffinch (*Fringilla coelebs*). Swamp sparrows sing relatively simple songs, consisting of repetitions of several tonal elements (Clark et al., 1987). Chaffinches sing songs that are more varied in terms of tonality, but with an clear syntax: typically they start with a trill of several repeated syllables, followed by a “flourish”, which usually consists of a short phrase of high-pitch elements and a long buzzy element (Thorpe, 1954).

Methods:

Song data

Archived recordings were retrieved from three sources: the Cornell Lab of Ornithology Macaulay Library of Natural Sounds (n = 43 individuals; house finch), The Ohio State University Borror Laboratory of Bioacoustics (n=20 individuals; house finch), and xeno-

canto.org (n=13 individuals; 3 swamp sparrow, 3 chaffinch and 7 house finch). These data were supplemented with songs of 10 house finch individuals recorded by one of us (FCG) in the field in 2012 (New York City and vicinity). Individuals were recorded until they stopped singing or flew away. Each site was visited only once to minimize the chances of recording the same individual repeatedly. Within each site, the next individual to be recorded was at least 100 meters away from any other individual recorded. Tracks with multiple song bouts were excluded from analysis, except in cases where a second bird was reported as singing as soon as a first bird flew away. Table 1.1 lists detailed information for the recordings. The original field recordings were saved with a sample rate of 44100 Hz, sample width of 16 bits, through one channel (mono). The total number of frames varies from 200,000-1,000,000 per song (4.5s~22.7s). For each individual of house finch, we randomly picked 3 songs to make a larger song dataset for analysis.

Table 1.1. Bird song recordings used for analysis

House finch	State	Year	Number of individuals	Source ¹
	Hawaii,USA (HI)	1962, 1976-1977	10	C,X
	Mexico (MEX)	1957-1983	10	C,X
	Arizona,USA (AZ)	1959-1981	10	B,C,X
	California,USA (CA)	1959-1981	10	B,C,X
	New York (1970s), USA (NY1970s)	1975-1980	10	C
	New York (2012), USA (NY2012)	2012	10	F
	Oregon, USA (OR)	1989	10	B,C,X
	Texas, USA (TX)	1961	10	B,C,X
Chaffinch	Sweden	2014	3	X
Swamp Sparrow	Minnesota, USA	2014	3	X

¹B:Borror Laboratory; C: Cornell Lab of Ornithology; X: xeno-canto.org; F: FCG

Signal Extraction and Quantitative Description by FinchCatcher

Quantitative description of songs using FinchCatcher was carried out in six steps. Figure 1.1 summarizes the procedure.

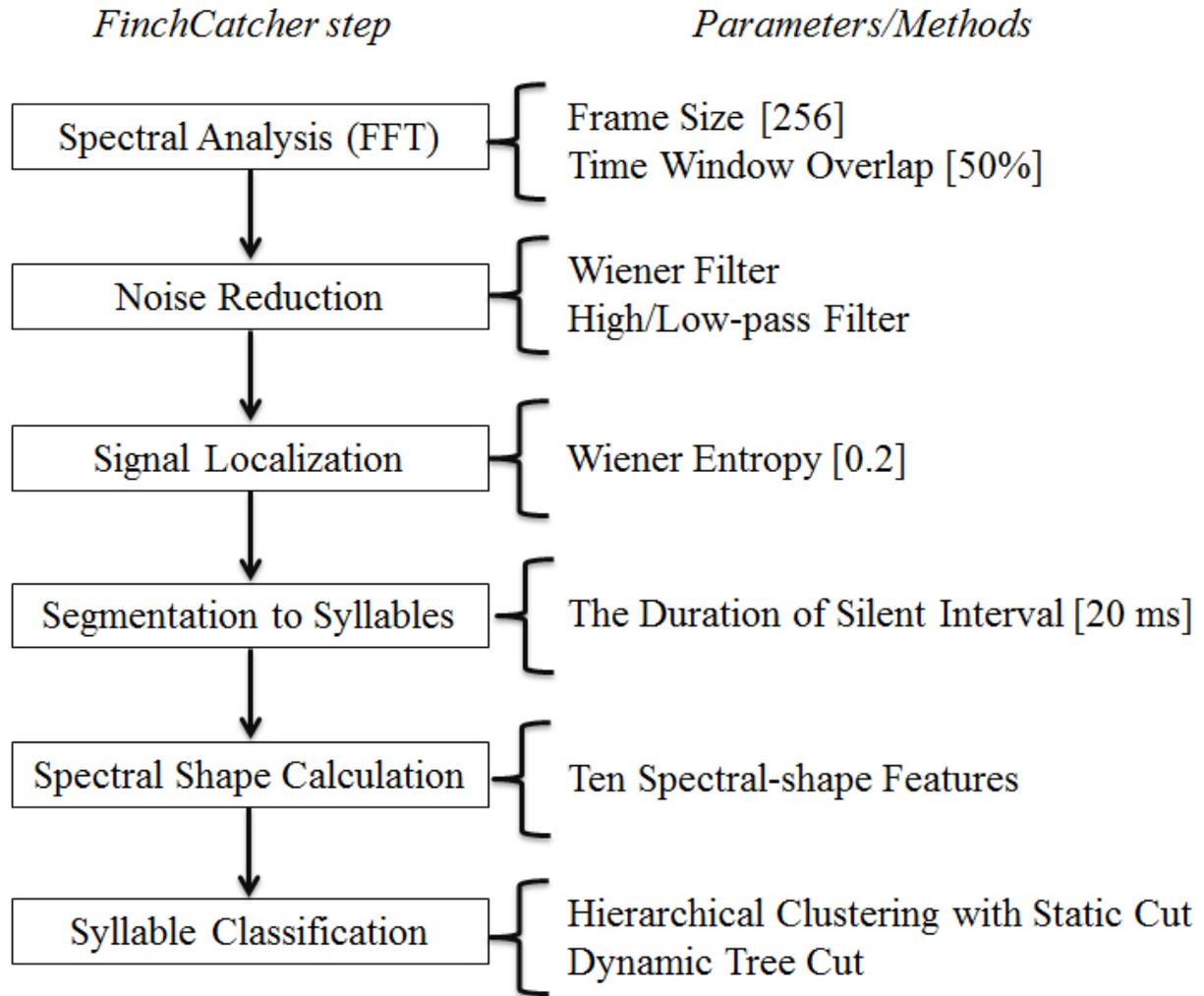


Figure 1.1. Summary of FinchCatcher’s quantitative analysis procedure.

The left column lists steps in order. The right column lists corresponding parameters for each step. Default values are in square brackets.

Spectral Analysis

The process of generating digital data is based on the idea of the spectrogram, which is a visual representation of sound with time on the x-axis and frequency (pitch) on the y-axis (Koenig, Dunn, & Lacy, 1946; Thorpe, 1958). To standardize intervals along time and frequency axes, the frame size, known as Fourier transformation length or time window, is set to a constant of any of two to the power of n (2^n). In our analysis, the constant by default is 256, which means that every 256 frames (or $256/44100=0.0058s=5.8ms$ of song) are grouped as a time segment upon which Fast Fourier transformation (Frigo & Johnson, 2005) is performed. The corresponding frequency resolution is 172 Hz. The next window starts by default 2.9 ms after the beginning of the previous one and therefore has a 50% overlap. These two parameters (frame size and overlap) are user-modifiable in FinchCatcher. Note that increases in frame size will result in decreases in bandwidth (increase in resolution in the frequency axis) and increases in the time constant (decrease in resolution in the time axis).

Noise Reduction

Identification of song signals under natural conditions is often hindered by the presence of extrinsic noise (Fantana & Kozhevnikov, 2014). In order to achieve reliable acquisition of signals, noise reduction algorithms have recently become part of digital acoustic coding systems. Such systems work within a trade-off between signal quality and noise reduction (Vary & Martin, 2006). In an ideal scenario, the noise signal is maximally attenuated in the output signal, while the focal signal is minimally distorted. In addition, any noise remaining in the processed signal should not sound unnatural. Noise reduction can be performed in many different ways,

such as Spectral Subtraction (*Boll, 1979*), Wiener Filter (*Grover & Hwang, 1992*), the Weighting Rules of Ephraim and Malah (*Ephraim & Malah, 1985*) and Psychoacoustical Weighting Rules (*Gustafsson, Martin, Jax, & Vary, 2002*). Basically, different spectral regions of the mixed signal of sound and noise are attenuated with different factors. FinchCatcher employs Wiener Filter, for the following reasons: (1) it assumes that signal and (additive) noise are stationary with known spectral characteristics, as those in natural recordings often are; (2) the occurrence of unnatural noise can be effectively reduced; and (3) it is computationally inexpensive, a characteristic especially important for large sample sizes. This method returns a relatively clear and clean spectrogram that facilitates further analysis. In addition, users can remove low-frequency/high-frequency recording noise with a high/low-pass filter.

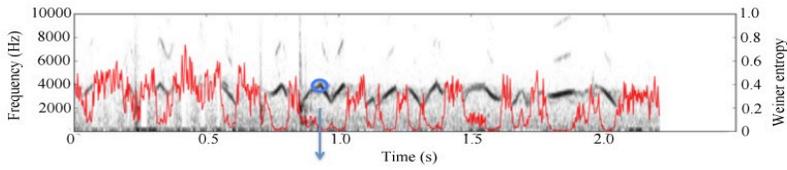
Signal Localization

To locate signals, FinchCatcher estimates Wiener entropy (or spectral flatness) at each time point between high-pass and low-pass thresholds (if present). Wiener entropy is a measure commonly used in digital signal processing to estimate the noisiness of the signal (*Dubnov, 2004; Johnston, 1988; Tchernichovski et al., 2000*). It quantifies how tone-like (as opposed to noise-like) a sound is, by dividing the geometric mean of the power spectrum by the arithmetic mean of the power spectrum. In time windows without signal, the power spectrum should be of high spectral flatness (approaching 1 for white noise), indicating the spectrum has a similar amount of power in all spectral bands. In contrast, the power spectrum of a bird song is of low spectral flatness (approaching 0 for a pure tone), indicating that spectral power is restricted to one single band or a small number of bands, like a mixture of sine waves. Hence, any time window with bird signals will return low value for Wiener entropy, while others containing only

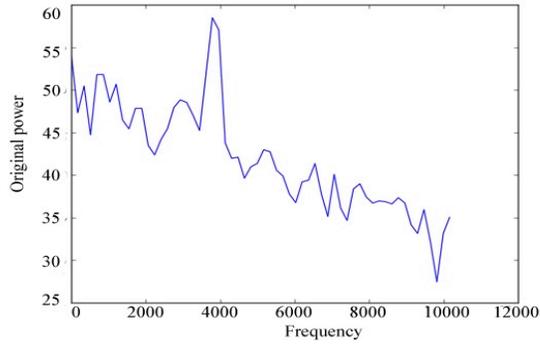
noise will return a high value, as presented in Fig. 1.2. A threshold of 0.2 for Wiener entropy is sufficient to extract all time points with signals for our test data. This threshold is user modifiable in FinchCatcher.

After determining time frames with signals, FinchCatcher extract the frequency trace over time by assessing spectral derivatives in the frequency axis at each time point. Spectral derivatives estimate the change of power in the time-frequency plane. One can measure change along time axis, along frequency axis or at any arbitrary direction. In SAP, spectral derivatives are used to track frequency traces in the spectrogram and thought to be superior to traditional sonogram because they show a sharp transition at bird signals, based on which a contour that more accurately defines the spectral information in a song is obtained (Tchernichovski, Lints, Deregnaucourt, Cimenser, & Mitra, 2004). FinchCatcher calculates spectral derivatives along the frequency axis of time points passing the Wiener entropy filtering, standardizes them, and evaluates the significance of change of power by Z-statistics. A significant change corresponds to a sharp transition from positive to negative values at bird signals (Fig. 1.2), as suggested by Tchernichovski et al. (2004). Frequencies at which such transition occurs are locations of signals on the spectrogram.

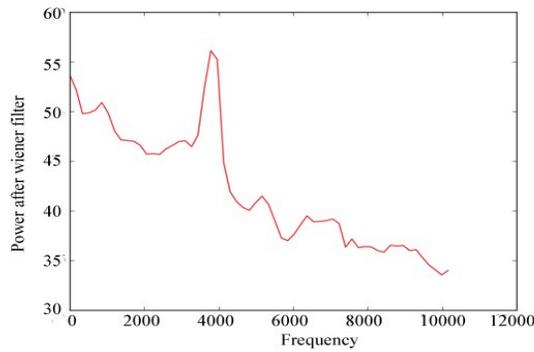
a) Wiener entropy of pure-tone sounds



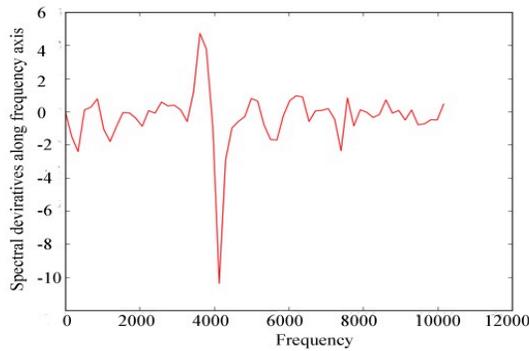
b) Power spectrum of a time frame with bird signals



c) Power spectrum of a time frame with bird signals after noise reduction



d) Spectral derivatives along the frequency axis at the selected time frame



e) Significance of power change by Z-statistics

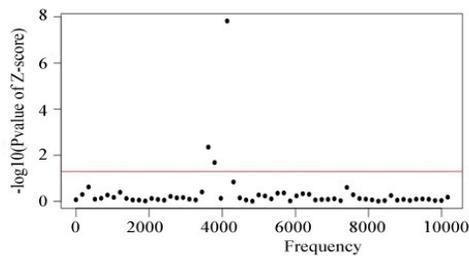


Figure 1.2. Illustration of signal localization for pure-tone sounds

a) Wiener entropy (red lines) associated with a recording of house finch song, superimposed on a spectrogram. Time points with bird signals correspond to low values of Wiener entropy. One signal point in time (at 0.93s and 3962 Hz) is used as an example to illustrate how FinchCatcher retrieves signals through **c)** the application of Wiener filter **b)** to the original spectrogram, **d)** calculating spectral derivatives **e)** and evaluating the significance of power change by Z-statistics. The red line indicate $y = -\log_{10}(0.05) = 1.3$ (remove figure e)

Segmentation to Syllable Units

Acoustic signals can often be segmented into individually distinct units (phonemes, syllables, elements) that are separated by silent intervals. If desired by the user, FinchCatcher can separate songs into syllables. The typical duration of silent intervals is species-specific, and so can be set by the user. For the house finch, adjacent elements tend to be separated by at least 20 ms (Mundinger, 1975); in FinchCatcher, 20ms is the default value.

Spectral Shape Calculation

FinchCatcher measures ten acoustic features in order to characterize and compare songs (Table 1.2).

Table 1.2. Descriptions of ten features

Variables	Description
Average frequency (Hz)-AF	The average frequency of the syllable
Start frequency (Hz)-SF	The frequency at the beginning of the syllable
End frequency (Hz)-EF	The frequency at the end of the syllable
Bandwidth (Hz)-BW	The difference between highest and lowest frequencies of the syllable
Frequency-time excursion length-FL ¹	Arc length of a syllable treated as continuous by connecting amplitude peaks in the spectrogram, using line segments to create a polygonal path
Duration (ms)-DU	Time duration of the syllable
Highest frequency (Hz)-HF	The highest frequency of the syllable
Lowest frequency (Hz)-LF	The lowest frequency of the syllable
Concavity-CO	Number of changes in sign of slope of frequency change throughout the syllable
Average Slope –AS ¹	The average slope of frequency change through the syllable

1. This measurement could be biased by different scales of the time vs. frequency axes; thus I standardized frequency so that its population mean equals that of duration

Syllable Classification

FinchCatcher classifies signals by syllable (or song) shape. Since multiple features are not all measured on the same scale, standardization is necessary to guarantee an unweighted contribution of each feature. In our analysis, all features are standardized by their corresponding standard deviation. FinchCatcher uses the adjusted values of features to classify note or syllable types, based on unsupervised hierarchical clustering using Euclidean distance. FinchCatcher uses the “hclust” function from the R package “stats”, which merges objects in an agglomerative manner. This method makes no assumptions about the number of clusters in the data, so it is appropriate for our analyses because we do not have prior knowledge of syllable types. However, a hierarchical clustering analysis results in not one but several levels of clusters, each nested within the next to form a tree-like topology. Other methods of determining the optimal number of clusters, such as *Akaike Information Criterion* (AIC) or *Bayesian Information Criterion* (BIC) will not work well for most bird songs, as they would recommend far too many clusters. These two methods are optimized for easily clustered data with discrete separation among groups (D. Anderson, 2008), whereas bird songs exhibit more continuous variation. To handle this issue, FinchCatcher provides an alternative way to determine clusters using dynamic tree cuts (Langfelder, Zhang, & Horvath, 2008) (“dynamicTreeCut” in R packages). Whereas the common method uses a constant height threshold, FinchCatcher’s dynamic hybrid cut builds the clusters in a bottom-up manner. First, it respects the order of hierarchical structures in the dendrogram and detects groups of syllables with high similarity by applying a stringent threshold. Second, all previously unassigned objects are reapointed to defined clusters in the first step based on their distances to these clusters. Thus the dendrogram structure is ignored and only dissimilarity information is used (Langfelder et al., 2008).

Application of FinchCatcher

We used FinchCatcher to process all recordings for statistical properties of spectral-shape features. We used default values for all parameters except for duration of silence, which we changed from 20ms to 10ms for swamp sparrow songs. After syllable separation, we used different features for the purpose of syllable classification in the different species. For swamp sparrow songs, we simply used “start and end frequencies”. For chaffinch songs, we selected “concavity” and “frequency-time excursion length”.

For house finch songs, we measured all ten features. For this species’ songs we checked that our comparison algorithm produced reasonable measures of the similarity of syllables and songs by comparing the automated classification with those concluded by human observation. One hundred syllables were classified into 17 distinct types by one of us (CJ), and were re-processed for statistical properties of shape by FinchCatcher. Then we assessed the performance of our algorithm to classify syllable types in several ways. First, we considered whether the algorithm placed visually similar song types in the same cluster. Secondly, we asked whether neighboring types appear more similar to each other than distal pairs in the dendrogram. After comparison validation, FinchCatcher processed 240 recordings of house finch songs for statistical properties, using default values for all parameters.

We assessed similarity in syllable lexicon between localities. To determine the degree of syllable sharing, I used Jaccard’s similarity coefficient (S_j), a method commonly used to measure

similarity for binary data (presence/absence) in acoustic analysis (Krebs, 1989; Podos, Peters, Rudnicki, Marler, & Nowicki, 1992). S_j was calculated as follows:

$$S_j = a / (a + b + c)$$

where a = the number of shared syllables in the lexicons of both localities, b = the number of unique syllables in the first locality, c = the number of unique syllables in the second locality. The similarity score ranges from 0 (not sharing at all) to 1 (perfect sharing). Nevertheless, the maximum possible S_j for the comparison would be < 1 because sizes of syllable lexicon are unlikely to be equal between localities. So I adjusted S_j to account for differences in size as follows:

$$S_j(\text{adj}) = a / (a + b + c - d)$$

where d is the difference in detected lexicon size between localities. I calculated $S_j(\text{adj})$ for each locality compared with all other locality (for a total of 28 comparisons).

Results:

House Finch

Comparison Validation

The algorithm performed accurately relative to human classification (Fig. 1.3). Only two syllables (F-1 and F-2) were misclustered with other types. Closer study found that these two instances partially matched Type R, but were much shorter in duration. This suggests that assigning an increasing weight to “duration” would be advantageous. However, some clusters exhibit greater variation within than between types. In this sample, most of such variable clusters

This dendrogram return a typical tree structure of nested clusters with clear outliers (I-1 and N-1 in Fig. 1.3). Applying a constant height threshold would have generated many clusters containing outliers. Dynamic tree cut identified those outliers first, and then recursively determined clusters depending on the remaining structures in the tree. The resulting clusters were more consistent with human classification than a standard tree cut would have been.

Testing previously published observations

We analyzed house finch recordings from eight geographic localities (including two from New York in different years). Since song types from different localities are quite distinctive (Fig 1.4), we focus on similarity in syllable lexicon. In our sample the average number of syllables is 10.2 per song, for a total of 2448 syllables from 240 songs. Excluding 32 outliers filtered by dynamic tree cut, the rest were classified into 73 types. The result of cluster analysis based on S_j(adj) values of eight localities accords well with published reports (Fig. 1.5). In particular, relatively higher similarities in syllable lexicon are observed between one of the house finch's ancestral localities (California) and its descendent population in New York and Hawaii.

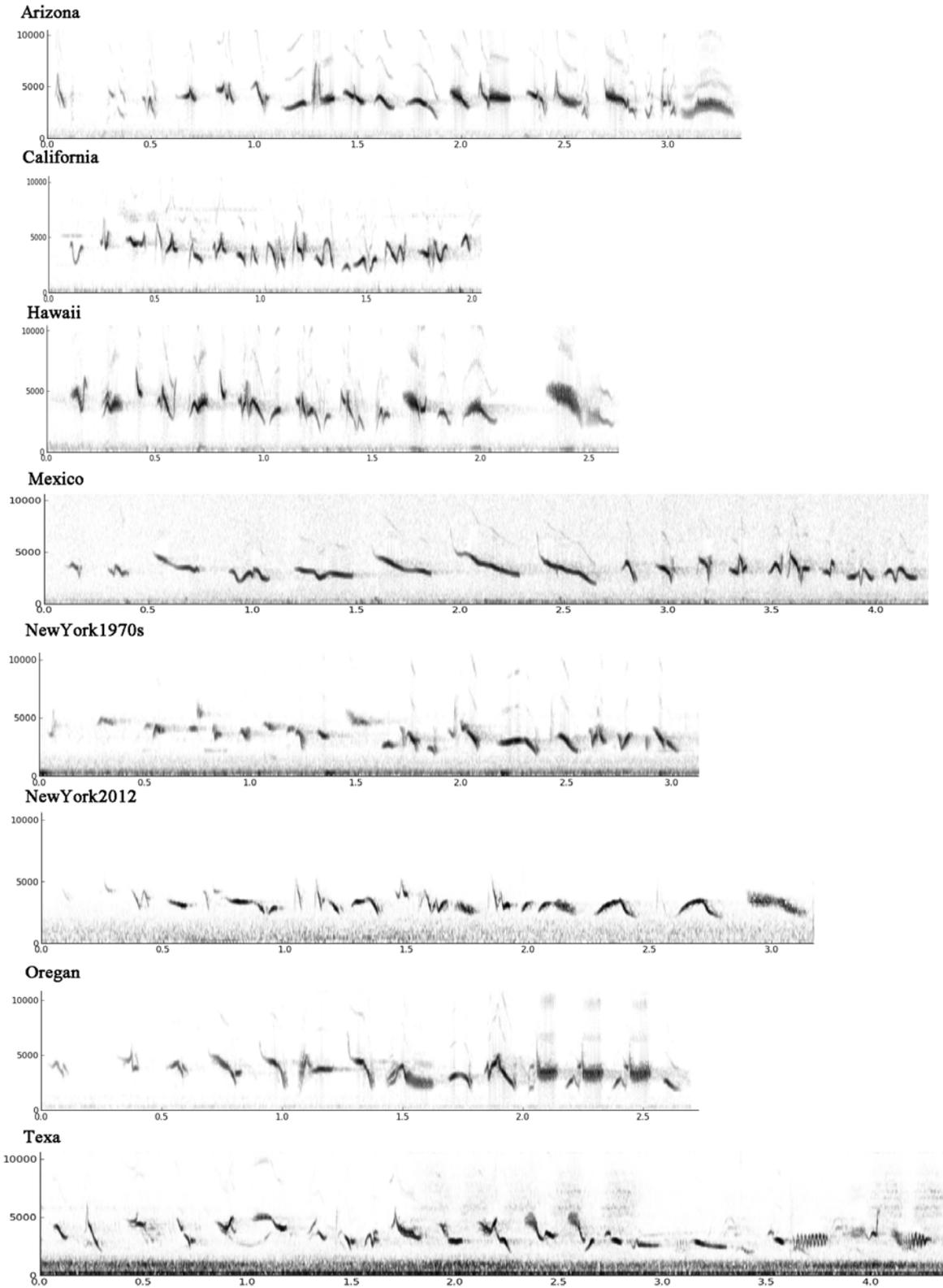


Figure 1.4. Representative songs of house finch from eight localities

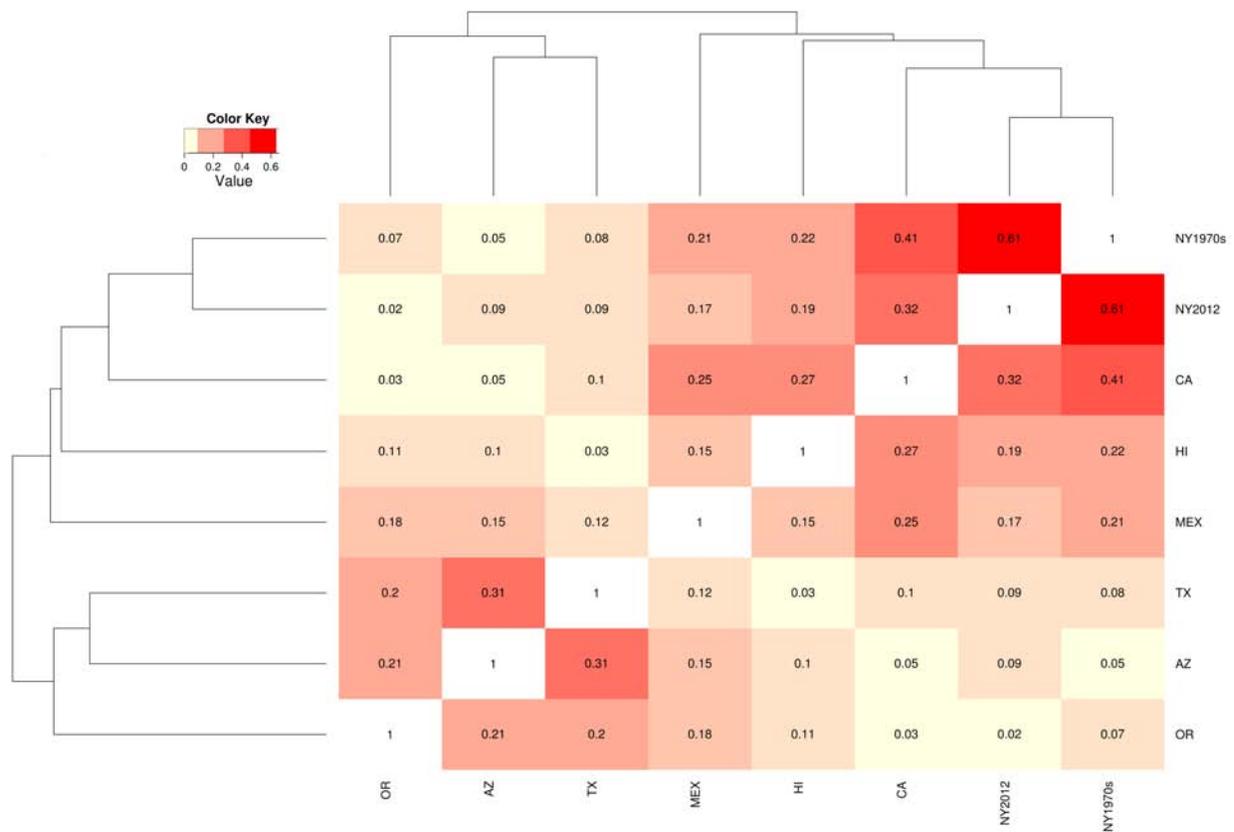


Figure 1.5. Heatmap of UPGMA cluster analysis of house finch songs from eight localities based on similarity in syllable lexicon

Pairwise $s_j(\text{adj})$ values are shown in the cell. Red cells indicate higher similarity while yellow indicate lower similarity in syllable lexicon. Songs from the introduced New York population in different years are the most similar to each other. Relatively higher similarities are found between one of the house finch’s ancestral localities (California) and its descendent population in New York and Hawaii.

Swamp sparrow

Swamp sparrows sing relatively simple songs, repeating two or more simple song elements, or notes (Fig. 1.6a, c and d). A few simple spectral-shape features are sufficient to distinguish distinct types. FinchCatcher retrieved bird signals from recordings despite high acoustic space

densities and phonologically similar elements (Fig. 1.6b). Syllable classification based on cluster analysis of both start and end frequencies of each note returned six distinct types that accord well with human classification (Fig. 1.6e). This result corroborated the known simple syntactical structure of swamp sparrow song. In this species, different combinations of elements characterize different songs. For instance, in the examples analyzed here, element types A, E, and F are unique to the first, second, and third song respectively (Fig. 1.6).

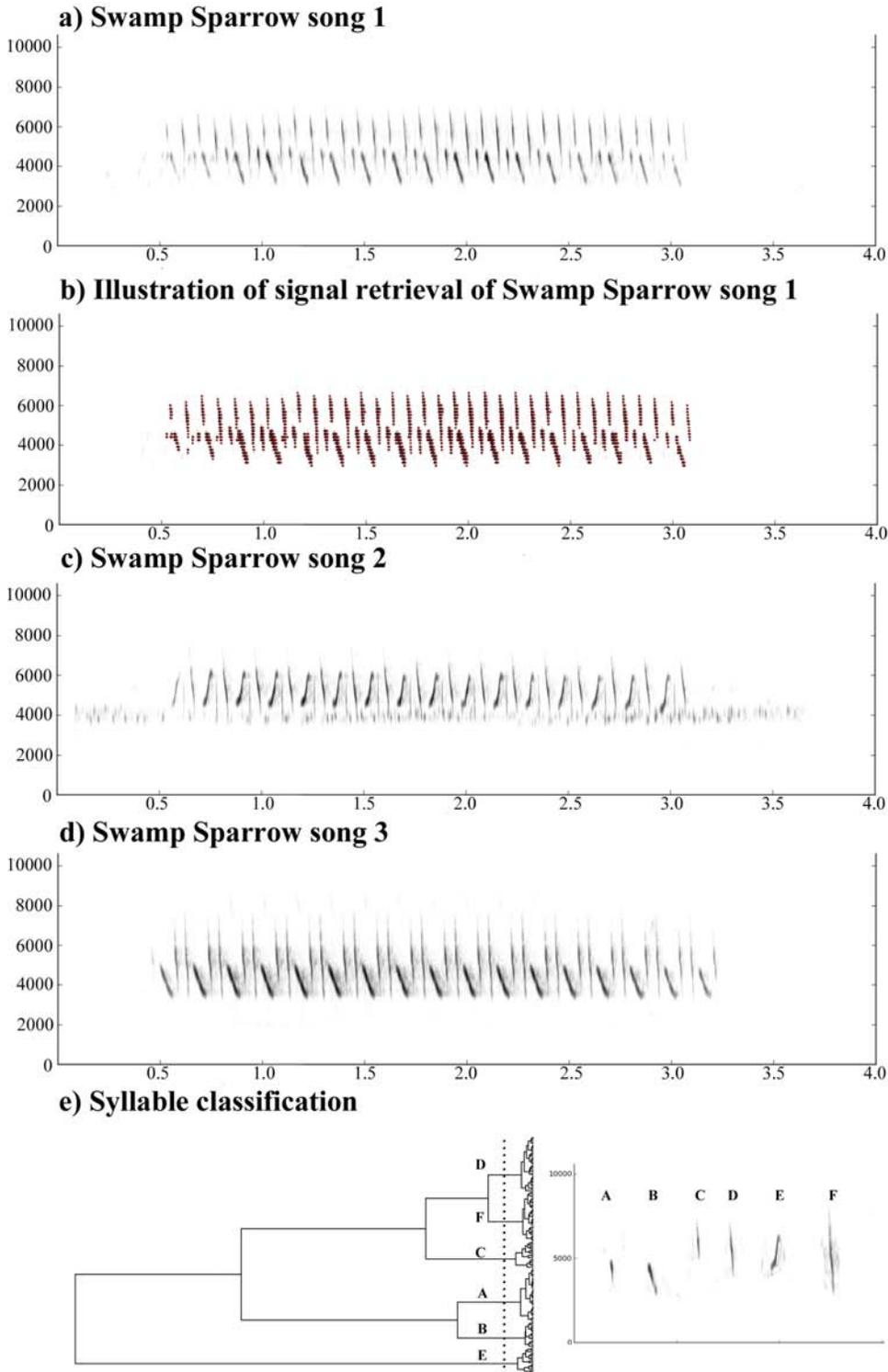


Figure 1.6. Bioacoustical analysis of three swamp sparrow songs

a) the spectrogram of one song; **b)** illustration of signal retrieval of the song from (a); **c-d)** the spectrograms of a second and a third song; **e)** classification of elements from all three songs based on two simple spectral-shape features: start and end frequency.

Chaffinch

Chaffinch songs are more complex than Swamp Sparrow songs: in addition to long buzzy sounds at the end of songs, song elements exhibit high frequency modulation (Fig. 1.7a, c and d). Still, FinchCatcher captured song elements well (Fig. 1.7b). For chaffinch song elements, complex spectral-shape features better identify different syllable types than do simple features. We chose two features, “concavity” and “frequency-time excursion length”, for classification. The result was consistent with subjective manual classification. More importantly, the hierarchical tree reflects element complexity, with simple elements (D, E and F types in Fig. 1.7e) separated from complex elements containing large frequency modulation (G type in Fig. 1.7e).

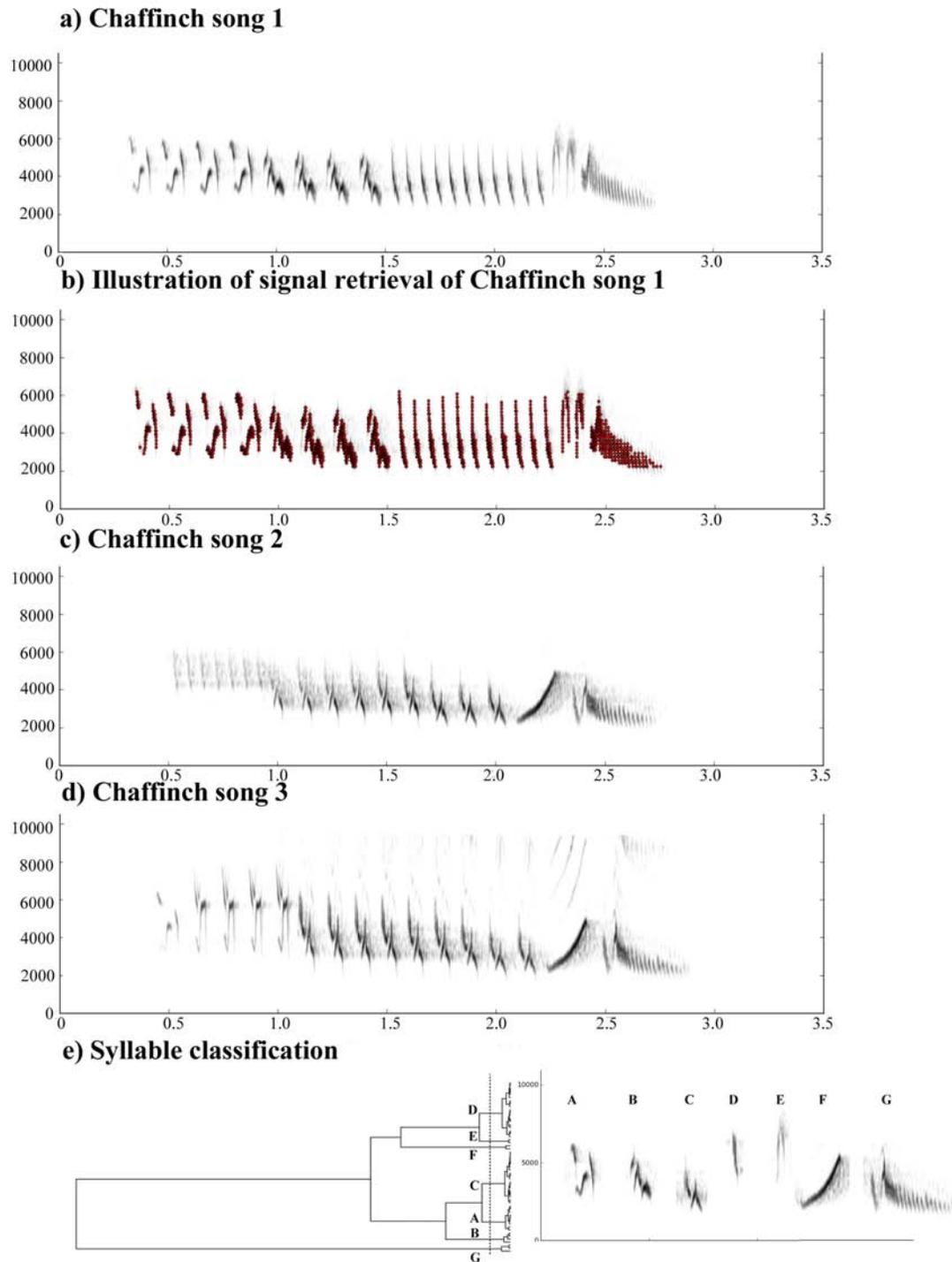


Figure 1.7. Bioacoustical analysis of three Chaffinch songs

a) The spectrogram of one song; **b)** illustration of signal retrieval of the song in (a); **c-d)** the spectrograms of two more songs; **e)** classification of elements from all three songs based on two complex spectral-shape features: concavity and frequency-time excursion length.

Discussion:

We present a computational framework, FinchCatcher, to process field recordings of bird songs automatically, and to quantify and compare songs based on the output. Retrievals of signals from distinct backgrounds generated unbiased measurement of acoustic features across recordings. Dynamic tree cut excluded a large number of outliers, ensuring classification of the same syllables. In addition to commonly used acoustic features, we introduce additional complex features that could be informative in classifying syllable patterns. The measurements of any number of these features can be integrated into a general score for syllable classification. In this way, each acoustic feature can be assessed separately for biological importance in particular species or ecological situations.

To test the effectiveness of our program, we analyzed songs of three species. For swamp sparrow songs, we used two simple features to distinguish continuous song elements. Then we identified syllable types as non-redundant aggregates of several elements and the smallest units of repetition in the song. For example, only four distinct elements (note types) are recovered from one swamp sparrow song, such that this song is composed of seventeen repetitions of syllable type “ACBD”, where A, B, C, and D are note types (Fig. 1.6). FinchCatcher reveals a similar syntactical rule for the second and the third songs, which are sixteen repetitions of syllable type “ED” and seventeen repetitions of syllable type “BFF”, respectively (Fig. 1.6). From this perspective, FinchCatcher not only provides quantitative classification of syllable types, but also information regarding internal organization of notes for each syllable, which will enable researchers to track song differences objectively during song development and evolution. For chaffinch songs, despite complexity in song structure, FinchCatcher reliably classified individual

notes using two features. This analysis preserves well the classic dichotomy between trill and flourish in chaffinch song, a distinction of communicative importance because the birds themselves distinguish songs based on the relative lengths of these two parts (Leitão & Riebel, 2003). Moreover, these two features also permit measurements of song performance, a trait that can influence female choice of males (Ballentine, Hyman, & Nowicki, 2004; J. Byers, Hebets, & Podos, 2010), as well as male competition (Moseley, Lahti, & Podos, 2013).

House finch song has previously been shown to exhibit diversity in many respects, with large sizes of syllable and song types in a given population (Mundinger, 1975). Our analysis corroborated these qualitative observations by returning large numbers of syllable types within individual songs and substantial diversity of syllable between individuals. Interestingly, songs from the introduced New York population in different years are the most similar to each other, suggesting that temporal change may be minor compared with spatial change (Derryberry, 2011). Several new hypotheses are also indicated by this analysis (Fig. 1.5), including that the house finch spread into the U.S. in two events: one into California, and another further eastward into Texas and Arizona; and that it is this second population from which birds in the northwest are descended. Larger sample sizes of broader geographical coverage would be necessary to test these hypotheses rigorously. Finally, known points of divergence (in this sample, 1940 between California and New York; and differences between 1975 and 2012 in New York) can provide estimates of a “cultural clock”, analogous to a molecular clock in genetic phylogenies.

The main value of FinchCatcher is to provide a quick and quantitative way of classifying song elements, facilitating the study of various features of interest in bird song, even if recorded in the field. Our program is a particularly effective tool to analyze songs at different levels, as a variety of features can be investigated independently or in interaction (a list that will grow in subsequent versions). FinchCatcher gathers more information than the values of assessed features employed in the sample analyses here. For instance, it retains syllable order and measures the time interval between syllables. Such information, as part of a chosen subset of features, can be fruitful when approaching new biological questions. For instance, time intervals and frequency shifts between ending points of the precedent syllables and starting points of the following syllables can be assessed as a factor of song performance (Cardoso, 2014; Geberzahn & Aubin, 2014).

The body of software available for acoustic analysis is growing rapidly. FinchCatcher, while being similar to some recent programs in the underlying principle of automated measurement, is a worthwhile complement to existing software for well-studied species, and is more suitable for large dataset or for field recordings. First, unlike many programs optimized for recordings in laboratory conditions, FinchCatcher provides an efficient way to address ambient noise in field samples. Second, retrieval of signals is based on a sharp change of power between signal and background. Such contrast might not be visually apparent without careful calibration of the spectrogram (Zollinger, Podos, Nemeth, Goller, & Brumm, 2012). Third, FinchCatcher features a new method of classification using dynamic tree cut, which will be especially useful for large datasets and complex topologies. As the number of recordings increases, the majority of identified syllables will likely belong to a small number of types. Unsupervised clustering techniques would result in nested structures, based on which a static tree cut will be unlikely to

define appropriate clusters. The height threshold for dynamic tree cut should be determined by the user. This threshold and other features can be altered or recombined without changing the subsequent framework for analysis. Likewise, although in our sample analyses we assigned unbiased weights to all analyzed features, the weighting scheme can be modified by the user according to preference or prior knowledge. Future developments in FinchCatcher will incorporate our accumulating knowledge of the acoustic properties of more species' vocalization. In addition, we would like to extend applicability of FinchCatcher to broadband and harmonic songs. We will develop a database to store analyzed syllable types, against which new syllables can be compared to determine if they are any of known types or novel types. In sum, FinchCatcher provides speed, flexibility, and biological relevance, whether in the furtherance of existing research questions or in development of new ones.

References:

- Aldrich, J. W., & Weske, J. S. (1978). Origin and evolution of the eastern House Finch population. *The Auk*, 95(3), 528-536.
- Anderson, D. (2008). *Model Based Inference in the Life Sciences: A Primer on Evidence*. Springer.
- Anderson, S. E., Dave, A. S., & Margoliash, D. (1996). Template-based automatic recognition of birdsong syllables from continuous recordings. *The Journal of the Acoustical Society of America*, 100(2), 1209-1219.
- Angle, G., & Coskun, H. (2014). A complete syllable dictionary for *Serinus canarius*. *Ecological Informatics*, 20, 67-75.
- Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. *Animal Behaviour*, 77(4), 973-978.
- Ballentine, B., Hyman, J., & Nowicki, S. (2004). Vocal performance influences female response to male bird song: an experimental test. *Behavioral Ecology*, 15(1), 163-168.
- Belthoff, J. R., & Gauthreaux Jr, S. A. (1991). Partial migration and differential winter distribution of House Finches in the eastern United States. *Condor*, 93(2), 374-382.
- BirdLife International. (2013). *Carpodacus mexicanus*. *IUCN Red List of Threatened Species*. Version 2013.1.
- Bitterbaum, E., & Baptista, L. F. (1979). Geographical variation in songs of California House Finches (*Carpodacus mexicanus*). *The Auk*, 96(3), 462-474.
- Boll, S. (1979). Suppression of acoustic noise in speech using spectral subtraction. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 27(2), 113-120.
- Bookstein, F. L. (1996). Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology*, 58(2), 313-365.
- Brumm, H., & Slabbekoorn, H. (2005). Acoustic communication in noise. *Advances in the Study of Behavior*, 35, 151-209.
- Byers, B. E., Belinsky, K. L., & Bentley, R. A. (2010). Independent Cultural Evolution of Two Song Traditions in the Chestnut-Sided Warbler. *The American Naturalist*, 176(4), 476-489.
- Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour*, 79(4), 771-778.
- Cardoso, G. C. (2014). Studying the silent side of birdsong. *BMC Biology*, 12(1), 62.
- Clark, C. W., Marler, P., & Beeman, K. (1987). Quantitative analysis of animal vocal phonology: an application to swamp sparrow song. *Ethology*, 76(2), 101-115.
- Cortopassi, K. A., & Bradbury, J. W. (2000). The comparison of harmonically rich sounds using spectrographic cross-correlation and principal coordinates analysis. *Bioacoustics*, 11(2), 89-127.

- Daou, A., Johnson, F., Wu, W., & Bertram, R. (2012). A computational tool for automated large-scale analysis and measurement of bird-song syntax. *Journal of Neuroscience Methods*, 210(2), 147-160.
- Derryberry, E. P. (2010). Male response to historical and geographical variation in bird song. *Biology Letters*, 7, 57-59.
- Dubnov, S. (2004). Generalization of spectral flatness measure for non-gaussian linear processes. *Signal Processing Letters, IEEE*, 11(8), 698-701.
- Elliott, J. J., & Arbib, R. S. J. (1953). Origin and status of the house finch in the eastern United States. *The Auk*, 70, 31-37.
- Ephraim, Y., & Malah, D. (1985). Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 33(2), 443-445.
- Fantana, A. L., & Kozhevnikov, A. (2014). Finding motifs in birdsong data in the presence of acoustic noise and temporal jitter. *Behavioral Neuroscience*, 128(2), 228.
- Fee, M. S., & Scharff, C. (2010). The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR Journal*, 51(4), 362-377.
- Fernández-Juricic, E., Poston, R., De Collibus, K., Morgan, T., Bastain, B., Martin, C., . . . Treminio, R. (2005). Microhabitat selection and singing behavior patterns of male house finches (*Carpodacus mexicanus*) in urban parks in a heavily urbanized landscape in the Western US. *Urban Habitats*, 3(1), 49-69.
- Frigo, M., & Johnson, S. G. (2005). The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2), 216-231.
- Geberzahn, N., & Aubin, T. (2014). Assessing vocal performance in complex birdsong: a novel approach. *BMC Biology*, 12(1), 58.
- Giret, N., Roy, P., Albert, A., Pachet, F., Kreutzer, M., & Bovet, D. (2011). Finding good acoustic features for parrot vocalizations: The feature generation approach. *The Journal of the Acoustical Society of America*, 129(2), 1089-1099.
- Grover, R., & Hwang, P. (1992). Introduction to random signals and applied Kalman filtering. *Wiley, New York*.
- Gustafsson, S., Martin, R., Jax, P., & Vary, P. (2002). A psychoacoustic approach to combined acoustic echo cancellation and noise reduction. *Speech and Audio Processing, IEEE Transactions on*, 10(5), 245-256.
- Hamilton, T. (1991). Seasonal movement of house finches in the midwest. *North American Bird Bander*, 16, 119-122.
- Hultsch, H., & Todt, D. (1981). Repertoire sharing and song-post distance in nightingales (*Luscinia megarhynchos* B.). *Behavioral Ecology and Sociobiology*, 8(3), 183-188.
- Johnston, J. D. (1988). Transform coding of audio signals using perceptual noise criteria. *Selected Areas in Communications, IEEE Journal on*, 6(2), 314-323.

- Kakishita, Y., Sasahara, K., Nishino, T., Takahasi, M., & Okanoya, K. (2007). Pattern extraction improves automata-based syntax analysis in songbirds *Progress in Artificial Life* (pp. 320-332): Springer.
- Khanna, H., Gaunt, S., & McCallum, D. (1997). Digital spectrographic cross-correlation: tests of sensitivity. *Bioacoustics*, 7(3), 209-234.
- Koenig, W., Dunn, H., & Lacy, L. (1946). The sound spectrograph. *The Journal of the Acoustical Society of America*, 18(1), 19-49.
- Kroodsma, D. E., Miller, E. H., & Ouellet, H. (1982). *Acoustic Communication in Birds* (Vol. 2): Academic press New York.
- Lachlan, R. F. (2007). Luscinia: a bioacoustics analysis computer program. Version 1.0 [Computer program].
- Lachlan, R. F., Verhagen, L., Peters, S., & Cate, C. (2010). Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taenopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *Journal of Comparative Psychology*, 124(1), 92-108.
- Lahti, D. C., Moseley, D. L., & Podos, J. (2011). A tradeoff between performance and accuracy in bird song learning. *Ethology*, 117(9), 802-811.
- Laiolo, P., Vögeli, M., Serrano, D., & Tella, J. L. (2007). Testing acoustic versus physical marking: two complementary methods for individual-based monitoring of elusive species. *Journal of Avian Biology*, 38(6), 672-681.
- Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*, 24(5), 719-720.
- Leitão, A., & Riebel, K. (2003). Are good ornaments bad armaments? Male chaffinch perception of songs with varying flourish length. *Animal Behaviour*, 66(1), 161-167.
- Leonardo, A. (2004). Experimental test of the birdsong error-correction model. *Proceedings of the National Academy of Sciences of the United States of America*, 101(48), 16935-16940.
- Marler, P., & Pickert, R. (1984). Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Animal Behaviour*, 32(3), 673-689
- Meyer, J., Dentel, L., & Meunier, F. (2013). Speech Recognition in Natural Background Noise. *PloS One*, 8(11), e79279.
- Moseley, D. L., Lahti, D. C., & Podos, J. (2013). Responses to song playback vary with the vocal performance of both signal senders and receivers. *Proceedings of the Royal Society B: Biological Sciences*, 280(1768), 20131401.
- Mundinger, P. (1975). Song dialects and colonization in the house finch, *Carpodacus mexicanus*, on the east coast. *Condor*, 77, 407-422.
- Mundinger, P., & Lahti, D. (2014). Quantitative integration of genetic factors in the learning and production of canary song. *Proceedings of the Royal Society B: Biological Sciences*, 281(1781), 20132631.

- Nagy, C. M., & Rockwell, R. F. (2012). Identification of individual eastern Screech-Owls *Megascops asio* via vocalization analysis. *Bioacoustics*, *21*(2), 127-140.
- Payne, R. B. (1996). Song traditions in indigo buntings: origin, improvisation, dispersal, and extinction in cultural evolution. *Ecology and Evolution of Acoustic Communication in Birds*, 198-220.
- Podos, J., Huber, S. K., & Taft, B. (2004). Bird song: the interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics*, *35*, 55-87.
- Podos, J., Lahti, D. C., & Moseley, D. L. (2009). Vocal performance and sensorimotor learning in songbirds. *Advances in the Study of Behavior*, *40*, 159-195.
- Potamitis, I., Ntalampiras, S., Jahn, O., & Riede, K. (2014). Automatic bird sound detection in long real-field recordings: Applications and tools. *Applied Acoustics*, *80*, 1-9.
- Read, A. F., & Weary, D. M. (1990). Sexual selection and the evolution of bird song: a test of the Hamilton-Zuk hypothesis. *Behavioral Ecology and Sociobiology*, *26*(1), 47-56.
- Saar, S., & Mitra, P. P. (2008). A technique for characterizing the development of rhythms in bird song. *PLoS One*, *3*(1), e1461.
- Somervuo, P., Harma, A., & Fagerlund, S. (2006). Parametric representations of bird sounds for automatic species recognition. *Audio, Speech, and Language Processing, IEEE Transactions on*, *14*(6), 2252-2263.
- Tachibana, R. O., Oosugi, N., & Okanoya, K. (2014). Semi-Automatic Classification of Birdsong Elements Using a Linear Support Vector Machine. *PLoS One*, *9*(3), e92584.
- Tchernichovski, O., Lints, T., Deregnaucourt, S., Cimenser, A., & Mitra, P. (2004). Studying the song development process: rationale and methods. *Annals of the New York Academy of Sciences*, *1016*(1), 348-363.
- Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour*, *59*(6), 1167-1176.
- Thorpe, W. H. (1954). The process of song-learning in the chaffinch as studied by means of the sound spectrograph. *Nature* *173*, 465 - 469
- Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs*. *Ibis*, *100*(4), 535-570.
- Thorpe, W. H. (1961). *Bird-song: the Biology of Vocal Communication and Expression in Birds*. Oxford, England: University Press
- Vary, P., & Martin, R. (2006). *Digital Speech Transmission: Enhancement, Coding and Error Concealment*: John Wiley & Sons.
- Wellock, C. D., & Reeke, G. N. (2012). Quantitative tools for examining the vocalizations of juvenile songbirds. *Computational Intelligence and Neuroscience*, *2012*, 2.
- Wu, W., Thompson, J. A., Bertram, R., & Johnson, F. (2008). A statistical method for quantifying songbird phonology and syntax. *Journal of neuroscience methods*, *174*(1), 147-154.

Zollinger, S. A., Podos, J., Nemeth, E., Goller, F., & Brumm, H. (2012). On the relationship between, and measurement of, amplitude and frequency in birdsong. *Animal Behaviour*, 84(4), e1-e9.

CHAPTER 2:

Quantifying spatial variation in house finch (*Haemorhous mexicanus*) song

Abstract:

The study of geographic variation in acoustic behavior can provide insight into evolutionary processes in nature that interact to produce highly diverse vocal signals. I examined geographic variation in the songs of the house finch (*Haemorhous mexicanus*), a non-territorial passerine that has expanded its range dramatically in the past several decades following an introduction event from California to New York. Significant behavioral changes have been reported during the subsequent growth of the Eastern population, such as migration, which might have a profound impact on the geographic pattern of song. I analyzed songs from 96 individuals in western Long Island, an area overlapping the putative site of introduction of the founding propagule for the Eastern house finch population. While overall no similar song types were observed across individuals, I found significantly greater syllable and phrase sharing between nearby birds (less than 12 km apart) than those at greater distances. Degrees of such sharing differed regionally, resulting in localized syllable compositions, though there was no evidence for discrete syllable dialects. Discriminant function analysis revealed that three acoustic variables—number of concavities, minimum frequency, and duration of silent intervals—were related to recording location. Spatial differentiation was neither simply clinal change with geographic distance, nor discrete dialects. I suggest several potential explanations for the observed geographic pattern, including a multiple founder effect, differential and migration dispersal pattern, and developmental stress.

Keywords: geographic variation, migration, dispersal, developmental stress, house finch, song diversity, cultural evolution

Introduction:

Understanding the origins and maintenance of vocal diversity is fundamentally important to acoustic biology (Podos & Warren, 2007). (Campbell et al., 2010). Geographical variation in vocalizations has been reported in many songbirds (Catchpole & Slater, 2008; Cunningham, Baker, & Boardman, 1987; Mundinger, 1982; D. A. Nelson & Soha, 2004; Podos, 2007; Warren & Nelson, 2002; Wright, Rodriguez, & Fleischer, 2005). Acoustic variation can be attributed largely to vocal imitation, a learning behavior through which many birds learn to vocalize. In some laboratory experiments, degraded acoustic structure was observed in birds in the absence of learning models (Marler & Tamura, 1964). This seminal observation highlights the role that learning can play in the generation of vocal diversity. New songs or song elements may be results of learning “errors” or improvisations, and imitation can facilitate transmission of these novel patterns of vocal structure to younger conspecifics and therefore contribute to song “cultural evolution” through the combined effects of drift and selection (Baker & Cunningham, 1985; Payne, 1981), including cultural drift (Lemon, 1975), genetic drift (Podos, Huber, & Taft, 2004), cultural selection (Slabbekoorn & Smith, 2002), natural selection (Nowicki, Westneat, & Hoese, 1992) and sexual selection (Panhuis, Butlin, Zuk, & Tregenza, 2001). These effects are unlikely to be mutually exclusive. Instead, they are probably all involved in the evolutionary process with varying degrees of importance under different scenarios. Some have argued that adaptive processes are necessary to facilitate formation and maintenance of geographic pattern (Baker & Cunningham, 1985), while others have suggested that spatial diversification can

emerge as an epiphenomenon or byproduct of other, selectively neutral, processes (Andrew, 1962).

Dialect, a particular form of geographic acoustic variation, has been well described across a range of geographic scales (Leader, Wright, & Yom-Tov, 2000; Marler & Tamura, 1964; Mundinger, 1975). Song dialects may be formed in adjacent populations of birds, with typically sharp boundaries showing steep clines with stepped variation, instead of gradual and shallow clines or random variation (Podos & Warren, 2007). No dialect has yet been found in a species without imitative vocal learning (Podos & Warren, 2007). The other factor that appears to contribute to dialect formation is dispersal pattern, based on the assumptions that dialects may emerge if male birds are philopatric, or if they do disperse but prefer to settle with populations with similar song patterns (Marler & Tamura, 1964). The interplay of learning strategies, dispersal patterns, and subsequent geographic song patterns can be highly variable, even within a single species (Kroodsma, 1996; Slater, 1989). Therefore, given the difficulty of assessing the influences of multiple interacting factors, a known population history, such as a founder event, can be of great help in distinguishing the possible mechanisms, especially through the provision of a timescale (Podos et al., 2004; Wilkins, Seddon, & Safran, 2013).

The house finch (*Haemorhous mexicanus*) is a non-territorial and socially monogamous passerine (Hill, 1993). Native to southwestern North America, the house finch was introduced onto Long Island, New York, around 1940 (Elliott & Arbib Jr, 1953), and has since rapidly expanded its range; it currently occupies virtually the entire continental United States (Sauer et

al., 2008). Previous studies have suggested large morphological and behavioural differences between eastern and western house finches (Able & Belthoff, 1998; Aldrich & Weske, 1978; Badyaev & Hill, 2000), including the rapid evolution of migratory behaviour in the introduced Eastern population. The house finch was once considered non-migratory, as the native populations of western North America are reported to be sedentary (Hill, 1993; Thompson, 1960). Eastern house finches, however, have been found to be partial migrants: up to half of the eastern population migrates, and migrants have been observed to remain in the breeding areas in some winters (Able & Belthoff, 1998). Wang et al. (2003) showed significant genetic differentiation between eastern and western house finch populations using amplified fragment length polymorphism (AFLP), but found no substantial loss of AFLP diversity in the East, suggesting that the rapid growth of the eastern population has neutralized bottleneck effects of the founder event(s).

Typically, the male house finch sings a long, rambling, hoarse warble composed of a series of short notes from the large repertoire of syllables of each individual (Bitterbaum & Baptista, 1979; Mundinger, 1975; Pytte, 1997; Thompson, 1960). The song learning mechanism of the male house finch has been briefly introduced (Miller, 1921). Juvenile learning usually occurs prior to natal dispersal (Mundinger, 1975). Male house finches do not defend resources or territories, so song does not function in territory defense (Hill, 1993). However, male house finch song does function in mate attraction: males sing at faster rates when they approach females (Bitterbaum & Baptista, 1979), females prefer longer songs and songs performed at a faster rate (Nolan & Hill, 2004), and male house finches singing songs with elaborate features sire more extra-pair young (Mennill, Badyaev, Jonart, & Hill, 2006). Geographic variation of

house finch song has been reported in a few areas (Bitterbaum & Baptista, 1979; Mundinger, 1975; Tracy, Zasadny, Erickson, & Siemers, 2009). Geographic variation at the syllable level is well documented as well (Tracy & Baker, 1999; Tracy et al., 2009). In general, male song structure has been suggested to differ between eastern and western populations: eastern songs might be more stereotyped and less syntactically variable than western songs, and the eastern syllable lexicon size might be smaller than in the west (Bitterbaum & Baptista, 1979; Mundinger, 1975; Pytte, 1997).

In this study, I investigated microgeographic variation in house finch song in western Long Island. In general, given the fact that Dr. Mundinger reported dialect in Westchester County, New York (~25km away from this study's focal area) (Mundinger, 1975) and dispersal and migration might increase cultural diversity by increasing the exchange of song elements, I expected to observe dialect areas too, within which, however, reduction in the degree of sharing of songs and song elements were found. More specifically, I had three aims:

(1) To assess the extent of sharing of songs and song elements among individuals. The study area was one of the introduction sites of the eastern population of the house finch in 1940. Song element sharing among individuals would likely have been high shortly after introduction due to a founder effect. However, I predicted that song sharing would be low in recent years, for two reasons. First, diversity in song elements should increase along with the rapid growth of the eastern population, neutralizing the founder effect. As population density increased, a greater number of tutors would be available to each young learner, increasing the variety of song elements likely to be learned. Second, an increase in acoustic diversity would be further augmented by the documented increase in migratory behavior, as migratory birds have been

shown to overproduce song and syllable types compared to their sedentary conspecifics in a number of songbirds (D. Nelson, Marler, & Morton, 1996). To achieve this aim, I examined pairwise similarities of song and syllable repertoires between individuals.

(2) To describe the geographic pattern of song and song element variation. Specifically, I determined the form of geographic variation prevalent in my study area: dialect variation, random variation or gradual clines. Distinct song dialects with sharp boundaries were reported previously in Westchester County, New York (~25km away from this study's focal area) (Mundinger, 1975). Within a given dialect area, the songs of the local population were stable over a short period of time, and clearly distinguishable from neighboring dialects (Mundinger, 1975). The present study area likely displayed a similar pattern in the 1970s (Mundinger, unpublished data). However, dialect variation might not be well maintained over time, due to increased migratory activity and greater dispersal distance (Paradis, Baillie, Sutherland, & Gregory, 1998), both of which factors would in turn complicate any female strategy of choosing local males based on their songs. Therefore, I expected a mixture or intermediate pattern between dialect and clines in the contemporary New York population. To test this, I examined variation in song element similarity (at the level of syllable and at the level of bigram phrase) across the study area, and performed cluster analysis of individuals based on the syllable composition of their songs.

(3) To test if particular vocal parameters displayed significant correspondence with the general geographic pattern. A number of studies have shown that different vocal parameters or portions of songs can serve different functions, and change at different paces and under the influence of different factors (Sung & Handford, 2006; Williams, Levin, Norris, Newman, & Wheelwright, 2013). For instance, some song features might serve mainly to announce species or

individual identity, while others might indicate male quality. Although links between specific functional mechanisms and geographic patterns in bird song have proven difficult to demonstrate, an association between specific vocal parameters and geographic variation might provide a first step towards the discovery of underlying mechanisms that shape the local landscape of acoustic variation. For this purpose, I performed discriminant function analysis on several features of the vocal profile of individuals and related the results to geographic variation.

Methods:

Study Area and Song Sample

I analyzed songs of wild house finches recorded in a roughly 34 km (north to south) by 40 km (east to west) area in western Long Island, New York: from Long Beach in the south (40.589° N 73.667° W) to Port Washington in the north (40.895° N 73.634° W), and from Bay Ridge in the west (40.631° N 74.036° W) to East Meadow in the east (40.729° N 73.566° W) (Figure 2.1).

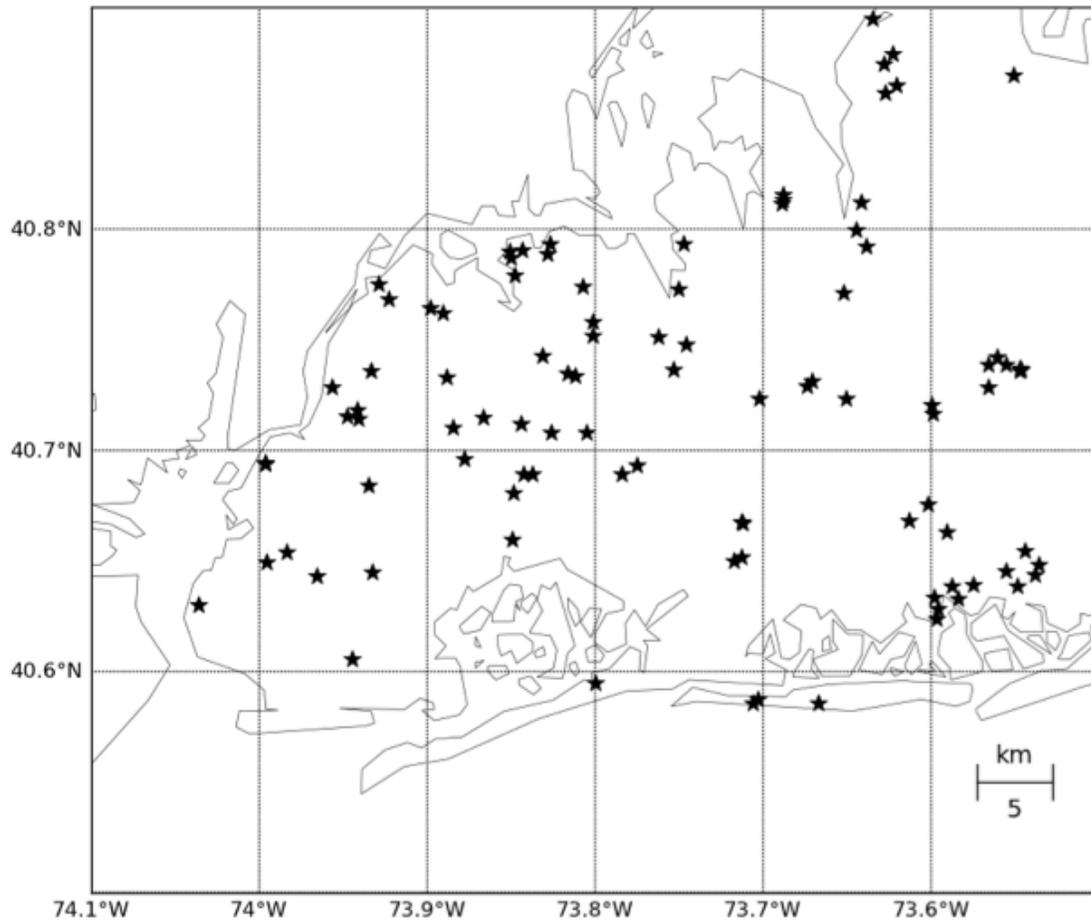


Figure 2.1. Sampling localities of 96 individuals in western Long Island

The songs were recorded in the field in April-June 2012 by Frances Geller. Individuals were recorded until they stopped singing or flew away. Each site was visited only once to minimize the chances of recording the same individual repeatedly. Within each site, each individual recorded was at least 100 meters away from any other individual recorded. Tracks with multiple song bouts were excluded from analysis, except in cases where a second bird was reported as singing as soon as the first bird flew away. I analyzed only individuals with five or more high-quality song recordings: 96 individuals (mean \pm SD of 10.2 \pm 4.5 songs per individual; the total

number of songs is 981). Recordings were saved at a sample rate of 44100 Hz, sample width of 16 bits, through one channel (mono).

Syllable Classification

A syllable is defined here as a contiguous signal separated from others within a song by at least 20 ms (Mundinger, 1975). Identifying syllable types in house finches has proved not amenable to an approach based on the visual inspection of song spectrograms, because variation in syllable structure appears to be continuous, which would result in increased subjectivity in visual classification and probably a decrease in reproducibility among different investigators. I therefore based syllable classification on an automated similarity measurement procedure (FinchCatcher) that compares a number of shape-related acoustic features to generate overall similarity scores between syllables (see Chapter 1). Then I performed unsupervised hierarchical clustering (“hclust” function from the R package “stats”) using Euclidean distance, on this similarity matrix. This method makes no assumptions about the number of clusters in the data, so it is appropriate for my analyses because I do not have prior knowledge of syllable types. However, a hierarchical clustering analysis results in not one but several levels of clusters, each nested within the next to form a tree-like topology. Many methods of determining the optimal number of clusters, such as *Akaike Information Criterion* (AIC) or *Bayesian Information Criterion* (BIC) did not work well for my data, as they returned far too many clusters for subsequent analysis. These two methods are optimized for easily clustered data with discrete separation among groups (Anderson, 2008), whereas the bird songs in my dataset exhibited more continuous variation. To handle this issue, I adopted an alternative method of determining clusters, dynamic tree cut (Langfelder, Zhang, & Horvath, 2008) (“dynamicTreeCut” in R

packages). In contrast to the method that applies a constant height threshold to cut dendrograms into clusters, the dynamic tree cut builds the clusters in a bottom-up manner. First, it respects the order of hierarchical structures in the dendrogram and detects groups of syllables with high similarity by applying a stringent threshold. Second, all previously unassigned objects are reapointed to defined clusters in the first step based on their distances to these clusters. Thus the dendrogram structure is ignored and only similarity information is used (Langfelder et al., 2008). The results of classification were examined by eye.

To determine the number of songs that would be necessary to recover an individual's entire repertoire of syllables, I plotted the cumulative number of distinct syllable types observed against the n^{th} song sampled for each individual, including only those individuals in the dataset with more than 20 recorded songs.

Song Sharing Between Individuals

Sharing of the same song type between individuals was determined using the 75% similarity criterion (Mundinger, 1975; Tracy & Baker, 1999). Specifically, I calculated the edit (Levenshtein) distance (Ristad & Yianilos, 1998) between the pair of song sequences, by which the minimum number of operations required to transform one string into the other were returned. Then, the similarity score of the pair of sequences was estimated as follows:

$$S=1-(E-D)/L,$$

where E = edit distance between two sequences, D = the difference in numbers of syllables between two songs and L = the length of the shorter of the two songs. This equation allowed me to adjust similarity scores to account for the difference in lengths of songs, in order to ensure that

a given song would not be misclassified as a different type than a more or less complete version of the same song. The paired songs with $S > 0.75$ were considered shared. I calculated the highest similarity scores between any songs in the recorded repertoire of both paired individuals.

Syllable Sharing Between Individuals versus Distance

To determine the degree of syllable sharing among birds, I used Jaccard's similarity coefficient (S_j), a method commonly used to measure similarity for binary data (presence/absence) in acoustic analysis (Krebs, 1989; Podos, Peters, Rudnicki, Marler, & Nowicki, 1992). S_j was calculated as follows:

$$S_j = a / (a + b + c)$$

where a = the number of shared syllables in both individual repertoires, b = the number of unique syllables in the first individual, c = the number of unique syllables in the second individual. The similarity score ranges from 0 (not sharing at all) to 1 (perfect sharing).

Nevertheless, the maximum possible S_j for the comparison would be < 1 because sizes of syllable repertoires are unlikely to be equal between individuals. So I adjusted S_j to account for differences in sample size as follows:

$$S_j(\text{adj}) = a / (a + b + c - d)$$

where d is the difference in detected repertoire size between individuals. I calculated $S_j(\text{adj})$ for each individual compared with all other individuals (for a total of 4560 comparisons). I performed a Mantel test (Legendre & Fortin, 1989) to determine whether similarity in the sharing of syllable repertoires declined with distance. I specifically examined the extent of

sharing of syllable repertoires within a bounded interval of geographic distances. By this means, I determined whether and at what geographic distance syllable repertoire similarity decreased dramatically. I also corrected between-individual similarity for geographic distance (using the residual term in the mantel model) to assess difference in syllable sharing between individuals regardless of distance. To test for patterns of localized syllable composition, I performed the “unweighted pair group method with arithmetic mean” (UPGMA) (Sokal, 1958) to group individuals based on similarity of their syllable repertoires. Similarities between individuals were estimated by Pearson correlation on the presence/absence of all syllable types. All analyses were carried out using R package “ecodist”.

Phrase Sharing Between Individuals versus Distance

The mean±SD of expected chances of observing detected shared bigram phrases among individuals was 0.001 ± 0.002 . Taking into account the numbers of bigrams per individual (mean±SD of 161.3 ± 62.2), the mean of numbers of occurrence of shared bigram was lower than 1 (mean±SD of 0.11 ± 0.08). Therefore, observations of shared bigrams between individuals were unlikely due to random event.

I used the same methods to measure similarity of phrases (i.e. sequences of syllables) between individuals as I did for syllables. Specifically, I was interested in sharing of bi-gram (two syllable) sequences. The equation for this purpose is as follows:

$$SSj(adj)=a/(a+b+c-d)$$

Where a = the number of shared bi-gram sequences in both individual repertoires, b = the number of unique bi-gram sequences in the first individual, c = the number of unique bi-gram

sequences in the second individual, and d = the difference in the number of bi-gram sequences in the two individuals due to differences in song length. I studied bi-gram sequences since I performed a pilot study that demonstrated the rarity of sharing larger phrases (out of all pairwise comparisons between individuals ($N=4,560$), less than 0.2% had any shared trigrams).

Some studies have indicated that the basic unit of song learning could be the phrase (e.g. Lemon, 1975; D. A. Nelson, Marler, & Palleroni, 1995). Therefore, it is worth examining whether phrases are more likely to be shared between individuals than expected by chance. My sample recordings contain an average of 10 songs per individual, with a song having an average length of 12 syllables; thus there are 110 bi-gram phrases per individual. The average syllable repertoire (z) represented by our sampling is 70. I developed the null probability of bi-gram phrase sharing between two individuals as follows: (1) Shared bi-gram phrases must be comprised of shared syllables. For a given individual, the chance of having bi-grams comprised of shared syllables is the square of the proportion of shared syllables, or $(S_j(\text{adj}))^2$. The chance for two individuals to share this bi-gram is $(S_j(\text{adj}))^4$. The number of shared syllables in one individual's repertoire is $z * S_j(\text{adj})$. (2) The chance of having the same bi-gram between individuals equals the chance that both picked the same syllables twice in a row from the pool of shared syllables; hence, the probability is $(1/(z * S_j(\text{adj})))^2$. To sum up, the null probability of bi-gram phrase sharing between two individuals is

$$SS_j(\text{adj})_{\text{null}} = (S_j(\text{adj}))^4 * (1/z * S_j(\text{adj}))^2 = (S_j(\text{adj})/z)^2.$$

Thus, the null chance of shared bi-grams between individuals is positively correlated with the degree of syllable sharing and negatively correlated with syllable repertoire.

Discriminant Function Analysis of Vocal Profile

I measured ten features of each song (Table 2.1). The variables “Average frequency” and “End frequency” were highly correlated ($r=0.92$, $p<0.0001$); thus I excluded “End frequency” from analysis. The correlations among all other variables was much lower, ranging from -0.2 to 0.61 (mean \pm SD, 0.34 ± 0.23), which translated to a tolerance greater than 0.1. Tolerance is a measure of multicollinearity given by $1 - R^2$, where R^2 is the coefficient of determination from a regression of one variable on all the other variables. A tolerance of less than 0.1 indicates a multicollinearity problem (O’Brien, 2007).

To test for a geographic signature evident in this measured vocal profile, I applied discriminant function analysis (DFA) (McLachlan, 2004). Vocal variables not normally distributed were normalized using the Box-Cox transformation (Box & Cox, 1964). The *a priori* classification probabilities were set to an equal weight for each individual. DFA also provided the posterior classification probabilities for assignment of each song to any individual. Based on these, I also performed the leave-one-out cross-validation (jack-knife validation) procedure (Manly, 2004). This procedure involved designating a single song (out of 981 total) as the validation data, and all remaining songs as the training data to evaluate the performance of the discriminant function. This process is repeated such that each song is used once as the validation data. Computational implementation was performed in the R environment using “lda” function from the library “MASS”.

Table 2.1. Measured features of the house finch vocal profile

Variables	Description
Average frequency (Hz)-AF	The average frequency of the entire song
Start frequency (Hz)-SF	The frequency at the beginning of the song
End frequency (Hz)-EF	The frequency at the end of the song
Bandwidth (Hz)-BW	The average difference between highest and lowest frequencies of each syllable in the song
Frequency-time excursion length-FL	Arc length of a song treated as continuous by connecting amplitude peaks in the spectrogram, using line segments to create a polygonal path
Duration (ms)-DU	Time duration of the song
Highest frequency (Hz)-HF	The average highest frequency of each syllable of the song
Lowest frequency (Hz)-LF	The average lowest frequency of each syllable of the song
Concavity-CO	The average number of changes in sign of slope of frequency change per syllable throughout the song.
Intersyllable-IS	The average duration of silence between consecutive syllables in the song

Results:

Syllable Classification

A syllable catalog of 275 different syllable types was compiled out of 11968 syllables (Figure 2.2), excluding 175 syllables which were identified as outliers or could not be classified because of poor quality in recordings. The most common type was present in the syllable repertoires of 73 individuals (76% of all individuals).

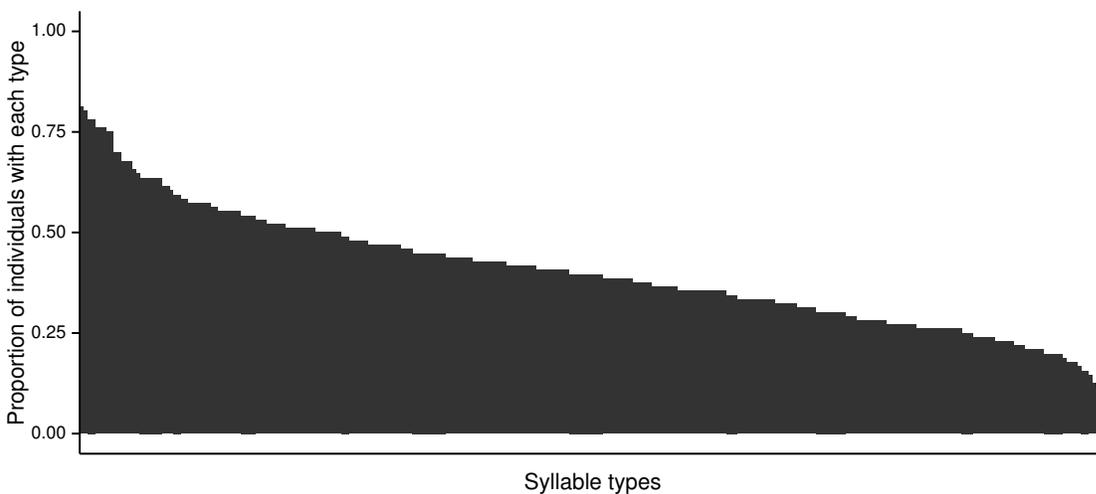


Figure 2.2. Bar plot showing frequencies of the 275 identified syllable types in the population

Types are sorted in a decreasing manner according to their frequencies.

Song Sharing Between Individuals

Songs typically ranged from 1 to 4 s (mean \pm SD of 2.47 \pm 1.1s) in length and were comprised of 6 to 35 syllables (mean \pm SD of 12.32 \pm 5.6). Each individual had 1 to 7 different song types (mean \pm SD of 3.1 \pm 1.3), excluding song types that occurred only once. There was variation within the same song type of each individual. Insertion or deletion of one to three syllables were

common within a song, and in many cases songs were not sung to completion, as suggested previously (Bitterbaum & Baptista, 1979; Mundinger, 1975; Tracy & Baker, 1999)

No pair of individuals shared a song type (Figure 2.3), even between individuals geographically close to each other. For instance, Figure 2.4 illustrates the most common song types from two pairs of nearby individuals (0.25km and 0.12km respectively). Similarity of songs between even these nearby birds was low.

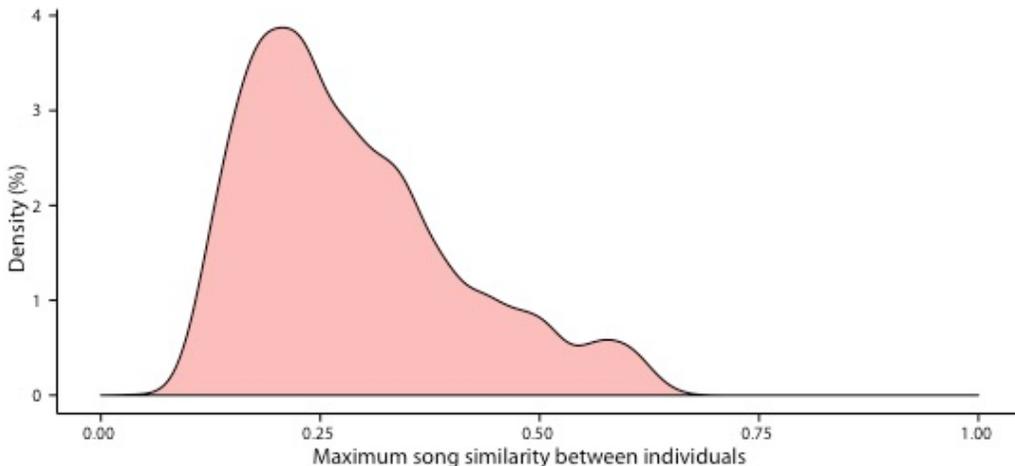


Figure 2.3. Density plot exhibiting the distribution of maximum song similarity between individuals

Similarity ranges from 0 (no sharing of any one syllable) to 1 (the exact same song type, in terms of shared syllables). Higher similarity scores between individuals would have indicated sharing of song types between individuals. According to 75% similarity criterion, no song sharing were found between individuals

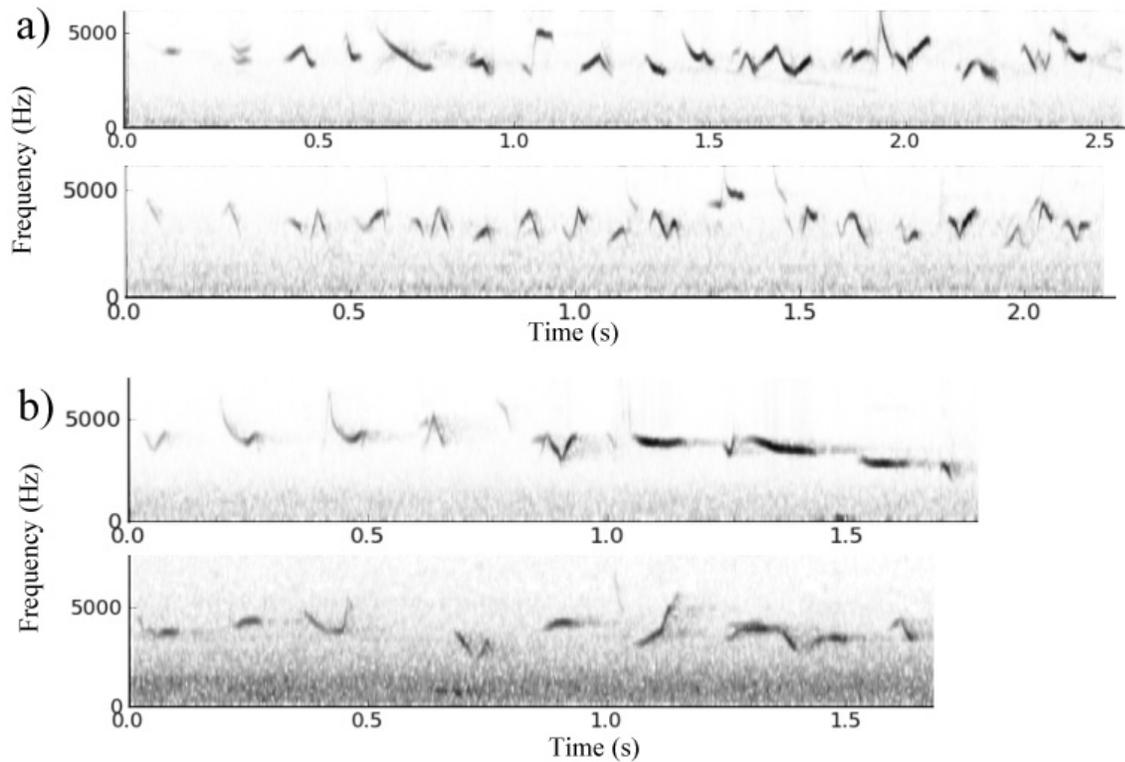


Figure 2.4. Examples of limited sharing of songs between nearby individuals

The distance between the two individuals in Panel **a** is 0.25 km, and in Panel **b** is 0.12 km.

Analysis of Syllable Repertoires

I plotted the cumulative number of syllable types detected in an individual against the n^{th} song sampled for this individual, for ten individuals from which at least 20 high-quality recordings were made (Figure 2.5). After approximately the tenth song sampled, few syllable types were detected. However, individuals varied in how quickly the cumulative curves leveled off before the tenth song. Some birds approached quickly to the maximum of their repertoire sizes in less than five songs, while some leveled off slowly. For these ten sampled birds, on average 76% (61 - 90%) of the total detected types were retrieved by the fifth song. In addition, there was no significant difference between $S_j(\text{adj})$ values between these ten individuals calculated using all

syllables, as compared to syllables from the first five songs (Wilcoxon signed-rank test, $S = 123.5$, $P=0.187$). Therefore, I decided to use five songs as the threshold to retain birds for subsequent analysis.

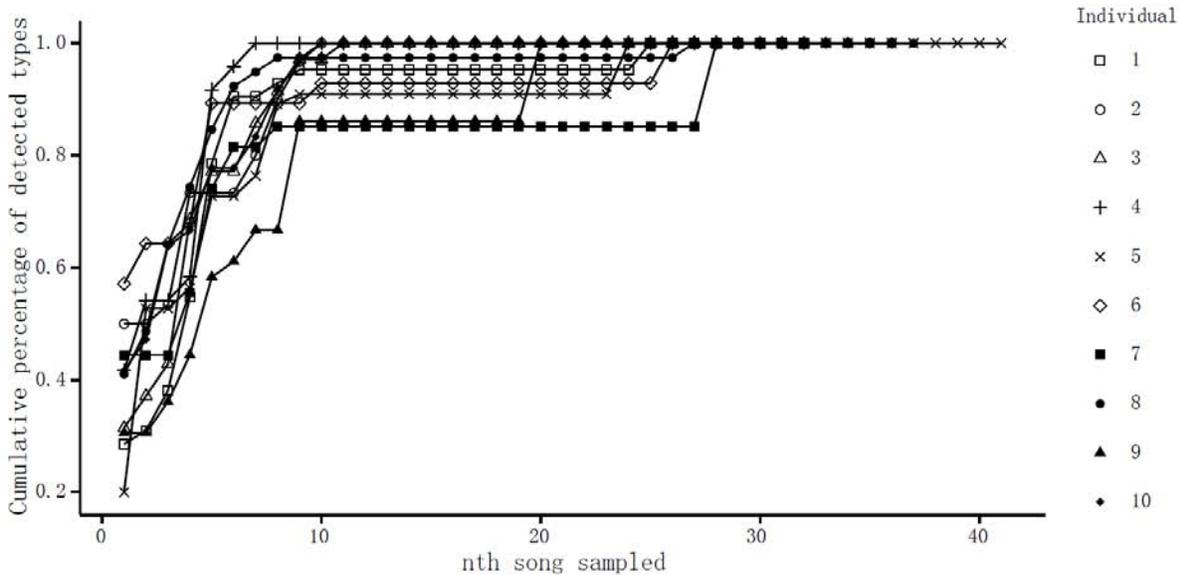


Figure 2.5. Cumulative percentage of syllable types detected by nth song

Cumulative percentage of syllable types plotted against the n^{th} song sampled for ten different birds, which were thoroughly sampled. For most birds in my sample, few new types were detected after tenth song sampled, and at least 61% of the total syllable repertoire was detected within the first five songs.

Syllable Sharing between Individuals versus Distance

$S_j(\text{adj})$ values ranged from 0.1 to 0.85. Generally, $S_j(\text{adj})$ values were significantly negatively associated with distance between individuals (Mantel test, $r=-0.6063$, $P=0.001$, number of permutations=999) (Figure 2.6a). A rapid decline of $S_j(\text{adj})$ values was observed at a short distance, followed by a leveling off at a greater distance (Figure 2.6b). $S_j(\text{adj})$ values within

5km and between 5 to 10 km were significantly different from those of the rest classes of distance (adjusted P by bonferroni < 0.001), while no significant difference in $s_j(\text{adj})$ values were found among other classes of distance.

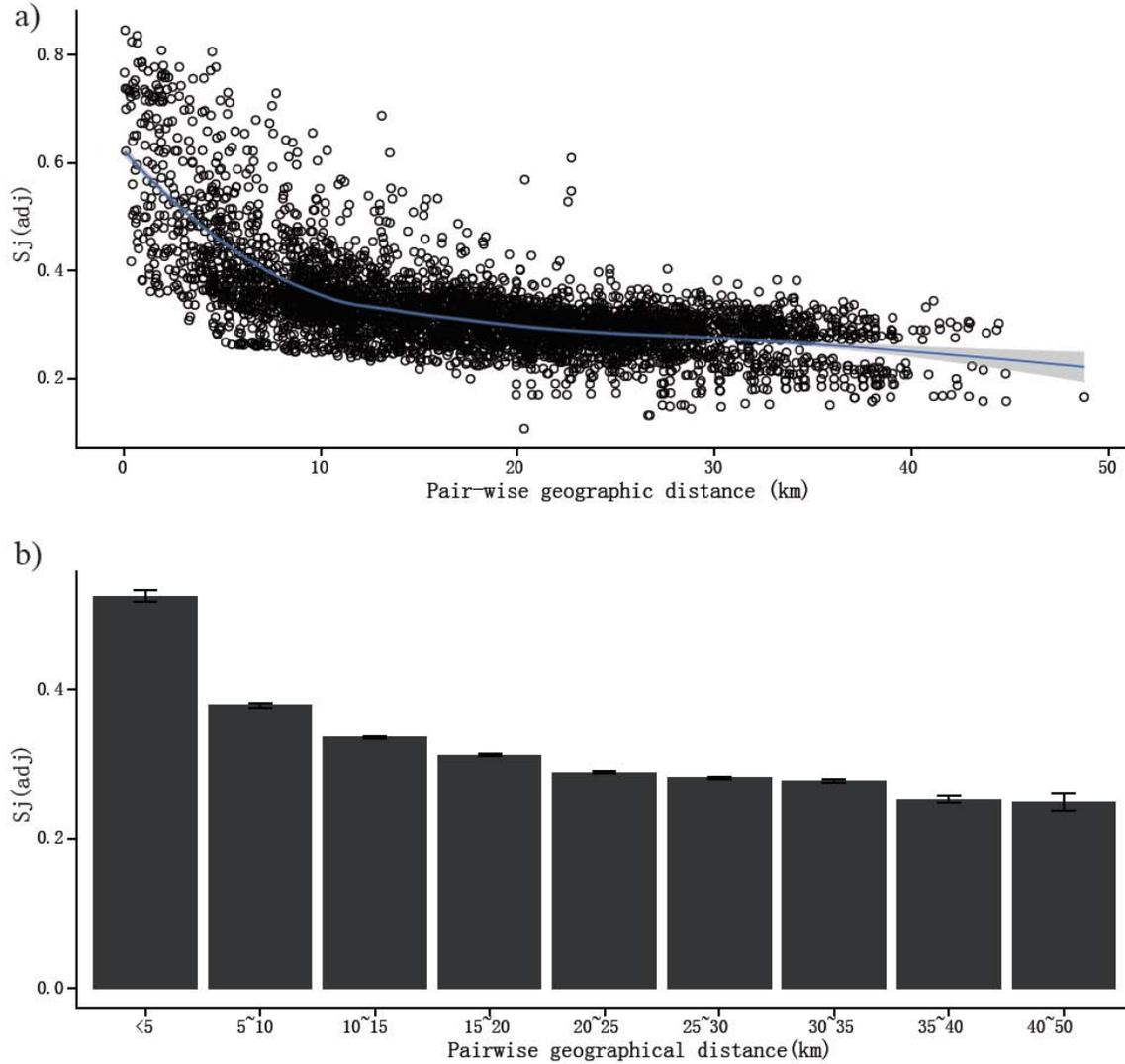


Figure 2.6. Syllable sharing between individuals versus distance

a) Scatterplot showing relationship between $S_j(\text{adj})$ values for syllable sharing and geographic distance. The blue line was plotted based on loess regression. Grey areas indicate the 95% confidence interval. Similarity of syllable repertoires decreased exponentially with distance between individuals. **b)** Barplot showing means and standard deviations of $s_j(\text{adj})$ at different classes of distance. A rapid decline of $S_j(\text{adj})$ values was observed at a short distance. $S_j(\text{adj})$ values within 5km and between 5 to 10 km were significantly different from those of the rest classes of distance (adjusted P by bonferroni < 0.001).

To illustrate spatial differentiation in degree of syllable sharing, I connected each pair of individuals in a map by a line, the color of which was determined by a ranking of S_j values (Figure 2.7). Specifically, lines connecting pairs of individuals with S_j values in the 75th percentile or higher were colored red, between the 50th - 75th percentile were yellow, between 25th - 50th percentile were green, and the rest were blue. Pairs of individuals within 10 km of each other generally had a higher degree of syllable sharing in the west of the sampling area (Figure 2.7a,c). This pattern held even after $S_j(\text{adj})$ values were corrected by distance (residuals) (Figure 2.7b,d).

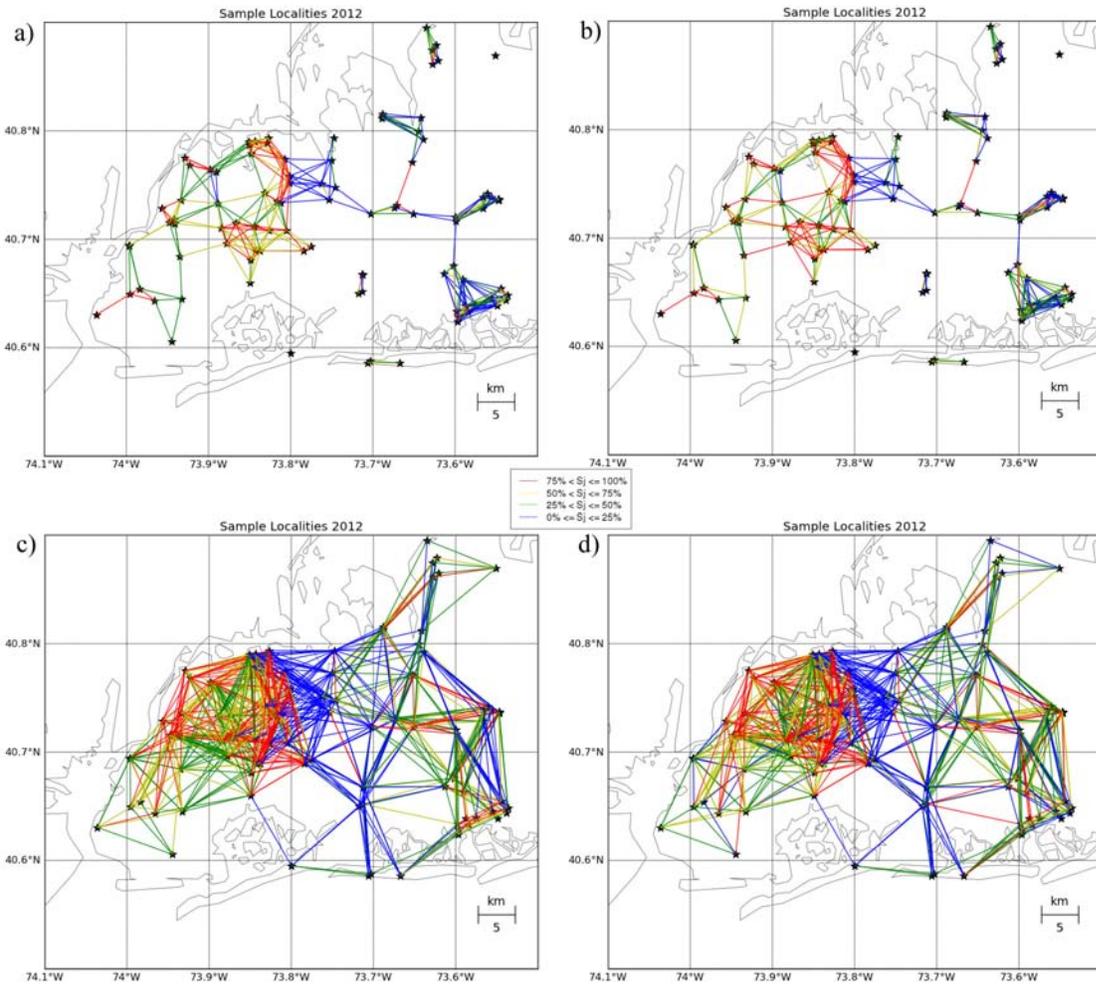


Figure 2.7. Extent of syllable sharing between pairs of individual house finches in the map

Extent of syllable sharing between pairs of individual house finches, according to an adjusted Jaccard's coefficient ($S_j(\text{adj})$). Each pair is connected by a line, the color of which was determined by a ranking of the corresponding $S_j(\text{adj})$ value, as follows. Pairs of individuals with an $S_j(\text{adj})$ larger than the 75th percentile are connected by red lines, those between the 50th and 75th percentile by yellow lines, those from 25th – 50th percentile with green lines, and the rest with blue lines. **a)** Locations of 331 pairs whose distance is less than 5km; colors were determined by ranking of original $S_j(\text{adj})$ values. **b)** Locations of 331 pairs whose distance is less than 5km; colors were determined by ranking of corrected $S_j(\text{adj})$ values for distance; **c)** Locations of 708 pairs whose distance ranged from 5km to 10km; colors were determined by ranking of original $S_j(\text{adj})$ values. **d)** Locations of 708 pairs whose distance ranged from 5km to 10km; colors were determined by ranking of corrected $S_j(\text{adj})$ values for distance. In general, individual pairs from the western area appeared to have a higher degree of syllable sharing.

UPGMA clustering of individuals based on their syllable repertoires separated them into three groups (Figure 2.8a). Interestingly, these three groups were also separated from each other geographically (Figure 2.8b), indicating localized syllable compositions. At the syllable level, if there were dialects I would expect to see syllable types or phrases shared widely among individuals within a region but not between regions. However, there was no such evidence for discrete syllable dialects. All 275 types could be found in region 2 and 3 (Region 1 was not considered since the number of individuals was too small to reveal the whole lexicon in this region).

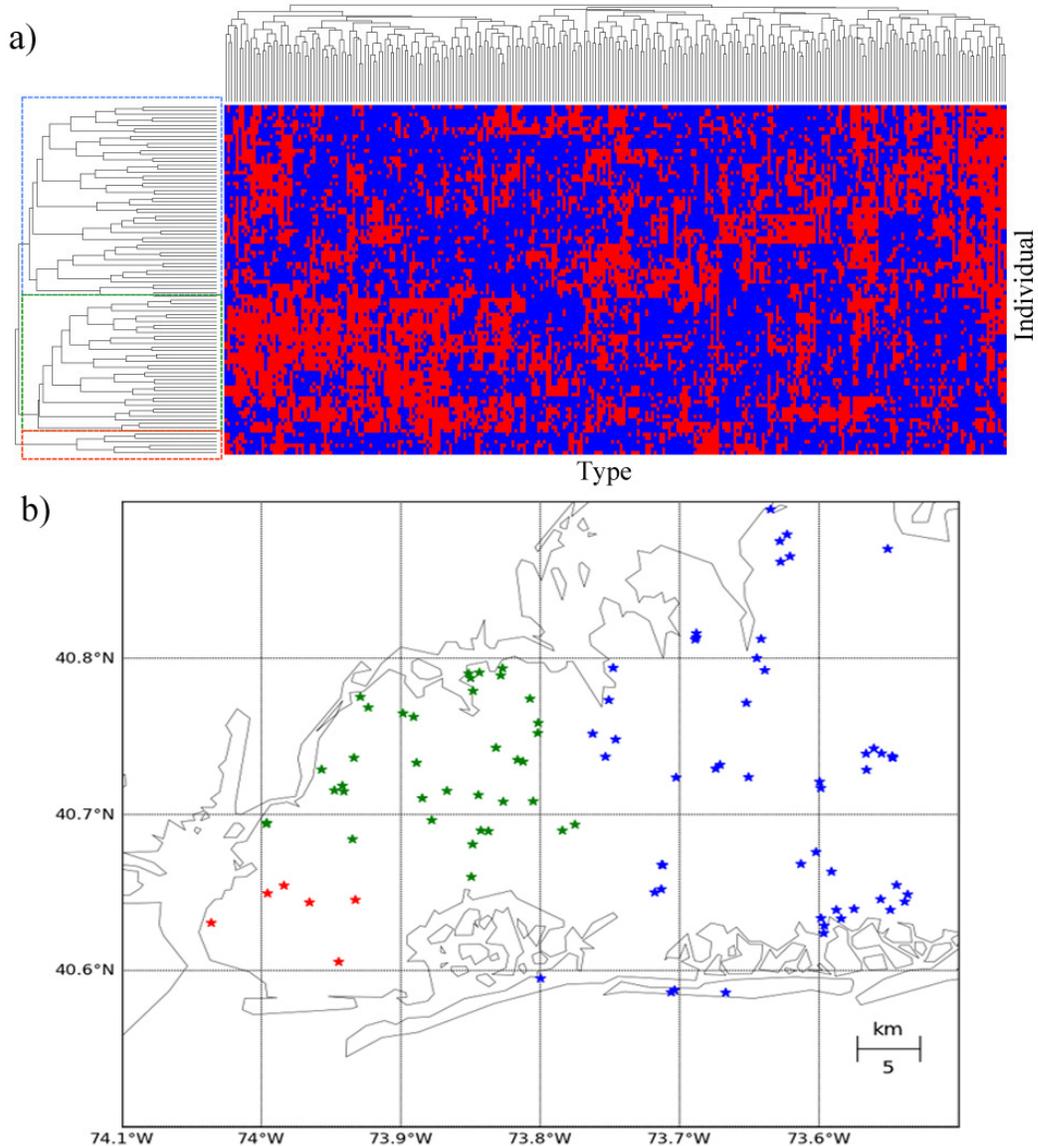


Figure 2.8. Localized syllable composition based on UPGMA clustering

a) UPGMA clustering based on similarities of syllable repertoires. Similarities between individuals were estimated by Pearson correlation on presence/absence of all syllable types. A heat map was generated to illustrate the clustering result, in which each row represents an individual bird, and each column a syllable type. Cells in red indicate presence of the syllable types in the corresponding individual's repertoire, and blue cells indicate absence. Individuals sharing most types were grouped together. Three major groups were formed for individuals on the basis of syllable sharing, as indicated in the hierarchical tree to the left. Syllable types present in a similar set of individuals were likewise grouped together. **b)** Location of individuals clustered into three groups. Three groups of individuals were colored as red (region 1), green (region 2) and blue (region 3), corresponding to results of **a**.

Phrase Sharing Between Individuals versus Distance

Of 4,560 comparisons, 119 pairs shared at least one bigram. Phrase sharing was only found between individuals within 20km, with one exception: one shared phrase was found between two individuals 26km apart (Figure 2.9). Limited sharing of phrases (bigrams) was observed between the three regions of distinct syllable composition identified above (Figure 2.10). Only one phrase was shared between region 2 and 3. Any observed sharing of bigrams between individuals is unlikely by chance (for these pairs the values of $SSj(\text{adj})$ were non-zero), as the maximum of the null probability of shared bi-gram between two individuals is $(0.83/70)^2 \approx 0.0001$ (0.83 is the maximum value of $Sj(\text{adj})$ and 70 is the average syllable repertoire size), which is one order of magnitude lower than the minimum value of $SSj(\text{adj})$ (0.0012) between individuals sharing at least one bi-gram.

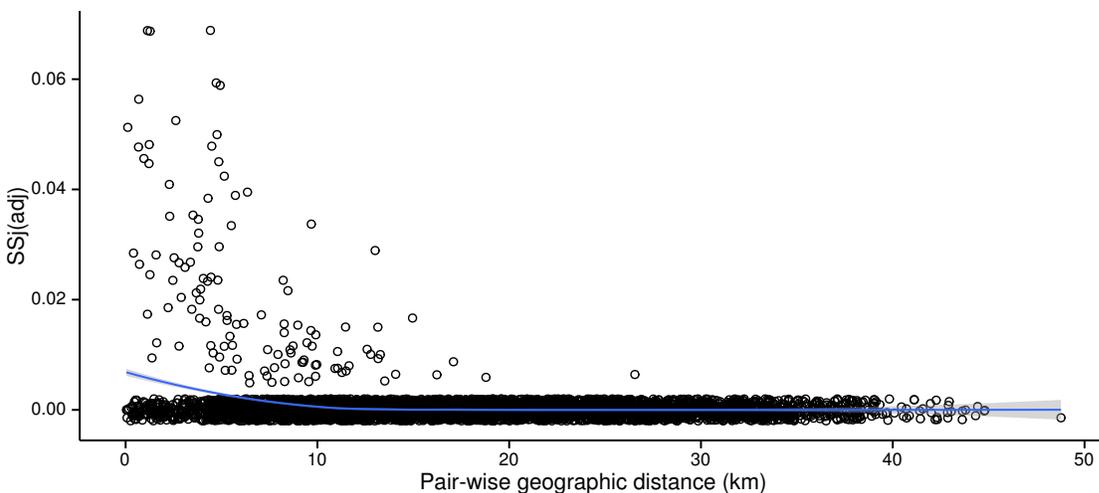


Figure 2.9. Scatterplot showing the relationship between $SSj(\text{adj})$ values for phrase sharing and geographic distance

The blue line was plotted based on loess regression. Grey areas indicate 95% confidence intervals. Only 119 pairs of individuals shared one or more bigrams. Nearly all phrase sharing occurred between individuals within 20 km; in one instance, individuals more than 20 km apart shared a phrase. I jittered points at $y=0$ to unhide them.

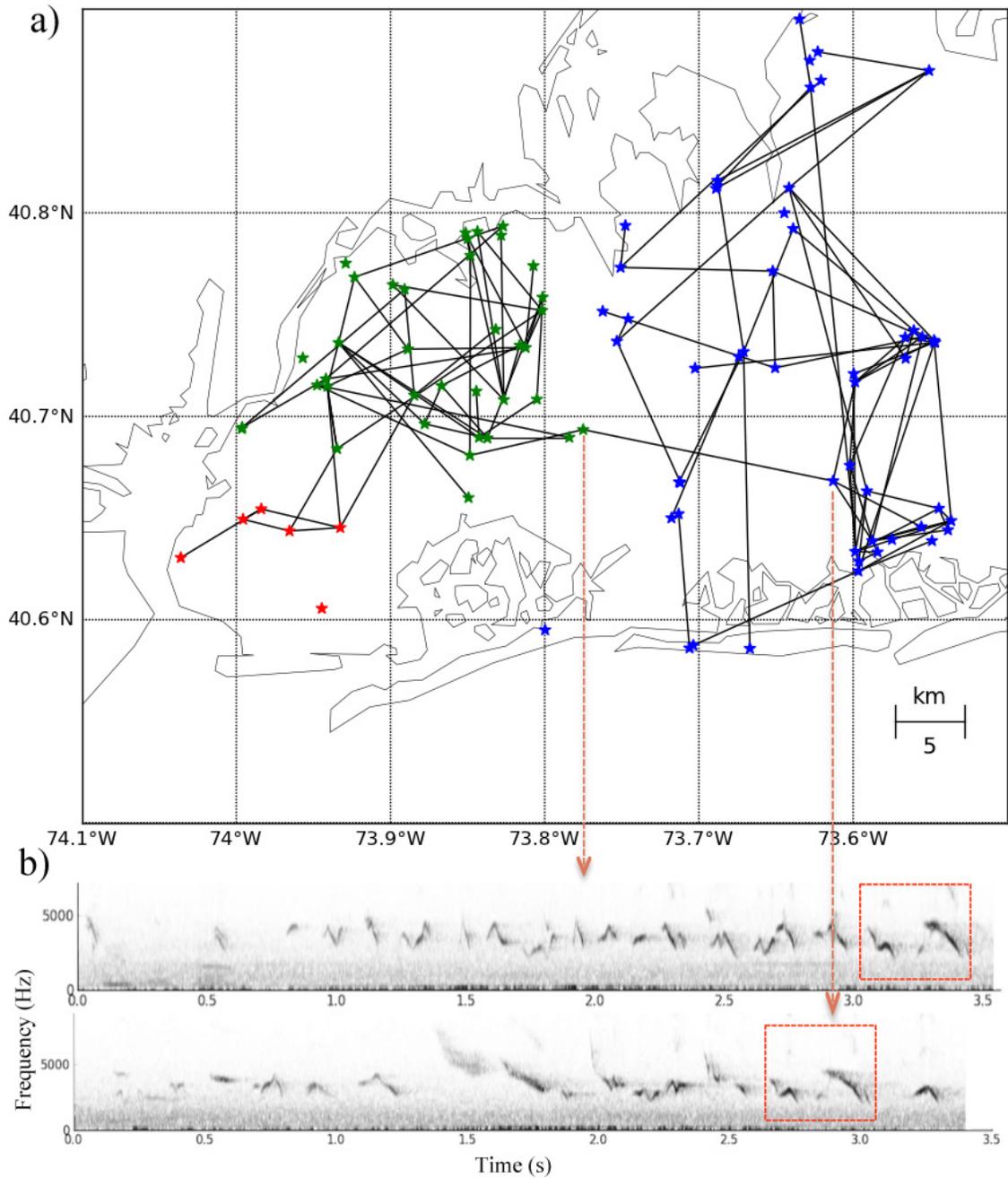


Figure 2.10. Sharing of phrases in the map

a) Locations of pairs of individuals ($N=119$) that shared at least one bigram. Individuals are colored according to the results of clustering shown in Figure 2.8. Most phrase sharing occurred within identified regions. Only one phrase was shared between region 2 and 3. **b)** Two spectrograms show the shared bigram (highlighted by red boxes) between the pair of individuals between region 2 and 3.

Discriminant function analysis of vocal profile

The DFA indicated that the individuals differed significantly in their vocal profile (MANOVA Wilks' Lambda $F=7.6312$, $P < 2.2e-16$). The concavity value, the intersyllable distance, and the lowest frequency of the song were the most significant variables for discriminating between individuals; they comprised the first root that accounted for more than 50% of the variance (Table 2.2). These three variables also exhibited geographical pattern. For example, songs of individuals in region 2 generally had higher concavity values than those from the other 2 regions (Figure 2.11). After assigning equal prior probability for individuals, 74.7% of cross-validated songs were correctly classified into their original individuals, suggesting that vocal profile can be used to distinguish songs of individual house finches in this study area.

Table 2.2. The linear discriminant function coefficient based on vocal profile

Variables	Discriminant roots							
	LD1	LD2	LD3	LD4	LD5	LD6	LD7	LD8
CO	0.655	0.290	0.173	-0.194	0.147	-0.588	-0.148	-0.148
SF	-0.004	-0.659	0.813	-0.066	-0.191	-0.468	-0.559	-0.016
AF	-0.026	-0.224	-0.911	0.585	0.002	-0.179	0.333	-0.086
HF	-0.002	0.130	-0.929	-0.003	-0.189	0.490	0.210	-0.022
FL	0.035	-0.785	-0.019	-0.011	0.718	0.484	-0.105	0.205
LF	0.122	0.230	0.079	0.064	-0.155	0.904	0.036	-0.022
IS	-0.223	-0.738	0.326	0.003	-0.871	-0.201	-0.034	0.371
DU	0.013	-0.695	-0.160	0.000	0.125	-0.185	0.069	-0.047
BW	0.016	0.172	-0.037	-0.002	0.491	0.213	0.001	-0.019
% of variance	56.86%	18.46%	8.50%	5.01%	4.27%	2.92%	2.50%	2.20%

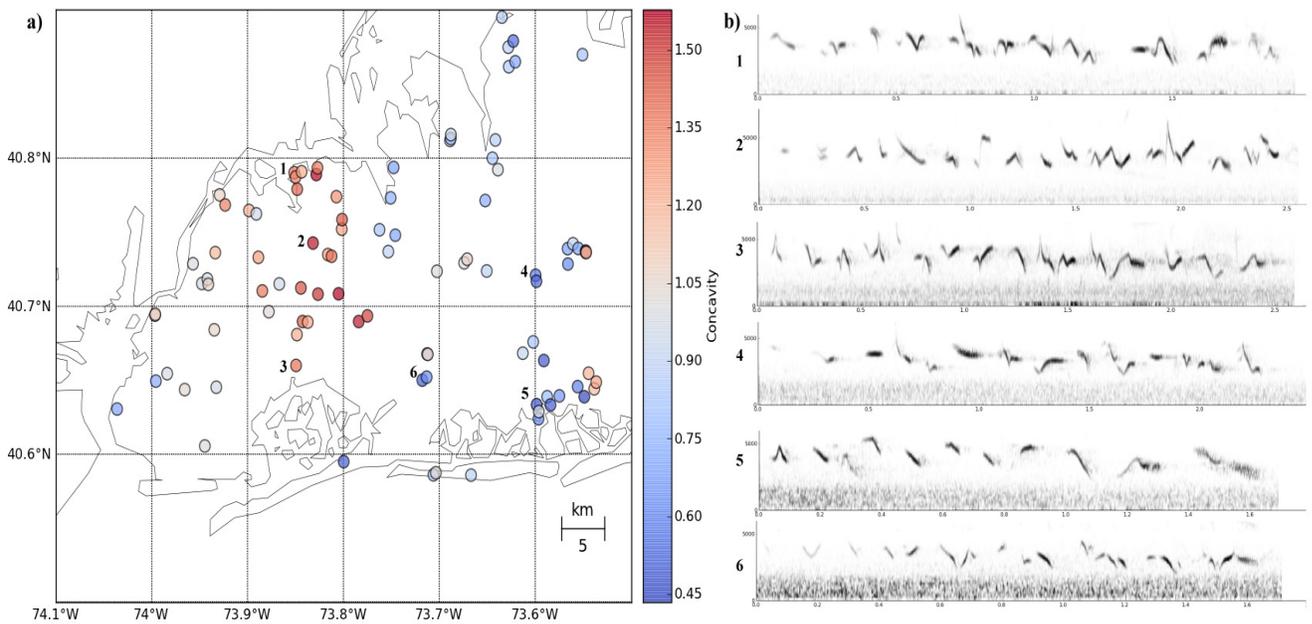


Figure 2.11. Geographic pattern of concavity

a) Average values of concavity for house finch individuals. Red represents a higher average concavity value, while blue represents a lower concavity value. This spatial pattern corresponds well to observed patterns based on similarities in syllable repertoire. **b)** Spectrograms of common song types from six individuals marked with numbers in panel **a**.

Discussion:

I examined the geographic variation of house finch songs within a relatively small region (34 km by 40 km). This region overlaps with the putative introduction site of the eastern population in Long Island (Elliott & Arbib, 1953), and is close to the study area of Mundinger (1975, 1982) in Westchester County (~25 km away) four decades earlier. In a previous analysis, Mundinger (1975) suggested high stereotypy and low syntactic variation in songs of house finches in the eastern population. Surprisingly, I observed no shared song types between individuals in this study, even among nearby birds (distance < 0.5 km). If the absence of song type sharing did not date back to the time when the founders of this region settled, this substantial syntactical change at the song level suggests that at least part of the eastern population is undergoing rapid cultural

evolution. Distinguishing vocal mechanisms underlying cultural evolution requires temporal data spanning a sufficiently long period, which will be addressed in the next chapter.

Geographic acoustic variation in birds typically has been evaluated at the level of whole songs; however, the same syllables may be found between songs of different individuals that do not share the same song types. Thus, a population might exhibit great variation at the song level, yet exhibit a high degree of sharing of syllables. Some studies have indicated that the basic unit of song learning could be the syllable or phrase, instead of a complete song (Lemon, 1975; D. A. Nelson, Marler, & Palleroni, 1995). In addition, lower-order units (e.g. syllables) are commonly more stable than higher-order units (songs and syllable sequences) over a long time period (Ellers & Slabbekoorn, 2003; D. A. Nelson, Hallberg, & Soha, 2004). These results highlight the importance of examining geographic variation at the syllable and phrase levels. In the present study, the similarity of syllable repertoires between individuals was negatively associated with distance. Some pairs of nearby birds, without common song types, shared up to 60% of their syllable type repertoires. No syllable type was found among all individuals, but relatively distant birds in my study area had a small number of syllable types in common. My findings were similar to the sharing pattern of house finch songs reported in California (Bitterbaum & Baptista, 1979) and Colorado (Tracy & Baker, 1999), in which no distinct dialect boundaries occurred and syllable repertoires became gradually less similar with distance. Moreover, this finding leads to an inference that syllables or phrases might be learned, instead of whole songs.

Nevertheless, my results were neither in accordance with a clinal pattern in which gradual and shallow change is observed across the entire range of the studied area, nor consistent with the classic definition of dialects (Marler & Tamura, 1962; Mundinger, 1982), on which acoustic variation is characterized by sharp transitions between localities and high consistency within localities. Despite the fact that the degree of syllable and phrase sharing decreased with distance, I detected three regions, within each of which individuals' syllable repertoires matched more closely to each other than to individuals from the other two regions, and phrase sharing was more likely to be found. These results indicated that less syllable and phrase variation existed within a local region than between distant areas. However, the common syllable types within a region were not exclusively absent in other regions (Figure 2.8b), and some within-region pairs of individuals exhibited lower sharing than some between-region pairs.

As complex as the geographic pattern is, at least three hypotheses might explain the mixture of divergent and convergent findings in this study. One hypothesis is a multiple founder effect resulting from a small number of founders that varied among each other in syllable composition. In previous studies, house finches in Westchester County exhibited distinct song dialects over just a few kilometers on a side, with clearly established local boundaries separating dialects (Mundinger, 1975). However, the degree of syllable sharing within and between separated dialect areas was not well described. If distinct syllable composition was also observed between dialect areas, dialect boundaries could be explained at the level of syllable differences as well. Whether dialects were distinguishable on the basis of song type or syllable repertoire, the multiple founder effect hypothesis could be tested by performing genetic analyses to examine whether individuals from different areas were genetically distinct.

The second hypothesis is differential dispersal or migratory patterns, which could be a result of change in migratory behavior in at least part of the eastern population of the house finch. The house finch used to be considered sedentary based on studies of its ancestral population in the Southwest. However, partial migration was observed in the eastern population, where an increasing proportion of house finches have been migrating (Able & Belthoff, 1998). Song learning in the house finch is not restricted to the period prior to natal dispersal (Mundinger, 1982), and imitative learning is not restricted to parents. Hence, limited or biased movement is critical to song dialect evolution, where song similarity arises from a bias towards learning in the natal area (Marler & Tamura, 1962). Migration, by which individuals will likely encounter individuals from other neighborhoods even if they do settle in exactly the same place as they had left as a juvenile, would lead to mixing of syllable and song types, resulting in an obliteration of dialects. In addition, migration can result in increased dispersal distance because males going away for the winter are probably less likely to settle in their natal area than males that spent their first winter there (Paradis et al., 1998). Such an increase in dispersal distance would further increased mixing of syllables and song types, as males might continue learning songs after landing in places where the predominant songs are different from those in their own natal neighborhood. A comparative study of the process of song development between migratory and sedentary subspecies of white-crowned sparrow showed that migratory subspecies tended to overproduce song and phrase types during song development, most of which were likely to be lost after settlement (D. Nelson et al., 1996). In this study, it is likely that individuals from some regions (e.g., region 3) are undergoing increased exchange of songs and song elements with conspecifics outside of their natal area due to migration or increased dispersal distance. As a consequence, migratory individuals might display a lower degree of sharing with adjacent birds.

To test this hypothesis, birds could be banded to track migration of individuals from different regions.

The last hypothesis could be based on the “developmental stress hypothesis” (Nowicki, Peters, & Podos, 1998), which proposes that poor nutrition or other early developmental privation would compromise birds’ ability in song learning. In turn, song could serve as an indicator of the early developmental history (Brumm, Zollinger, & Slater, 2009; MacDougall-Shackleton & Spencer, 2012). In this study, individuals from region 2 generally had higher number of concavity on average, a potential indicator of higher degree of song elaboration. The region is also more densely populated than the other two, implying relatively higher food resources. Moreover, the elaboration of songs is suggested to serve as an indicator of male quality and thus might be under the influence of sexual selection. If this is the case, individuals with a better developmental history will likely be favored by females, which have already been shown to prefer a high song rate (Mennill et al., 2006; Nolan & Hill, 2004). Testing the developmental stress / sexual selection hypothesis will be challenging, not only because it represents a range of more specific possibilities, but also because multiple causal factors can interact with each other. An effective test will first distinguish developmental effect on song elaboration from the effect of female preference and proposed regional differences in resource abundance. Developmental stress can be tested with laboratory studies involving imposed food deprivation in early development, or else observational correlations in the field between nestling or fledgling health and subsequent song elaboration. Regional resource abundance can be inferred from these same natural observations, or could be assessed more directly by measuring food availability. Lastly, the effect of local resources on song features could be tested by supplementing the food of wild breeding

house finches in a restricted area. Such field experiments would face the significant challenge of locating birds in their first breeding year after banding them as nestlings or fledglings.

Hypotheses two and three above thus provide potential explanations for the observed higher sharing and elaboration respectively in the songs of the western region of the study area compared to the eastern region. These observations also render unlikely an alternative, “byproduct” hypothesis, whereby elaborate songs exhibit more variation in acoustic features, and this increased variation led to less sharing because of the greater variety of possibilities open to learning birds. The results of the present study show the opposite—the region with higher sharing (region 2) has the more elaborate syllables.

To further explore the mechanisms underlying cultural evolution in this region, the next chapter provides a temporal comparative analysis of house finch songs spanning nearly four decades. This study will illuminate the precise changes that have occurred in the song of the house finch and perhaps indicate the factors that underlie the geographic differentiation in vocal profiles observed in this study.

References:

- Able, K. P., & Belthoff, J. R. (1998). Rapid 'evolution' of migratory behaviour in the introduced house finch of eastern North America. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1410), 2063-2071.
- Aldrich, J. W., & Weske, J. S. (1978). Origin and evolution of the eastern House Finch population. *The Auk*, 95, 528-536.
- Anderson, D. (2008). *Model Based Inference in the Life Sciences: A Primer on Evidence*: Springer.
- Andrew, R. J. (1962). Evolution of Intelligence and Vocal Mimicking Studies of large-brained mammals promise to elucidate some problems of human evolution. *Science*, 137(3530), 585-589.
- Badyaev, A. V., & Hill, G. E. (2000). Evolution of sexual dichromatism: contribution of carotenoid-versus melanin-based coloration. *Biological Journal of the Linnean Society*, 69(2), 153-172.
- Baker, M. C., & Cunningham, M. A. (1985). The biology of bird-song dialects. *Behavioral and Brain Sciences*, 8(01), 85-100.
- Bitterbaum, E., & Baptista, L. F. (1979). Geographical variation in songs of California House Finches (*Carpodacus mexicanus*). *The Auk*, 96, 462-474.
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26, 211-252.
- Brumm, H., Zollinger, S. A., & Slater, P. J. (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. *Behavioral Ecology and Sociobiology*, 63(9), 1387-1395.
- Campbell, P., Pasch, B., Pino, J. L., Crino, O. L., Phillips, M., & Phelps, S. M. (2010). Geographic variation in the songs of neotropical singing mice: testing the relative importance of drift and local adaptation. *Evolution*, 64(7), 1955-1972.
- Catchpole, C. K., & Slater, P. (2008). *Bird Song: Biological Themes and Variations*: Cambridge University Press.
- Cunningham, M. A., Baker, M. C., & Boardman, T. J. (1987). Microgeographic song variation in the Nuttall's White-crowned Sparrow. *Condor*, 261-275.
- Ellers, J., & Slabbekoorn, H. (2003). Song divergence and male dispersal among bird populations: a spatially explicit model testing the role of vocal learning. *Animal Behaviour*, 65(4), 671-681.
- Elliott, J. J., & Arbib Jr, R. S. (1953). Origin and status of the house finch in the eastern United States. *The Auk*, 70, 31-37.
- Hafner, D. J., & Petersen, K. E. (1985). Song dialects and gene flow in the white-crowned sparrow, *Zonotrichia leucophrys nuttalli*. *Evolution*, 39, 687-694.

- Hill, G. E. (1993). Male mate choice and the evolution of female plumage coloration in the house finch. *Evolution*, *47*, 1515-1525.
- Krebs, C. (1989). *Ecological Methodology*. Harper Collins Publishers.
- Kroodsma, D. E. (1996). Ecology of passerine song development. *Ecology and Evolution of Acoustic Communication In Birds*, 3-19.
- Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*, *24*(5), 719-720.
- Leader, N., Wright, J., & Yom-Tov, Y. (2000). Microgeographic song dialects in the orange-tufted sunbird (*Nectarinia osea*). *Behaviour*, *137*(12), 1613-1628.
- Legendre, P., & Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, *80*(2), 107-138.
- Lemon, R. E. (1975). How birds develop song dialects. *Condor*, *77*, 385-406.
- MacDougall-Shackleton, S. A., & Spencer, K. A. (2012). Developmental stress and birdsong: current evidence and future directions. *Journal of Ornithology*, *153*(1), 105-117.
- Manly, B. F. (2004). *Multivariate Statistical Methods: A Primer*: CRC Press.
- Marler, P., & Tamura, M. (1962). Song "dialects" in three populations of White-crowned Sparrows. *Condor*, *64*, 368-377.
- Marler, P., & Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science*, *146*(3650), 1483-1486.
- McLachlan, G. (2004). *Discriminant Analysis and Statistical Pattern Recognition* (Vol. 544): John Wiley & Sons.
- Mennill, D. J., Badyaev, A. V., Jonart, L. M., & Hill, G. E. (2006). Male house finches with elaborate songs have higher reproductive performance. *Ethology*, *112*(2), 174-180.
- Miller, L. (1921). The Biography of Nip and Tuck: A Study of Instincts in Birds. *Condor*, 41-47.
- Mundinger, P. (1975). Song dialects and colonization in the house finch, *Carpodacus mexicanus*, on the east coast. *Condor*, *23*, 407-422.
- Mundinger, P. (1982). Microgeographic and macrogeographic variation in the acquired vocalizations of birds. *Acoustic Communication in Birds*, *2*, 147-208.
- Nelson, D., Marler, P., & Morton, M. (1996). Overproduction in song development: an evolutionary correlate with migration. *Animal Behaviour*, *51*(5), 1127-1140.
- Nelson, D. A., Hallberg, K. I., & Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology*, *110*(11), 879-908.
- Nelson, D. A., Marler, P., & Palleroni, A. (1995). A comparative approach to vocal learning: intraspecific variation in the learning process. *Animal Behaviour*, *50*(1), 83-97.

- Nelson, D. A., & Soha, J. A. (2004). Perception of geographical variation in song by male Puget Sound white-crowned sparrows, (*Zonotrichia leucophrys pugetensis*). *Animal behaviour*, 68(2), 395-405.
- Nolan, P. M., & Hill, G. E. (2004). Female choice for song characteristics in the house finch. *Animal Behaviour*, 67(3), 403-410.
- Nowicki, S., Peters, S., & Podos, J. (1998). Song learning, early nutrition and sexual selection in songbirds. *American Zoologist*, 38(1), 179-190.
- Nowicki, S., Westneat, M., & Hoese, W. (1992). Birdsong: motor function and the evolution of communication. *Seminars in Neuroscience*, 4(6), 385-390.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690.
- Oden, N. L., & Sokal, R. R. (1986). Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Biology*, 35(4), 608-617.
- Panhuis, T. M., Butlin, R., Zuk, M., & Tregenza, T. (2001). Sexual selection and speciation. *Trends in Ecology & Evolution*, 16(7), 364-371.
- Paradis, E., Baillie, S. R., Sutherland, W. J., & Gregory, R. D. (1998). Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, 67(4), 518-536.
- Payne, R. B. (1981). Song learning and social interaction in indigo buntings. *Animal Behaviour*, 29(3), 688-697.
- Petrinovich, L., & Patterson, T. L. (1981). The responses of white-crowned sparrows to songs of different dialects and subspecies. *Zeitschrift für Tierpsychologie*, 57(1), 1-14.
- Podos, J. (2007). Discrimination of geographical song variants by Darwin's finches. *Animal Behaviour*, 73(5), 833-844.
- Podos, J., Huber, S. K., & Taft, B. (2004). Bird song: the interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics*, 35, 55-87.
- Podos, J., Peters, S., Rudnicki, T., Marler, P., & Nowicki, S. (1992). The organization of song repertoires in song sparrows: themes and variations. *Ethology*, 90(2), 89-106.
- Podos, J., & Warren, P. S. (2007). The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, 37, 403-458.
- Pytte, C. L. (1997). Song organization of house finches at the edge of an expanding range. *Condor*, 99, 942-954.
- Ristad, E. S., & Yianilos, P. N. (1998). Learning string-edit distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(5), 522-532.
- Sauer, J. R., Hines, J. E., Fallon, J., Pardieck, K., Ziolkowski Jr, D., & Link, W. (2008). *The North American Breeding Bird Survey, Results and Analysis 1966-2007*. Version, 5(15), 2008.

- Slabbekoorn, H., & Smith, T. B. (2002). Habitat-dependent song divergence in the little greenbul: An analysis of environmental selection pressures on acoustic signals. *Evolution*, *56*(9), 1849-1858.
- Slater, P. (1989). Bird song learning: causes and consequences. *Ethology Ecology & Evolution*, *1*(1), 19-46.
- Sokal, R. R. (1958). A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull*, *38*, 1409-1438.
- Sung, H.-C., & Handford, P. (2006). Songs of the Savannah Sparrow: structure and geographic variation. *Canadian Journal of Zoology*, *84*(11), 1637-1646.
- Thompson, W. L. (1960). Agonistic behavior in the House Finch. Part I: Annual cycle and display patterns. *Condor*, *62*, 245-271.
- Tracy, T. T., & Baker, M. C. (1999). Geographic variation in syllables of House Finch songs. *The Auk*, *116*, 666-676.
- Tracy, T. T., Zasadny, H. J., Erickson, J. G., & Siemers, J. L. (2009). Syllable and song sharing in a recently established population of House Finches (*Carpodacus mexicanus*). *The Auk*, *126*(4), 823-830.
- Wang, Z., Baker, A. J., Hill, G. E., & Edwards, S. V. (2003). Reconciling actual and inferred population histories in the house finch (*Carpodacus mexicanus*) by AFLP analysis. *Evolution*, *57*(12), 2852-2864.
- Warren, P. S., & Nelson, D. (2002). Geographic variation and dialects in songs of the bronzed cowbird (*Molothrus aeneus*). *The Auk*, *119*(2), 349-361.
- Wilkins, M. R., Seddon, N., & Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution*, *28*(3), 156-166.
- Williams, H., Levin, I. I., Norris, D. R., Newman, A. E., & Wheelwright, N. T. (2013). Three decades of cultural evolution in Savannah sparrow songs. *Animal Behaviour*, *85*(1), 213-223.
- Wright, T. F., Rodriguez, A. M., & Fleischer, R. C. (2005). Vocal dialects, sex-biased dispersal, and microsatellite population structure in the parrot *Amazona auropalliata*. *Molecular Ecology*, *14*(4), 1197-1205.

CHAPTER 3:

Temporal variability in house finch (*Haemorrhous mexicanus*) song:

Four decades of cultural evolution

Abstract:

Examining temporal stability of songs is critical to distinguishing functional mechanisms underlying culture evolution in oscines. In this study, I conducted a comparative analysis of songs of the house finch during an interval of 37 years (between 1975 and 2012). Recordings from both years were collected in western Long Island, a reported site of one of the founding propagules for the Eastern house finch population. No shared song types were found between the past and recent birds. Individuals from 1975 displayed a higher degree of song and syllable sharing, an effect that remained after controlling for geographic distance. Decreased repertoire size of syllable types per individual and rigidity of song syntax were observed in the contemporary songs. These observations imply a decrease in song performance and complexity. In addition, I observed dramatic changes in syllable composition over time. About half of syllable types defined in the past data were not present in the recent songs, which were nevertheless composed from a larger lexicon at the population level. Sharing of the syllable types that persisted over this time period varied according to localized regions of distinct syllable composition. These results illustrate the potential interplay between multiple drivers of spatial and temporal variation. I conclude by suggesting a complex interaction among partial migration, demographic factors, relaxed selection by females on song quality, and developmental stress underlying song evolution.

Keywords: cultural evolution, birdsong, dispersal, relaxed selection, plumage

Introduction:

The classic definition of culture in the animal kingdom is information that can be acquired or modified through social learning of complex behaviors, such as tool use, predator avoidance and vocal communication (Janik & Slater, 2000; Leadbeater & Chittka, 2007; Rapaport & Brown, 2008; Richerson & Boyd, 2008). Oscine birds acquire information via song learning. Recent studies on bird song focus on the lability of song structure in evolution (Janes & Ryker, 2013; Williams, Levin, Norris, Newman, & Wheelwright, 2013). The evolution of cultural traits in bird songs is often considered to be analogous to genetic evolution (P. Munding, 1982; Podos, Huber, & Taft, 2004). Two major hypotheses have been widely accepted to describe how song structure of birds in a population changes over time. The first one assumes a neutral process including mutation (Bentley, Hahn, & Shennan, 2004), where copy errors or improvisation introduce variation during imitative song learning (P. Munding, 1975; van Schaik, 2010), and cultural drift, where rare song and song elements may be lost by chance, analogous to genetic drift and mutation respectively. Under this hypothesis, song divergence is induced by random events, and accumulated via cultural transmission from generation to generation (Payne, 1996). Such a neutral process has been supported by a number of studies examining cultural change in bird songs, such as the middle segment of savannah sparrow (*Passerculus sandwichensis*) songs (Williams et al., 2013) and the unaccented song tradition in chestnut-sided warblers (*Dendroica pensylvanica*) (Byers, Belinsky, & Bentley, 2010). The second hypothesis emphasizes the selective or adaptive processes underlying cultural change. One of the most well known is sexual selection, as bird songs are often primarily produced by males and are heavily involved in male-male competition and mate attraction (Catchpole & Slater, 2008). For instance, consistency of syllable type delivery between song renditions has been related to mating success (Byers, 2007;

Botero et al., 2009). Also, certain song features such as complexity and rate of syllable repetition might be difficult to produce at a high level (Suthers and Goller, 1997; Podos et al., 2009), in which case a male's song performance can indicate quality and thus be preferred by females (Ballentine et al., 2004; Byers et al., 2010). Natural selection can also change bird songs, for instance indirectly through effects on bill morphology (Badyaev, Young, Oh, & Addison, 2008; Podos, 2001; Potvin, 2013). In many cases, a sort of cultural selection can operate, such that certain song features are more likely to be learned or otherwise achieve higher cultural success. For instance, song types or syllables best adapted to local habitats and environment, resulting in highest efficiency of transmission, can be favored (Cardoso & Atwell, 2011; Derryberry, 2009).

Assessing these hypotheses has proved difficult for several reasons. First, neutral and adaptive processes are not mutually exclusive, and their relative roles in evolution largely depend on historical and demographic factors in the population (e.g., founder effects) and the timescale of evolution (Podos et al., 2004; Wilkins, Seddon, & Safran, 2013). Second, cultural evolution is likely to be linked with biological evolution, in which genetically transmitted learning predispositions can influence the former (Lachlan & Feldman, 2003; Lachlan & Servedio, 2004; Lachlan & Slater, 1999; Lahti, Moseley, & Podos, 2011; P. C. Munding & Lahti, 2014). Third, it is widely appreciated that cultural traits vary geographically (Marler & Tamura, 1964; D. A. Nelson, 1998; Podos, 2007), and change over time (Derryberry, 2009; Podos et al., 2004). However, although song learning and development has been extensively studied (Fee & Scharff, 2010; ILipkind & Tchernichovski, 2011; Tchernichovski, Lints, Deregnaucourt, Cimenser, & Mitra, 2004), empirical evidence demonstrating the interplay among drivers of spatial and temporal variation affecting cultural diversity is relatively limited (for exceptions, D. A. Nelson

and Marler (1994); Rivera-Gutierrez, Matthysen, Adriaensen, and Slabbekoorn (2010); Slabbekoorn, Jesse, and Bell (2003)). Taken together, we still know little about the functional mechanisms underlying patterns of cultural diversity in free-living populations.

The house finch (*Haemorhous mexicanus*) is a non-territorial and socially monogamous passerine (Geoffrey E Hill, 1993). Its range in eastern North America expanded rapidly after it was introduced into Long Island, New York, around 1940 from California (Elliott & Arbib Jr, 1953). Remarkable differences in morphology (e.g. plumage ornamentation) and behaviors (e.g. migratory behavior) have been suggested between eastern and western house finches (Able & Belthoff, 1998; Aldrich & Weske, 1978; Badyaev & Hill, 2000). Genetic studies (Wang, Baker, Hill, & Edwards, 2003) have shown that the eastern population has neutralized the bottleneck effects of the founder event and grown rapidly.

The male house finch song is a long, rambling, hoarse warble composed of a series of short notes (Bitterbaum & Baptista, 1979; P. Mundinger, 1975; Pytte, 1997; Thompson, 1960). Typically the syllable repertoire is large for each individual (P. Mundinger, 1975, 1982; Pytte, 1997; Tracy & Baker, 1999). House finch song is used primarily for mate attraction. Many studies have documented that sexual selection favors complex songs with greater syllable diversity (Bitterbaum & Baptista, 1979; Nolan & Hill, 2004; Searcy & Yasukawa, 1996). Nevertheless, theory proposes that the significance of cultural traits in sexual selection might be undermined by other sexually selected traits with greater detectability or with better adaptation to current environmental conditions, such as plumage ornamentation (Badyaev, Hill, & Weckworth, 2002).

In a previous study of contemporary songs in western Long Island in New York, I observed an absence of shared song types across individuals and suggested the geographic pattern, which was neither simply clinal change with geographic distance nor discrete dialects, might be due to partial migration, potential adaptive processes or multiple founder effects (see Chapter 2). Nevertheless, clarification of the underlying mechanism was hampered by lack of temporal evidence. Temporal stability in the acoustic structure of bird songs has proved to be a critical variable for the study of functional mechanisms underlying cultural evolution and the persistence of cultural variation (Podos & Warren, 2007; Rothstein & Fleischer, 1987). In this study I took advantage of house finch songs also recorded in western Long Island in New York in 1975, thus enabling me to perform a comparative analysis across an interval of 37 years in order to detect and measure cultural change of songs over time. If dispersal and migration were indeed major forces shaping the contemporary landscape of acoustic variation, I expected that songs from 1975 would display a relatively high degree of sharing in songs and song elements, and that the geographic pattern would be the form of dialect, as Dr. Mundinger reported in an area 25 km away from the focal area in this study. I had three specific aims:

(1) To assess the extent of song and song element sharing across individuals and describe the geographic pattern of song variation from 1975, as I did for 2012 (see Chapter 2). The study area was one of the introduction sites of the eastern population of the house finch in 1940. I predicted song element sharing among individuals would likely have been high shortly after introduction due to a founder effect. Distinct dialect areas with sharp boundaries were reported previously in Westchester County, New York (~25km away from this study's focal area) (P. Mundinger, 1975). I expected that the present study's focal area would display a similar pattern in the 1970s. To achieve this aim, I examined pairwise similarities of song and syllable

repertoires between individuals and examined variation in song element similarity across the study area.

(2) To examine how songs and song elements change over time. Two types of analyses are widely adopted to detect temporal patterns of cultural change. Some studies have evaluated change in cultural diversity and the geographic distribution of songs or song elements (Byers et al., 2010; Fayet, Tobias, Hintzen, & Seddon, 2014; Lachlan & Feldman, 2003; Lang & Barlow, 1997; Lynch & Baker, 1993, 1994), while others have tracked specific changes in the songs or the song elements over a long period of time (Byers et al., 2010; D. A. Nelson, Hallberg, & Soha, 2004; Williams et al., 2013). I performed both analyses in this study. I tested for difference in lexicon sizes and geographic patterns between years, and also tracked specific changes over time in terms of acoustic features, song complexity, and song type consistency.

(3) To distinguish between different diversifying influences and adaptive processes, and discuss how their relative importance plays out over the divergence of house finch songs in my study area. Despite accumulating evidence for the major roles of a number of driving forces in acoustic divergence, the interplay among these drivers of spatial and temporal variation shaping cultural diversity remain unclear in most studies, and are not entirely illuminated by the present study either. I discuss the kind of evidence that would be helpful in order to differentiate the relative significance of drivers of acoustic divergence, in turn providing a clearer hypothesis-testing framework for future work.

Methods:

Study Area

I collected songs of wild house finches previously recorded in western Long Island, New York, in 1975 (by P. C. Mundinger) and 2012 (by F. C. Geller). The songs from 1975 were recorded in a roughly 15km (north to south) by 42km (east to west) area that overlapped with the 2012 study area at Hewlett Bary Park (40.640° N 73.700° W) and Freeport (40.584° N 73.560° W): from Freeport (40.584° N 73.560° W) in the south to Bay Shore (40.717° N 73.243° W) in the north, and from Alantic Beach West (40.588° N 73.741° W) in the west to Bay Shore (40.717° N 73.243° W) in the east (Figure. 3.1). The songs from 2012 (the same as those analyzed in Chapter 2) were recorded in a roughly 34 km (north to south) by 40 km (east to west) area in western Long Island, New York: from Long Beach in the south (40.589° N 73.667° W) to Port Washington in the north (40.895° N 73.635° W), and from Bay Ridge in the west (40.631° N 74.036° W) to East Meadow in the east (40.729° N 73.566° W).

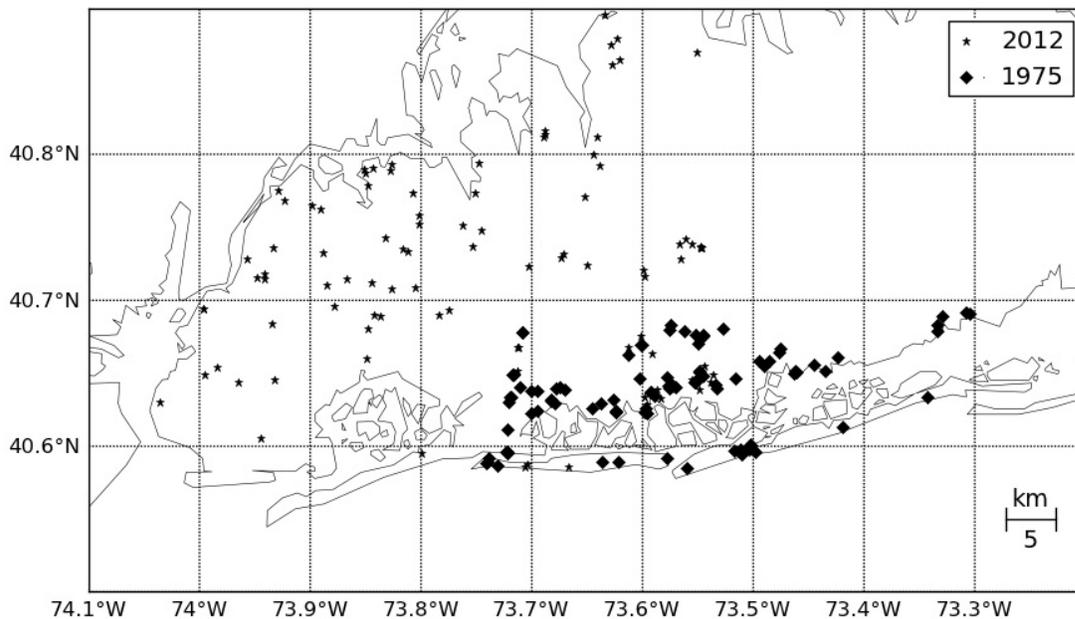


Figure 3.1. Sampling localities of 94 individuals (diamonds) from 1975 and 96 individuals (stars) from 2012 in western Long Island

Song Sample

The songs were recorded in the field in April-July of 1975 and 2012. Individuals were recorded until they stopped singing or flew away. To reduce the chance of duplication from the same bird, each site was visited only once. Within each site, only one individual were recorded within 100 meters (P. Mundinger, 1975). I excluded recordings with multiple song bouts. My previous study revealed that 5 songs could permit the retrieval of on average 75% of the full repertoire of syllables types of an individual house finch (see Chapter 2). Hence, in this study I decided to use five songs as the threshold sample size to select birds for subsequent analysis. By this cutoff, 94 individuals (mean±SD of 11.2±5.1 songs per individual; the total number of songs is 1042) and 96 (mean±SD of 10.2±4.5 songs per individual; the total number of songs is 981) were analyzed from 1975 and 2012, respectively.

Identifying Song and Syllable Types

Syllable Types: A syllable is defined here as a contiguous signal separated from others within a song by at least 20 ms (P. Mundinger, 1975). Identifying syllable types in house finches has proved not amenable to an approach based on the visual inspection of song spectrograms, because variation in syllable structure appears to be continuous, which would result in increased subjectivity in visual classification and probably a decrease in reproducibility among different investigators. I therefore based syllable classification on an automated similarity measurement procedure (FinchCatcher) that compares a number of shape-related acoustic features to generate overall similarity scores between syllables (see Chapter 1). To classify syllable types, first I performed unsupervised hierarchical clustering (“hclust” function from the R package “stats”) using Euclidean distance on these similarity scores. This method makes no assumptions about the number of clusters in the data, so it is appropriate for my analyses because I do not have prior knowledge of syllable types. Many methods of determining the optimal number of clusters, such as the *Akaike Information Criterion* (AIC) or Bayesian Information Criterion (BIC), which are optimized for easily clustered data with discrete separation among groups (D. Anderson, 2008), would not handle continuous variation well in my data; therefore I applied dynamic tree cuts to facilitate the determination of clusters (Langfelder, Zhang, & Horvath, 2008) (“dynamicTreeCut” in R packages). In contrast to the method that applies a constant height threshold to cut dendrogram for clusters, dynamic hybrid cut builds the clusters in a bottom-up manner. First, it respects the order of hierarchical structures in the dendrogram and detects groups of syllables with high similarity by applying a stringent threshold. Second, all previously unassigned objects are reapportioned to defined clusters in the first step based on their distances (the second threshold)

to these clusters. Thus the dendrogram structure is ignored and only the dissimilarity information is used (Langfelder et al., 2008). In this study, all syllables from both datasets were combined for classification. The results of classification were manually examined. In order to take advantage of results of acoustic analysis of song data from 2012 in the previous analysis (see Chapter 2), I compared defined types in this study with types reported previously using a distance-based approach. The distance between any two clusters of syllable types is taken to be the average of all distances between pairs of all objects in these two clusters. If the distance between types from the current analysis and the previous analysis was lower than the second threshold applied in the classification of combined dataset, I considered them as the same type.

Song Types: Different house finch songs are generally quite distinctive (see below). Visual matching has been proven a successful way to yield reproducible song type classification among different observers (P. Mundinger, 1982; Tracy & Baker, 1999). Nevertheless, quantitative measurement of similarities between different song types was impossible by this approach. To address this issue, I transformed each song into a sequence of strings, each of which indicates a distinct type defined by the syllable classification. Then I calculated the edit (Levenshtein) distance (Ristad & Yianilos, 1998) between each pair of song sequences, by which the minimum number of operations required to transform one string into the other was returned. Last, the similarity score of the pair of song sequences was estimated as follows:

$$S=1-(E-D)/L,$$

Where E = edit distance between two sequences, D = the difference in lengths (numbers of syllable) between two songs and L = the shorter length of any pair of songs in comparison. This

equation adjusts similarity scores to account for difference in lengths of songs in comparison, to ensure that an incomplete song would not be misclassified as a different song type from the same song completed. Traditionally, researchers have used a 75% similarity criterion to define distinct song types, whereby at least 75% of song elements are common and arranged in the same order between songs (P. Mundinger, 1975; Tracy & Baker, 1999). While this cutoff has proven useful to eliminate obvious misclassifications that can place some visually very different songs together, it is inevitably somewhat arbitrary and the resulting clusters might not reflect natural song-type categories. To deal with this problem, I clustered songs within an individual's repertoire using the single-linkage hierarchical clustering algorithm (nearest neighbour clustering) and measured the Global Silhouette Index (Rousseeuw, 1987) at different cut heights (from 0.05 to 0.65 at a interval of 0.05) to detect natural categories in the data. Different song types within an individual house finch are distinctive; therefore a range of cut heights might result in the same best classification result (the highest averaged Silhouette width). For example, two different types with between-type similarity scores lower than 0.2 and with-type similarity scores larger than 0.75 would be separated by any value of cut height from 0.2 to 0.75. I determined the best similarity criterion for song type distinction as the one that reaches the highest averaged Silhouette width (ASW) in a majority of individuals. Paired songs with $S >$ the best similarity criterion are considered the same song type.

Assessing Song and Syllable Sharing

Song Sharing: I averaged the similarity scores of all pairwise song comparisons between each pair of individuals to generate overall similarity scores between individuals. This measure was useful to find pairs of individuals sharing entire song repertoires, although this was a rare

occurrence. Usually only one song type was shared, so the highest similarity scores between individuals was generally sufficient to detect sharing of any song type.

Syllable Sharing (by distance): To determine the degree of syllable sharing among birds, I used Jaccard's similarity coefficient (S_j), a method commonly used to measure similarity for binary data (presence/absence) in acoustic analysis (Krebs, 1989; Podos, Peters, Rudnicki, Marler, & Nowicki, 1992). S_j was calculated as follows:

$$S_j = a / (a + b + c)$$

where a = the number of shared syllables in both individual repertoires, b = the number of unique syllables in the first individual, c = the number of unique syllables in the second individual. The similarity score ranges from 0 (no sharing) to 1 (complete sharing). Nevertheless, the maximum possible S_j for the comparison would be <1 because sizes of syllable repertoires are unlikely to be equal between individuals. So I adjusted S_j to account for differences in sample size as follows:

$$S_j(\text{adj}) = a / (a + b + c - d)$$

where d is the difference in detected repertoire size between individuals. I calculated $S_j(\text{adj})$ for each individual compared with all other individuals ($N=4371$ pair-wise comparison). I also applied a Mantel test within each year (Legendre & Fortin, 1989) to determine the relationship between the sharing of syllable repertoires and distance. To test for patterns of localized syllable composition, I performed the "unweighted pair group method with arithmetic mean" (UPGMA) (Sokal, 1958) to group individuals based on similarity of their syllable repertoires. Similarities

between individuals were estimated by Pearson correlation on the presence/absence of all syllable types. All analyses were carried out using R package “ecodist”.

Assessing Change over Time

Change in Syllable Sharing: To examine how the spatial structure of syllable sharing changed over time, I grouped average $S_j(\text{adj})$ values into 5-km distance intervals and then compared means between years at different distance classes. I performed t-tests on the results, followed by correction for multiple comparisons (Bonferroni, 1936).

Change in Syllable Lexicon: The syllable lexicon changed over time. I examined whether the presence of syllable types in both years could be explained by frequencies of types among individuals (i.e. the proportion of individuals having a given type in their syllable repertoire) in each year by logistic regression. I also determined whether persistent syllable types differed from other syllable types only present in one year, in 9 vocal parameters, including “Average frequency”, “Start frequency”, “End frequency”, “Bandwidth”, “Frequency-time excursion length”, “Duration”, “Highest frequency”, “Lowest frequency” and “Concavity” (see Table 3.1).

Change in Vocal Consistency: Vocal consistency is a measure of the tendency of birds to perform the same song/syllable type from rendition to rendition. I assessed song type consistency and its change over time. I estimated song type consistency/variability following the idea proposed by Anderson’s test of multivariate dispersion (M. J. Anderson, 2006): after a principal coordinates

ordination (PCO) of the dissimilarity matrix for all songs of a given song type, the distance between each song and the spatial median of this song type was calculated. Thus, the score of song type consistency/variability would be the averaged distance of each song to the spatial median. Larger distance to the spatial median represents greater variability. I averaged scores within each individual's repertoire to generate individual scores of song type consistency. These scores were then compared between years using a t test. To evaluate the minimum required sample size of songs within a type in order to precisely assess song type consistency, I selected 20 song types with more than 15 songs, randomly picked X ($X=2,3,4,5,6,7,8$) songs from each type, calculated scores of song type consistency for new sets of data within each type, and then compared the distribution of these scores with that of the scores based on full set of data from these 20 song types using the Wilcoxon signed-rank test. This procedure was repeated 100 times for each X. I determined the minimum required size as the smallest X at which the chance of yielding a significant difference between distributions of scores based on subsampled and full sets of songs is less than 5%.

Change in Song Complexity: Complexity can be measured as the repertoire size of either distinct song types or syllable types for a particular male. Considering that the sample size of songs recorded per individual is likely positively associated with the number of both song and syllable types, I applied multiple regression (independent variables: year and sample size per individual; dependent variable: Repertoire sizes of syllables or songs) analysis to test if there was significant change in song complexity at the level of both song and syllable.

Change in Acoustic Features: I define ten vocal parameters at the song level in Table 3.1. I compare means of each of these vocal parameters between eras with t tests, and I also compare their variability with Levene’s tests. “Frequency-time excursion length” could be biased by different scales of the time vs. frequency axes; thus I standardized frequency so that its population mean equals that of duration before estimating “Frequency-time excursion length”. Based on the concept of motor limits to birdsong production (Podos and Nowicki, 2004), several of these features might relate to motor performance: “Bandwidth”, “Frequency-time excursion length”, “Intersyllable distance”, and “Concavity”.

Table 3.1. Measured features of the house finch vocal profile

Variables	Description
Average frequency (Hz)-AF	The average frequency of the entire song (syllable)
Start frequency (Hz)-SF	The frequency at the beginning of the song (syllable)
End frequency (Hz)-EF	The frequency at the end of the song (syllable)
Bandwidth (Hz)-BW	The average difference between highest and lowest frequencies of each syllable in the song (The difference between highest and lowest frequencies of the syllable)

Frequency-time excursion length-FL	Arc length of a song (syllable) treated as continuous by connecting amplitude peaks in the spectrogram, using line segments to create a polygonal path
Duration (ms)-DU	Time duration of the song (syllable) without silent intervals
Highest frequency (Hz)-HF	The average highest frequency of each syllable of the song (The highest frequency of each syllable of the syllable)
Lowest frequency (Hz)-LF	The average lowest frequency of each syllable of the song (The lowest frequency of this syllable)
Concavity-CO	The average number of changes in sign of slope of frequency change per syllable throughout the song (Number of changes in sign of slope of frequency change throughout the syllable)
Intersyllable distance-IS	The average duration of silence between consecutive syllables in the song (not applicable at the level of syllable)

Results:

Definition of Song Type

Instead of applying the similarity criterion used by other scientists (P. Mundinger, 1975; Tracy & Baker, 1999) to define song type in house finch, I used Global Silhouette Index to determine a consistent cutoff to detect natural categories of song types within individuals (Figure 3.2).

Average silhouette width (ASW) is a measure of how appropriately the data has been clustered.

The higher score of ASW, the better separation among clusters. Therefore, maximum ASW indicates the best clustering quality and the optimal number of clusters. As suggested by other researchers of house finch (P. Mundinger, 1975), different song types are quite distinctive. The results suggested that a cut height of 0.3 (or a similarity criterion of 0.7), by which most individuals can have the best results of clustering of within-individual songs, is robust and most likely to generate natural groups of songs.

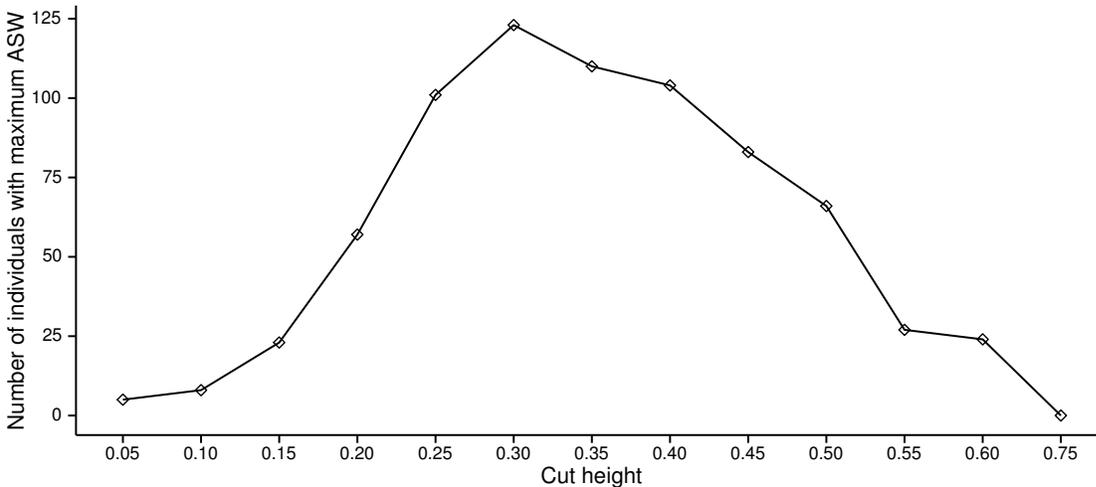


Figure 3.2. Validity of similarity criteria to define song type

Line plot to show by a given cut height the resulting classification of song data of how many individuals can achieve the maximum average silhouette width (ASW score). Maximum ASW indicates the best clustering quality and the optimal number of clusters. The peak is at $x=0.3$ and $y=123$. This means by the similarity criterion of 0.7 most individuals can reach the best separation of within-individual songs.

Songs of 1975

Songs lasted from 1 to 5 s (mean \pm SD of 2.33 \pm 1.1s) including intersyllable intervals, and were comprised of 6 to 31 syllables (mean \pm SD of 16.32 \pm 3.6). Birds sang 1 to 7 different song types (mean \pm SD of 3.3 \pm 1.4). However, the number of detected types depended on the number of recordings per individual (see below). As previously noted (Bitterbaum & Baptista, 1979; P. Mundinger, 1975; Tracy & Baker, 1999), song variation within a type was common, manifested as insertion or deletion of one to three syllables or incomplete songs (see below).

Sharing of song types was found according to the 70% similarity criterion (Figure 3.3b, 3.4a). Nevertheless, sharing of the entire detected song repertoires was rare, excluding pairs between

which one individual performed only one song type (Figure 3.3a). Only two pairs of individuals that sang more than one song type shared all of them with each other.

Out of 4371 pair-wise comparisons between individuals, 132 pairs shared at least one song type.

Among these, 91% (n=120) were between individuals less than 5km apart (Figure 3.4a).

However, even at such distances, males were much more likely to share no song types at all

(number of pairs sharing song type =120; number of pairs sharing no song type =659; Figure

3.4b). In total, 51 song types were shared between house finches (Figure 3.4c). However, only

37% (n=19) were shared by more than two individuals (Figure 3.4c).

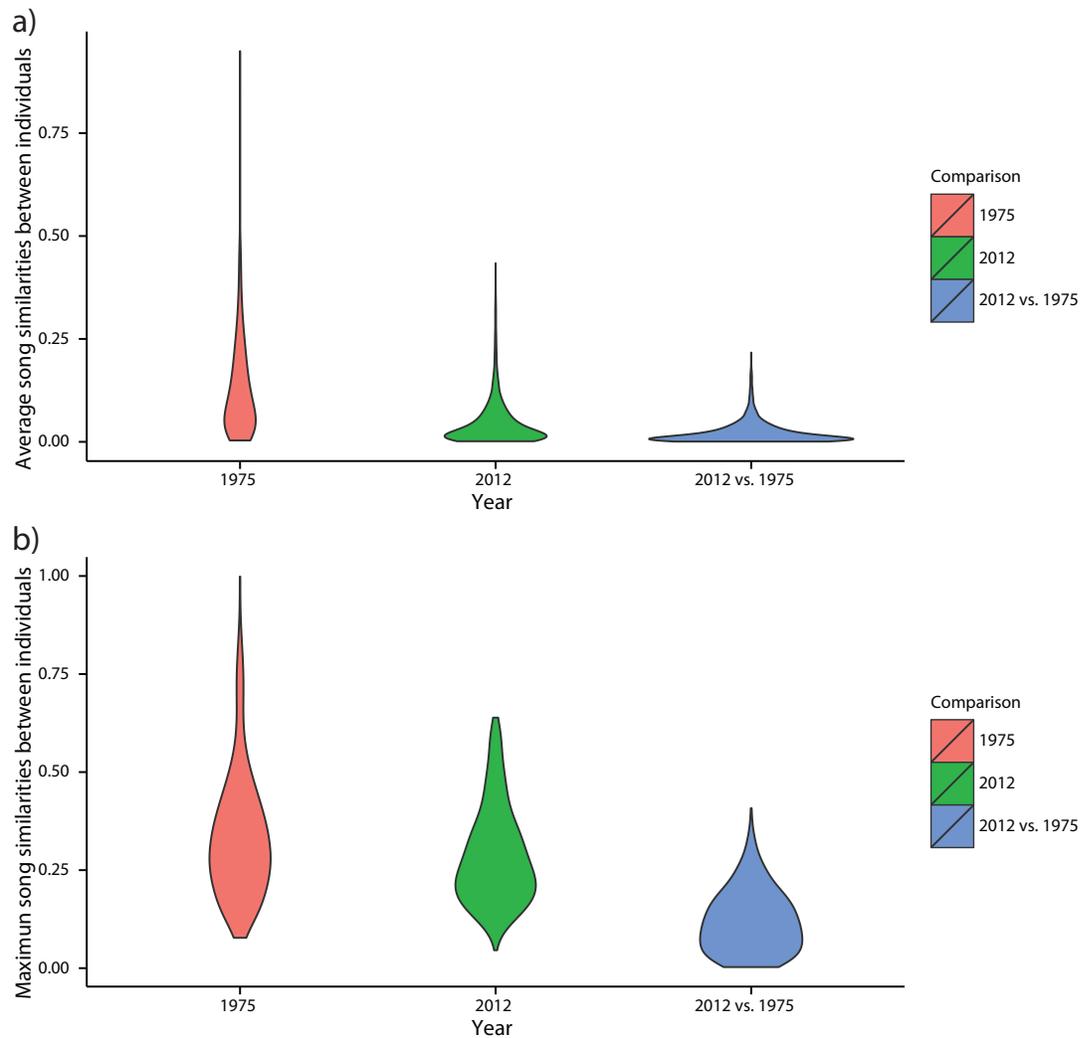


Figure 3.3. Violin plots showing distribution of song similarity scores between individuals

Violin plots showing distribution of song similarity scores between individuals within 1975 (Number of comparisons = 4371), within 2012 (Number of comparisons = 4560), and between 1975 and 2012 (Number of comparisons = 9024; only individuals from different years were compared). The violin plot is similar to a box plot, except that it also shows the probability density of the data at different values. **a)** Average song similarities between individuals; **b)** Maximum song similarities between individuals. Song sharing is found between individuals in 1975, but not in 2012. No shared song types are found between individuals from different years.

Difference in Song Sharing between Years

In contrast to the absence of sharing of types among individuals in 2012 (see Chapter 2), sharing of song types was not rare in 1975 (Figure 3.3b). In total, 51 types were shared between house finches (Figure 3.4c).

Difference in Song Similarity between Years

On average song similarity between individuals were higher in 1975 than in 2012 ($t=-43.7198$, $df=6467.142$, $P<0.0001$; Figure 3.3a). The average song similarity between individuals was still higher in 1975 even after I excluded shared song types in estimation of the average song similarity between individuals ($t=-39.2912$, $df=6467.142$, $P<0.0001$)

Temporal lability in Song Structure

Strikingly, no single song type was present in both years (Figure 3.3b), even when comparing individuals of different years from the same areas (distance < 5km) in (1) Long beach (number of birds from 2012=3, on average 12.3 songs per individual; number of birds from 1975 =8, on average 10.1 song per individual) (2) Freeport (number of birds from 2012=14, on average 14.5 songs per individual; number of birds from 1975 =15, on average 11.1 song per individual) and (3) Hewlett Bary Park (number of birds from 2012=3, on average 9.5 songs per individual; number of birds from 1975 =11, on average 13.1 song per individual) . This finding indicated that syllable sequences in house finch songs were not well preserved over a long period of time.

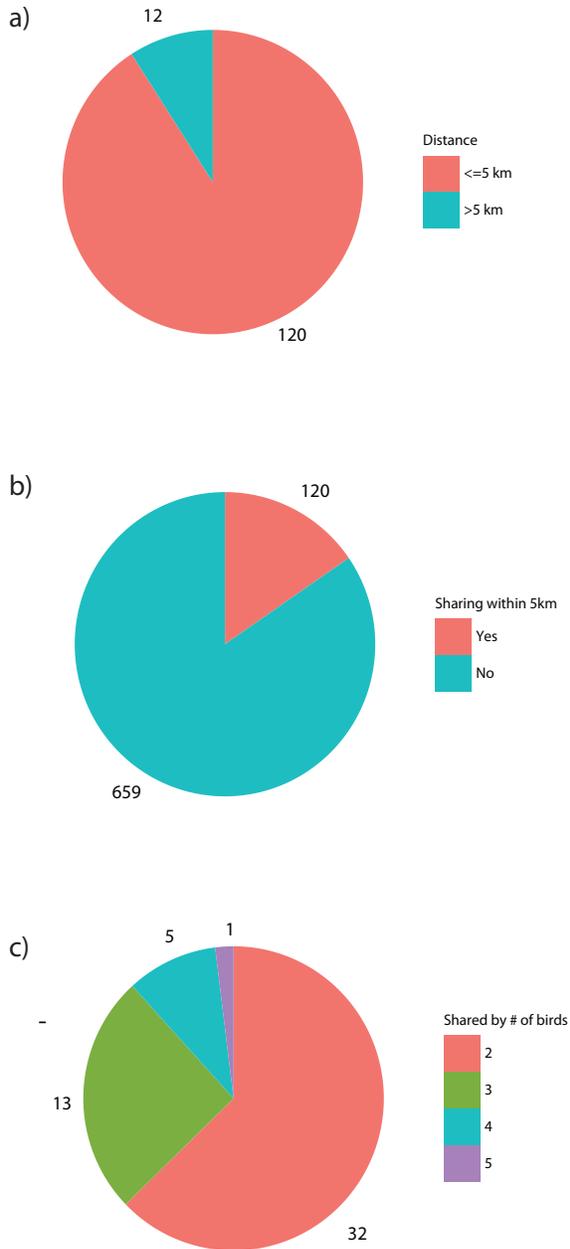


Figure 3.4. Pie charts illustrating song type sharing among house finches in Western Long Island in 1975.

Pie charts illustrating song type sharing among house finches (Number of individuals = 94; number of pair-wise comparisons = 4371) in Western Long Island in 1975. Counts of each marked category appear besides corresponding pie sections. **a)** Song type sharing by distance. **b)** The ratio of sharing versus non-sharing within 5 km. **c)** The number of birds sharing a given type.

Syllable Identification

A total of 407 distinct syllable types (out of 17005 syllables from 1975 plus 11968 syllables from 2012 = 28973) were defined in the classification analysis of the pooled sample of 1975 and 2012 songs, excluding 241 syllables whose quality was poor and or which were outliers in the classification. 238 and 287 syllable types comprised the 1975 and 2012 lexicons, respectively. Among these, 118 syllable types were present in both years (hereafter: persistent types: Figure 3.5a). Syllables were generally shared by a larger proportion of the population in 1975 than in 2012 ($t=7.4677$, $df=446.243$, $P=4.325e-13$) (Figure 3.5b,c). Persistent types were generally present in more individuals' repertoires in 1975 than types found only in 1975 (logistic regression: $z=5.935$, $P=2.93e-9$), but this trend was much weaker in 2012 and not significant (logistic regression: $z=1.883$, $P=0.0597$). This result suggests that persistence of types over time was due to a stochastic process, in which rare types that shared by few individuals were likely to be lost. I did not find significant differences in the mean or variance of any vocal parameter between persistent types and unique (year-specific) types in each year (P ranges from 0.1021 to 0.4812), suggesting that different syllable types are biologically equivalent.

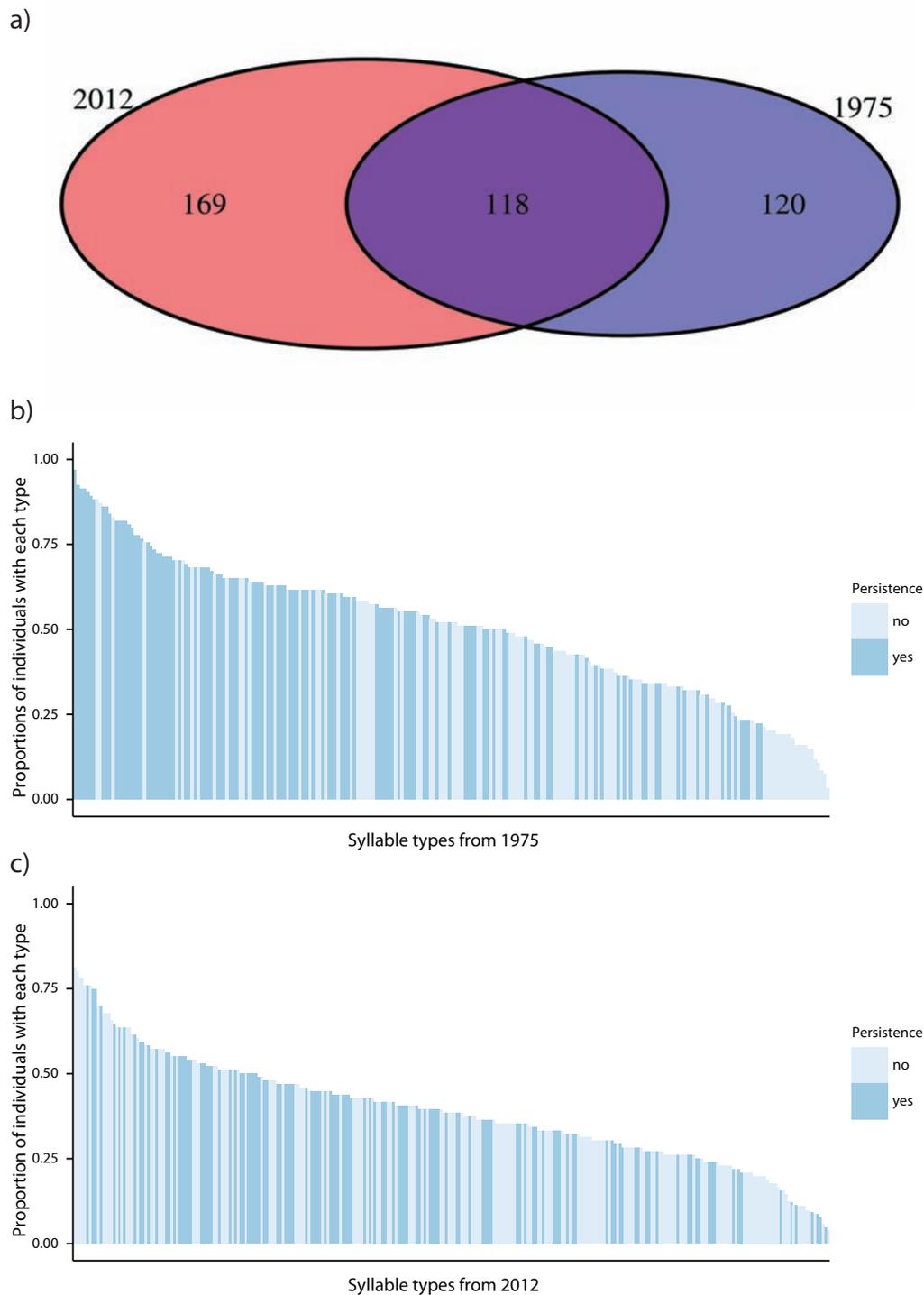


Figure 3.5. Results of syllable classification and frequencies of identified syllable types in each year

a) Venn diagram representing syllables present in the lexicons of both 1975 and 2012 and their intersection (persisting syllable types). **b)** Bar plot showing the proportion of the population

sharing each syllable type in 1975. Types were sorted according to their frequencies in the population. In total 238 types were detected. c) Bar plot showing the proportion of the population sharing each syllable type in 2012. Types were sorted according to their frequencies in the population. In total 287 types were detected. Types present in both years are colored dark blue, while types unique to one year are light blue.

Difference in Syllable Sharing between Years

In Chapter 2, I reported that sharing of syllable repertoires between individuals, measured by Jaccard's similarity coefficient ($S_j(\text{adj})$), declined with distance in 2012. I found a similar relationship between $S_j(\text{adj})$ and distance in the song data of 1975 (Mantel test, $r=-0.6129$, $P=0.001$, number of permutations=999) (Figure 3.6a). However, significantly higher sharing of syllable repertoires was observed in 1975 in every class of distance (adjusted $P<0.0001$ in all classes of distance) (Figure 3.6b). This result is consistent with the large proportion of population sharing syllable types in 1975.

Geographic Pattern of Syllable Composition in 1975

I performed UPGMA clustering based on individuals' syllable repertoires to detect groups of individuals with distinct syllable composition. Results of hierarchical clustering suggested individuals could be separated into two groups (Figure 3.7a). Discrete types were found in these two groups. However, individuals of different groups were not completely separated geographically (Figure 3.7b). Thus there was a region (from 73.6° W to 73.5° W) of mixed individuals of different syllable composition that could not be distinguished solely based on similarity in syllable repertoire. This appears to represent a "hybrid zone" between two otherwise discrete regions.

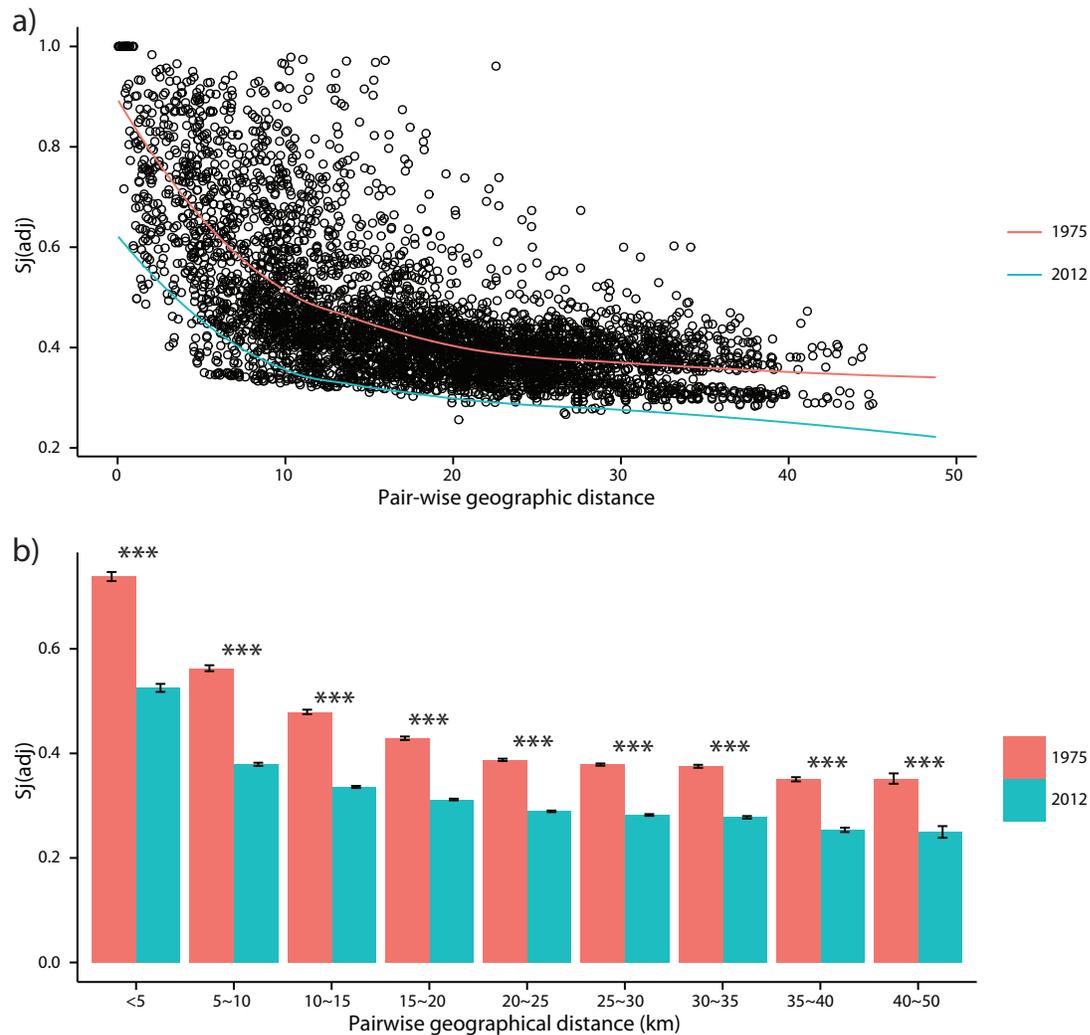


Figure 3.6. Difference in sharing of house finch syllable repertoires between 1975 and 2012

a) Scatterplot showing relationship between $S_j(\text{adj})$ values for syllable sharing by geographic distance for 1975 songs. The red line was plotted based on loess regression. Grey areas indicate 95% confidence of intervals. Similarity of syllable repertoires decreased exponentially with distance. For comparison, the loess regression line for the song data of 2012 was added (in green). **b)** Results of student t test comparing means of $S_j(\text{adj})$ values between years, grouped into 5-km distance intervals. Syllable sharing are higher in 1975 than in 2012 in every classed of distance. Stars above bars indicate significant differences after multiple-comparison correction (Bonferroni). ***: Adjusted $P < 0.0001$; **: Adjusted $P < 0.001$; *: Adjusted $P < 0.01$

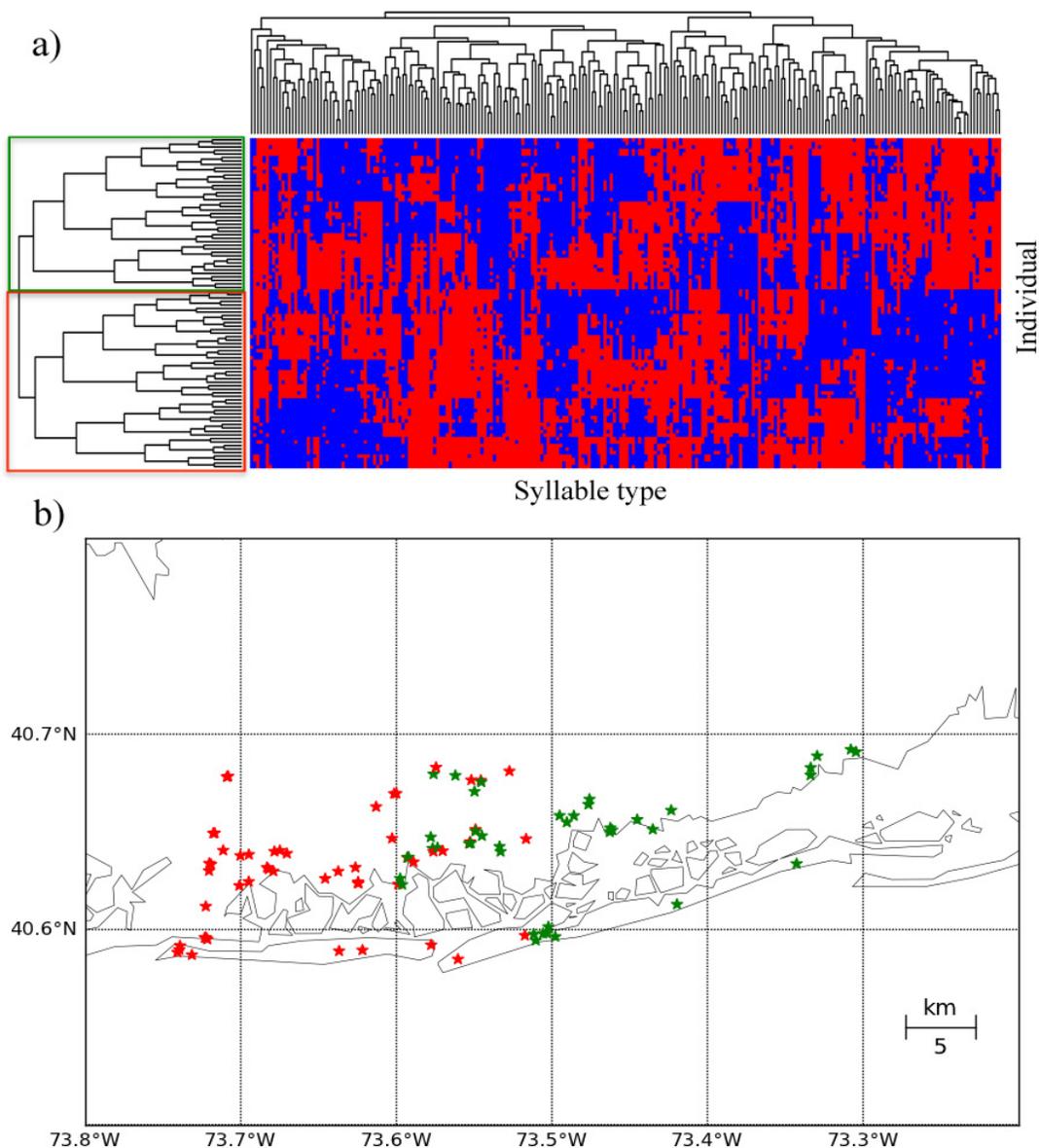


Figure 3.7. Geographic pattern of syllable repertoire sharing based on UPGMA clustering in 1975

a) UPGMA clustering based on similarities of syllable repertoires. Similarities between individuals were estimated by Pearson correlation on presence/absence of all syllable types. A heat map was generated to illustrate the clustering result, in which each row represents an individual and each column a syllable type. Cells in red indicate the presence of types in the corresponding individuals, and blue cells indicate absence. Individuals sharing most types were placed closely. Cluster analysis proposed two major groups of individuals, as indicated in the hierarchical tree to the left. Columns of syllable types present in the similar set of individuals were also clustered together. **b)** Location of individuals clustered into two groups. The two groups of individuals were colored as red (region 1), and green (region 2).

Change in song consistency

To determine the minimum sample size required within a type to accurately reflect within-individual song consistency across renditions of a song type, I simulated the distribution of scores for within-type consistency (measured as distance of each renditions of a song to the spatial median for all renditions) at a sample size from 2 to 8, from 20 defined song types of which more than 15 renditions each were recorded. Results showed that the distribution of scores for within-type consistency of simulated data at the sample size of 6 has a chance of only 4% to be significantly different from that based on the full set of songs from the same 20 types (Figure 3.8). Further increasing the threshold of sample size would be at a high cost of individuals that could be analyzed. Hence, I decided to include song types with at least 6 songs for this analysis. In the end, 51 (58 song types) and 42 (45 song types) individuals were qualified in 1975 and 2012, respectively. I observed a significant decrease in within-type consistency between years (t test: $t = -3.4817$, $df = 76.09$, $P = 0.0008$; Figure 3.9). This finding suggested that the structure of songs within a type was less rigid and consistent in 2012 than in 1975.

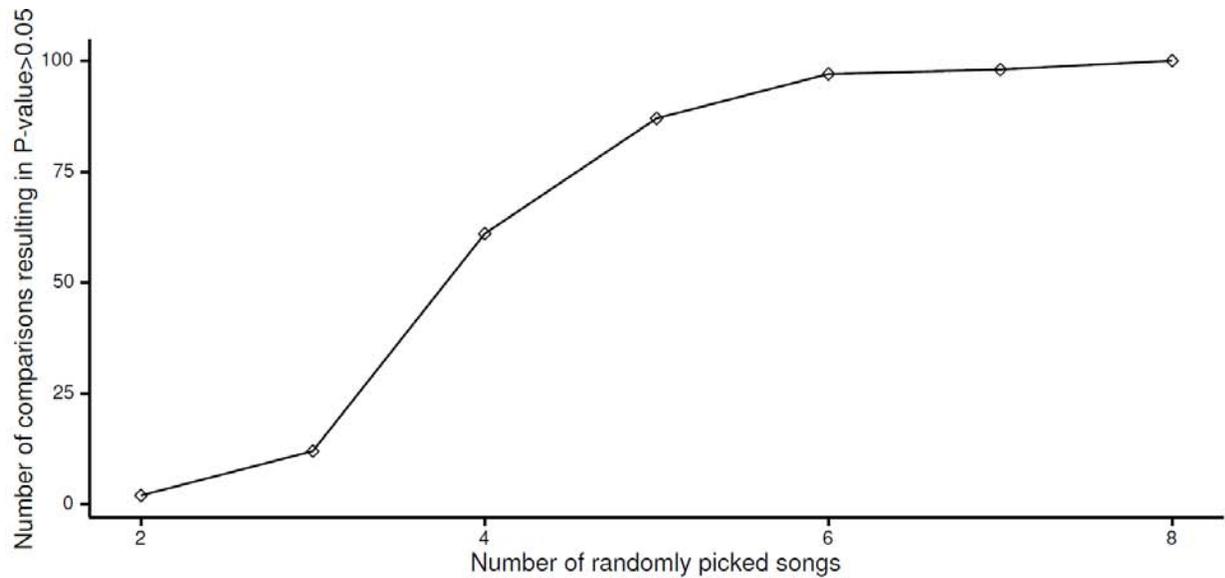


Figure 3.8. Validation of sample size to study song type consistency

Line plot to show how many times (Y-axis) out of 100 simulations the distribution of scores for within-type consistency of x ($x = 2$ to 8) randomly picked songs from each of 20 song types is not significantly different (Wilcoxon signed-rank test: $P > 0.05$) from that of the full set of songs from the same 20 types. The higher value of Y for a given x means the chance of obtaining a significantly different distribution of scores for within-type consistency based on x randomly picked songs compared to the full sample is smaller. When $x=6$, $y=97$. This indicates that the chance of obtaining a different distribution of scores for within-type consistency based on 6 randomly picked songs is 3% (less than 5%).

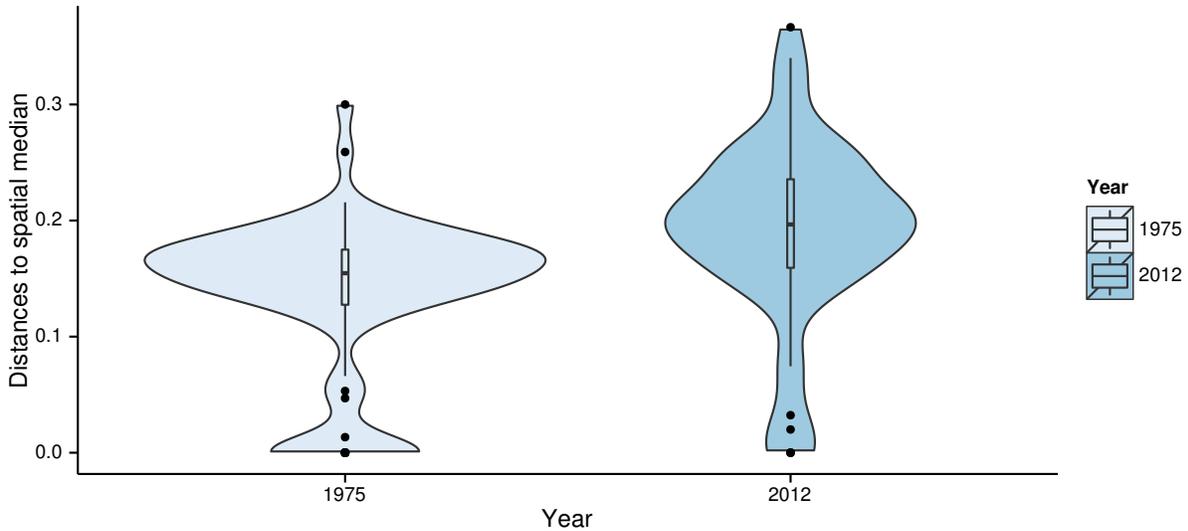


Figure 3.9. Within-individual song type consistency between years

Violin and bar plots to show distribution of distance to spatial median of defined types for individuals in 1975 or 2012 (scores of within-type consistency were first averaged within individuals whose had more than one qualified types). The violin plot is similar to box plots, except that it also shows the probability density of the data at different values. Song type consistency is significantly higher in 1975 than 2012.

Change in song complexity

The repertoire sizes of songs per individual were larger in 1975 than in 2012, though not significant, when the sample sizes of recorded songs per individual were controlled (multiple linear regression: $t(\text{year})=-0.640$, $P(\text{year})=0.523$; $t(\text{sample size})=4.258$, $P(\text{sample size})=3.26e-5$) (Figure 3.10a). Despite the fact that the numbers of detected song types were highly variable for a given sample size, the positive association between the sample sizes of songs per individual and the numbers of detected song types was significant, suggesting that the sample size (number of recorded songs) of many individuals in my study might be too small to retrieve the full repertoire, or a consistent set of song types. However, individuals' syllable repertoire size (1975: mean \pm SD of 80.3 \pm 12.7; 2012: mean \pm SD of 70.6 \pm 18.1) were significantly higher in 1975 than in

2012 (multiple linear regression: $t(\text{year})=-4.172$, $P(\text{year})=4.61e-5$; $t(\text{sample size})=1.971$, $P(\text{sample size})=0.0502$) (Figure 3.10b). Larger syllable repertoires were found in 1975, implying higher song complexity in 1975. The relationship between sample size and amount of detected syllable types per individual was not significant, consistent with my previous result showing that the first five songs could recover on average 75% of the detected syllable repertoire (see Chapter 2).

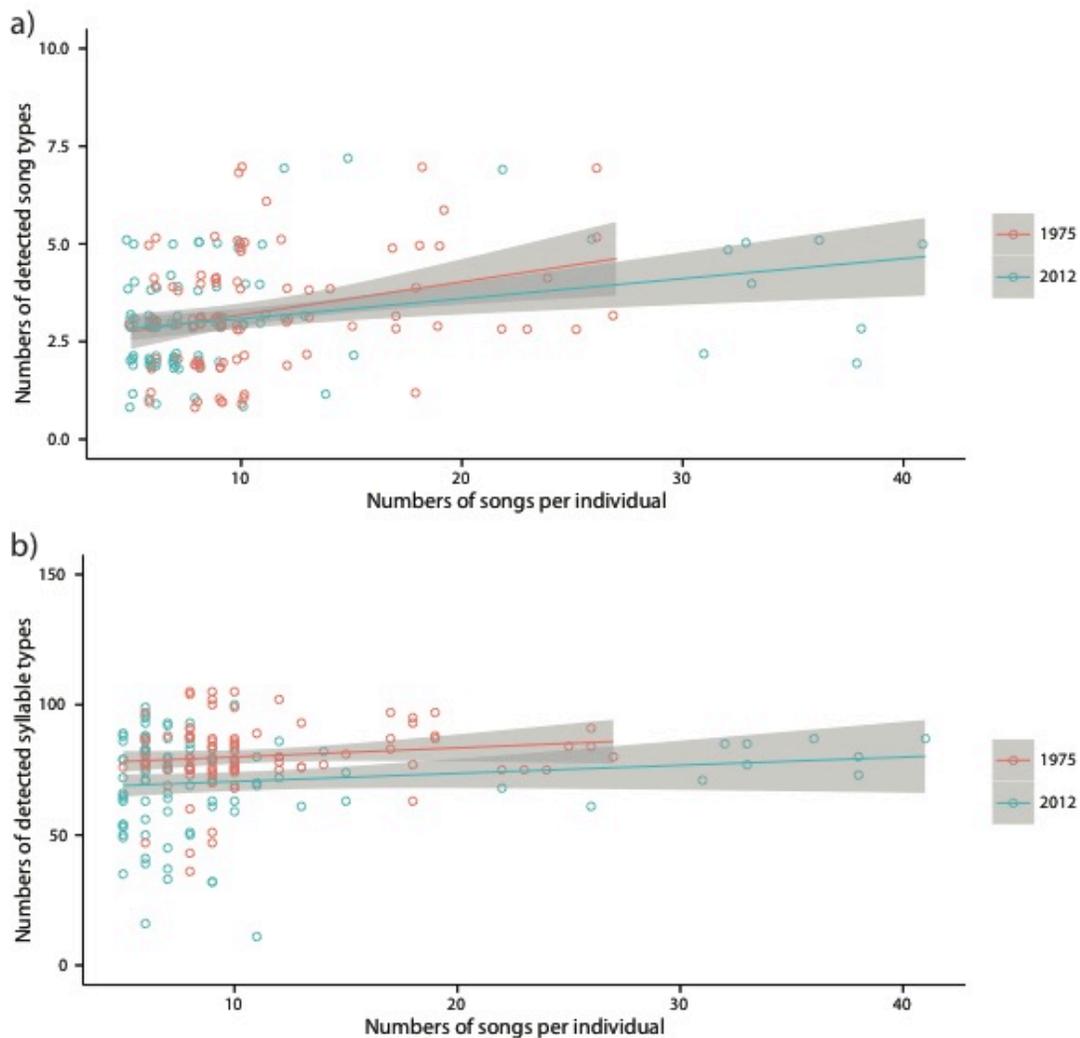


Figure 3.10. Song repertoire size and syllable repertoire size between years

a) Scatterplot showing the difference in song type repertoire size per individual between years, taking into account the number of songs recorded per individual. Data were jittered to unhide

overlapping points. Linear regression lines are represented for 1975 and 2012 data, with the 95% confidence interval in grey. **b)** Scatterplot showing the difference in syllable type repertoire per individual between years, taking into account the number of songs recorded per individual. Linear regression lines are represented for 1975 and 2012 data, with the 95% confidence interval in grey. Both repertoire sizes of song and syllable are higher for individuals in 1975, though significant difference is only found for syllable repertoire.

Change in Acoustic Features

Ten vocal parameters of songs were summarized within individuals and then compared between years (Figure 3.11). None of them exhibited significant difference in means. The smallest p-value was generated by comparison of minimum (lowest) frequency between years, which nearly reached the significance level without correction for multiple comparisons. “Concavity” and “Intersyllable distance” showed significantly greater variability in 2012 than in 1975 (Table 3.2). This result is in line with my previous finding (see Chapter 2) that the three variables “Concavity”, “Lowest frequency” and “Intersyllable distance” displayed high variation across the study range in 2012. (The P of Levene’s test for “Lowest frequency” = 0.08325, close to the threshold of significance).

Table 3.2. Statistics of tests comparing means and variability

Variables	Mean of 1975	Mean of 2012	Students' t test			Levene's test	
			T	DF	P-value	F-value	P-value
SF	3539.873	3458.545	0.6818	186.029	0.4962	1.0146	0.3151
EF	3231.582	3189.697	0.4668	185.475	0.6412	0.0646	0.7997
AF	3460.769	3498.111	-0.4464	186.78	0.6558	0.893	0.3459
HF	4083.61	4013.635	0.5965	187.778	0.5515	0.3454	0.5574
LF	2859.151	3017.517	-1.9515	172.116	0.05262	3.0325	0.08325
BW	1122.207	1172.754	-0.4789	187.966	0.6326	0.4421	0.5069
DU	1.555449	1.481697	0.4107	182.129	0.6817	0.1588	0.6907
FL	2.478321	0.3672681	-0.867	175.838	0.3871	0.855	0.3563
CO	0.9899601	0.921875	1.647	164.509	0.1015	27.607	4.006e-07***
IS	86.79493	93.13587	-0.9825	177.921	0.3272	11.627	0.0007954 ***

*** p-value < 0.0001

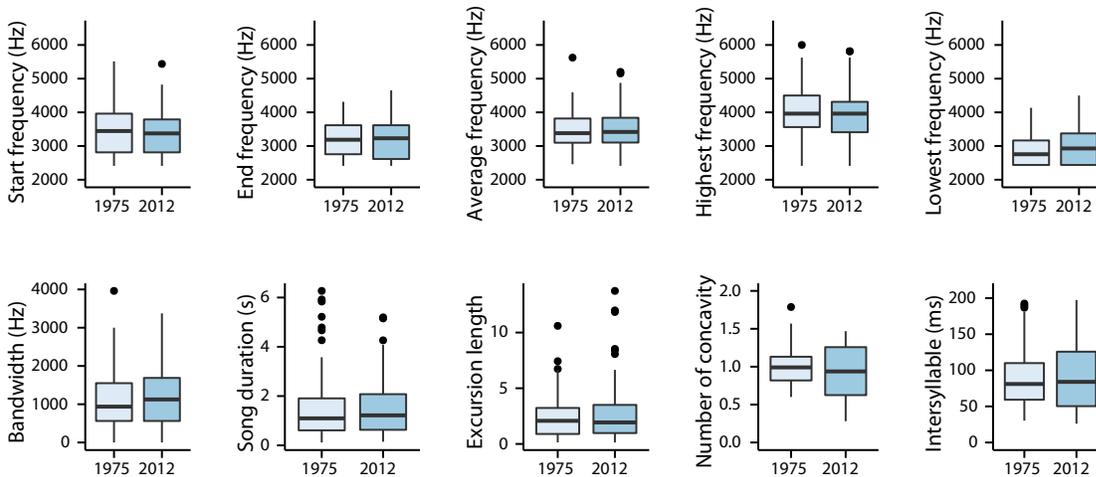


Figure 3.11. Boxplot showing tests comparing means and variability of ten vocal parameters between years

Three variables “Concavity”, “Lowest frequency” and “Intersyllable distance” displayed high variation across the study range in 2012

Discussion:

Song and Song Element Sharing across Individuals and Geographic Pattern in 1975

My analysis did not show a clear pattern of house finch song dialects as Dr. P. Mundinger (1975) found in a region 25 km away from the focal area of my study. Although the songs of a number of individuals were structurally similar with neighbors, more often distinct song types were found between neighboring individuals. At the level of syllables, despite the fact that two groups were distinguished based on similarity of syllable repertoires, members of different clusters actually differed in only a small number of group-specific types. Moreover, a roughly 5 km-wide hybrid region (from west to east) with a mixture of individuals of the two groups’ typical syllable compositions was defined. The findings of hybrid songs comprised of syllable sequences from

distinct regions of two clusters suggest that the hybrid zone could have been a contact region for two founding propagules.

The songs in my study area had evolved for roughly three decades since the colonization event on Long Island in the mid-1940s (Elliott & Arbib Jr, 1953). It is unlikely that the founders (back to mid-1940s) possessed all the cultural diversity that could be found on the east coast in 1975. After the colonization event, new songs or song elements as a result of copy errors during song learning were accumulated within local founder propagules, and eventually subsequent song divergence became sufficient to display distinct identities in local populations before they expanded and integrated with each other during population growth. I hypothesize that my study region might have hosted two founder propagules with highly similar syllable repertoires originally. Three or fewer decades of separation allowed them to develop a small set of unique types before they met in my study area.

On the other hand, limited and localized sharing of song types within each putative propagule area might compromise the grounds of this hypothesis. One possibility is that the issue of small sample size hampered my identification of a clear geographic pattern. P. Munding (1975) retrieved song repertoires based on an average of 17 songs, 1.5 times of the average sample size in my analysis. In addition, I detected a significant dependence of the numbers of detected song types on the number of recorded songs per individual (Figure 3.10a). Moreover, incomplete sharing of song repertoires between individual could likewise be a result of incomplete recording of song repertoires.

Another possible explanation for the difference between syllable and song type sharing could be different evolutionary rates at the level of syllable and song (Catchpole & Slater, 2008). In fact, lower-order units (e.g. syllables) are frequently more stable than higher-order units (songs and syllable sequences) over a long-term period (Ellers & Slabbekoorn, 2003; D. A. Nelson et al., 2004). For instance, some studies have shown that a bird can produce a limited number of syllables, while the recombination rate of these syllables may be relatively high (Lynch & Baker, 1994; Tubaro, 1991). This is also supported by my findings in this study that no shared song type was observed between two eras, while half of the syllables persisted. A final possibility is that the song tradition of the founder propagules in this study area might be different from Dr. Munding's study area, which exhibited stable song traditions.

Cultural evolution over time

Over the course of nearly four decades, the extent of sharing of songs and song elements, lability of song type, song and song element diversity and song performance of house finches in Western Long Island changed substantially. First, at the song level, song types present in the past were totally lost in the contemporary data. Moreover, contemporary birds displayed different song types from all other individuals, even neighboring ones. At the syllable level, half of the syllable types persisted over time, although the possibility could not be ruled out that the observed persistent types nowadays might not be homologous to those back then because of the turnover of syllable types. In addition, syllable diversity increased over time, which manifested as a larger lexicon size and a lower degree of syllable type sharing among individuals in the contemporary

population than in the past. These results suggest that syllable patterns in this house finch population were more stable than song patterns, and that overall similarity between individuals at the level of both song and song element is reduced today compared to several decades ago. No significant difference in the mean or variance of any vocal parameter between persistent types and unique (year-specific) types in each year suggested that different syllable types are functionally equivalent, as shown in a number of studies (e.g. (Catchpole & Rowell, 1993; Horn & Falls, 1988; Lynch & Baker, 1993; McGregor & Krebs, 1982; Payne, Payne, & Doehlert, 1988; Slater, Ince, & Colgan, 1980).

I found a larger lexicon in the contemporary data (20% more syllable types) than in the earlier data. This is unlikely to be due to sampling bias. The numbers of individuals (94 versus 96) and the numbers of recorded songs (1042 versus 981) were similar between 1975 and 2012.

Classification by visual inspection of sound spectrograms and by automated similarity analysis in this study was highly concordant (unpublished data). These results also rule out an artifact hypothesis: if songs from 1975 were less variable than contemporary songs more syllable types from 1975 might have been classified as the same, resulting in lower syllable diversity in the past. If this were the case, smaller individual syllable repertoire sizes would also have been observed in the past data. On the contrary, I found larger syllable repertoires in the past than the present. Therefore, the data indicate an increase in cultural diversity of house finch song in terms of syllable types.

Studies of acoustic variation in natural bird populations have focused on the lability of song structure in evolution (Janes & Ryker, 2013; Williams et al., 2013), with less attention paid to the potential importance of vocal features and song performance related to male quality and targeted by sexual selection. In this study, I specifically examined temporal changes in song complexity, certain acoustic features, and song type consistency that might function as measures of song performance. By applying a robust similarity criterion based on silhouette index rather than the 75% similarity criterion used by others to define song type, I showed that male song in the past exhibited less variation across renditions of a given type. In addition, I found a larger syllable repertoire size per individual in the past. These consistent findings in terms of both song consistency and complexity suggest that the contemporary songs have a poorer performance level than those of the past. No significant differences were observed in vocal features at the level of either syllable or song, which could be explained by one or more of the following: 1) these features were not under selection, but were mainly influenced by physiological processes that did not change over the time period; 2) different selection pressures might have counteracted each other; 3) the plasticity of these features was too high to show a consistent pattern of change; and 4) if selection on song features was relaxed, song features might still have persisted for a long period (Lahti et al., 2009). Despite no significant difference in the mean values of any of the vocal features between eras, results showed that two of them-- “Concavity” and “Intersyllable distance”, measures of syllable complexity and song rate respectively, were more variable nowadays. This implies that some of acoustic features potentially related to performance might be undergoing long-term diversification. This result is consistent with my previous observation that differences in concavity could be related to different regions with localized syllable composition and consequently different influences (see Chapter 2).

Potential Drivers of Spatial and Temporal Variation underlying Cultural Change

1. Demographic factors (greater population size and/or greater population density due to population growth): A genetic study (Wang et al., 2003) has shown that the eastern population has undergone rapid growth and has neutralized the bottleneck effect associated with the founder event. The frequency distribution of songs or song elements within populations has been suggested to fit a neutral model analogous to that for biological evolution in which there is equilibrium between mutation and drift (Byers et al., 2010; Lynch & Baker, 1993). Thus, I predicted greater song and syllable diversity as the population grows, because a denser population means more syllables that can be heard by any given bird, and a larger population means a larger number of syllables in the population as a whole, with new syllable types introduced by copying error and improvisation. Stochastic processes (e.g., drift and mutation) are probably responsible for the persistence or loss of syllable types over time in a frequency-dependent manner, such that rare syllable types in the past were unlikely to be copied by young conspecifics. My result was in accordance with this prediction: common types in the past were more likely than rarer types to persist through time (Figure 3.5b). At the song level, greater population size and/or greater population density today would increase the chance of learning from more tutors, leading to combining smaller chunks of different song types to form new song types.

2. Partial migration and/or increased dispersal distance: The eastern population of the house finch, unlike their sedentary source population, has been shown to be partially migratory: not all individuals migrate, and migratory birds do not migrate every year (Able & Belthoff, 1998). Migration can also result in increased dispersal distance because males going away for the winter

are probably less likely to settle in their natal area than males that spend their first winter there (Paradis, Baillie, Sutherland, & Gregory, 1998). Catchpole and Slater (2008) proposed that migratory songbirds were likely to display more elaborate songs than sedentary conspecifics based on the assumption that sexual selection is more intense in migratory birds. However, tests of this hypothesis have yielded mixed results showing that migratory birds do not necessarily perform complex songs (Byers, 2014; Collins, de Kort, Pérez-Tris, & Tellería, 2009), suggesting the relationship is inconsistent and might be variable. In addition to song complexity, migration plays an important role in shaping local cultural diversity (Fayet et al., 2014; D. Nelson, Marler, & Morton, 1996). In this study area, migration could have increased cultural diversity in the following manner. First, migratory birds could have introduced new songs and song elements from other populations or through a higher rate of song improvisation (Kroodsma, Liu, Goodwin, & Bedell, 1999). In my study, persistent syllable types were no longer dominant types in the contemporary data, which might be partially attributed to migration. Second, the known preference for local songs by female house finches (Hernandez & MacDougall-Shackleton, 2004) might have been relaxed by migration or increased dispersal distance, as limited movement is critical to maintaining local song types or dialects (Marler & Tamura, 1962).

3. Developmental stress: The developmental stress hypothesis proposes that poor early condition could adversely affect song complexity, singing performance (Spencer, Buchanan, Goldsmith, & Catchpole, 2004), and song learning (Nowicki, Searcy, & Peters, 2002). Thus, the accuracy of song learning and adult males' song repertoires could indicate male quality to the female. Developmental stress might therefore account for the greater variability in some vocal features, and reduced song performance as indicated by song type consistency and song complexity.

4. Relaxed selection by females on song performance: Colorful plumage and elaborate songs are the two most prominent and best known sexually selected display traits in the house finch. These two traits may either trade off with each other, as song complexity was found to be strongly negatively correlated with elaboration of plumage in cardueline finches (Badyaev et al., 2002), or evolve independently, supported by a recent study showing that plumage ornaments and song consistency function in different mating contexts (within-pair and extra-pair mating) (Taff et al., 2012). Given the fact that the house finch is socially monogamous (Geoffrey E Hill, 1993), I would expect that when two sexual traits of the male house finch are condition (e.g. resource) dependent, females should prefer the most conspicuous one, and the display of other traits would be reduced correspondingly. The plumage of house finches in New York is more colorful than that of their conspecifics in California (Geoffrey Edward Hill, 2002). This might be partly because colorful males were favored for sale as “Hollywood Finches” when introduced by merchants (Elliott & Arbib, 1953). Thus, “detectability” of the plumage of male house finches to females might have increased, or in other words, plumage might have become more conspicuous to females than song performance in the eastern population. Once a preference was expressed, selection would lead gradually to the further displacement of song performance by plumage, as predicted by Badyaev and Hill (2003); Geoffrey E Hill (1993). In sum, I hypothesize that in the study population females are emphasizing the most conspicuous trait-- plumage—which has resulted in long-term relaxed selection on the performance of male songs. If this has been occurring, greater variation across renditions of a given type and reduced song performance would be expected in the eastern population. In a pilot analysis testing differences in song complexity (syllable repertoire size) and song type consistency between house finches from NY

(1975, N=10) and CA (1982, N=10), I found higher complexity and consistency in songs of CA (data not shown).

Distinguishing between Drivers and Evaluating Their Roles in the Interaction

Distinguishing among potential mechanisms of the changes observed in this study would facilitate understanding of geographic variation and temporal stability in vocal signals. At the syllable level, the increase in syllable diversity might be influenced primarily by demographic factors and migration. However, there might be regional differences in their relative roles (Figure 3.12). For instance, in region 2 from 2012, persistent syllable types across both sample years still retained high sharing between individuals (Figure 3.12a), underscoring the major role of stochastic processes influenced by demographic factors. In region 3 of 2012, on the other hand, most types were shared only by a small proportion of birds, including types that persisted across both sample years (Figure 3.12b). This implies that the syllable pool in this area was diluted significantly by the introduction of new elements, perhaps through bird migration in region 3.

Poorer song performance in the recent dataset could be explained either by developmental stress or relaxed selection by females on the quality of male song. In my previous analysis (see Chapter 2), I proposed lower developmental stress in the region 2 to explain higher concavity. However, this region also exhibited lower song complexity in terms of syllable repertoire size, and lower song consistency across renditions than songs from 1975. This result suggests that different driving forces have influenced different aspects of song performance broadly considered.

The lack of persistence of any song type over time, and the lack of sharing of any song type between individuals in the contemporary population, were surprising to observe. All four hypothesized drivers could contribute to such a substantial increase in song diversity. Greater population size and/or greater population density (demographic factors) would increase the chance of learning from more tutors, leading to combining smaller chunks of different song types to form new song types. Additionally, migrants would bring new song types into the pool. Under relaxed selection or poor developmental condition, males might have a higher tendency to sing inconsistent songs, which increases the chance of copy error and improvisation, resulting in increased song diversity as was observed in 2012 relative to 1975. Demonstrating the relative roles of potential drivers in this regard would be difficult. I hypothesized that syllable types that have persisted through time would be well represented at the song level in individual repertoires if demographic factors have played a major role; whereas if migration has played a major role it would have brought new songs into the population, comprised of a large proportion of new syllables. Results showed a significantly higher proportion of persistent syllable types in each song per individual in region 2 compared to that in region 3 ($t=3.7173$, $df=87.885$, $P=0.0003537$) (Figure 3.13). This provides corroborative evidence at the song level to support the major roles of demographic factors in region 2 and migration in region 3. It is difficult to evaluate the relative roles of relaxed selection and developmental stress in increasing song diversity, but they are no doubt interacting with demographic factors and migration. For example, the interaction of migration and relaxed selection by females on local songs and/or high quality songs would have an additive effect on song diversity. If migrants tend to have more intense sexual selection, this situation would relax selection by females on male song, a situation that would be exacerbated if sexual selection on song were also relaxed because of the importance of bright plumage.

In summary, I have provided a framework for understanding the spatiotemporal variation in house finch song in Western Long Island, incorporating partial migration, population growth, relaxed selection by females on male song, and developmental stress playing interacting roles. I have also proposed a few ways to distinguish different factors (particularly demographic factors and migration), such as analyzing the prevalence of persistent syllable types at the syllable level (Figure 3.12), and the proportion of persistent syllable types within a song at the song level (Figure 3.13). I suggested relaxed selection of females on male songs and developmental stress to explain temporal change in song performance. In order to test the former hypothesis, the pilot analysis of differences in song complexity and song type consistency of individuals could be expanded to more areas within the range of the house finch. For instance, Hill (2002) found that the plumage of house finches in New York is the most colorful, and the plumage of house finches in Hawaii is the dullest, with birds from some other regions being intermediate. Comparing the songs of these populations could test the predicted negative correlation between plumage and song performance for house finches, providing an important insight into song evolution in the house finch. Though developmental stress and relaxed selection are both being hypothesized here to decrease song performance, they might influence different aspects of song performance. By experimentally testing which acoustic features change under poorer early condition in the lab, the results would help to (1) distinguish its role in explaining the data of this study, and (2) shed light on the underlying neurophysiological process responsible for such change. Features of decreased song performance that are not affected by developmental stress might be a result of relaxed selection by females, a hypothesis that could be tested by comparing female preferences for different song features in different populations.

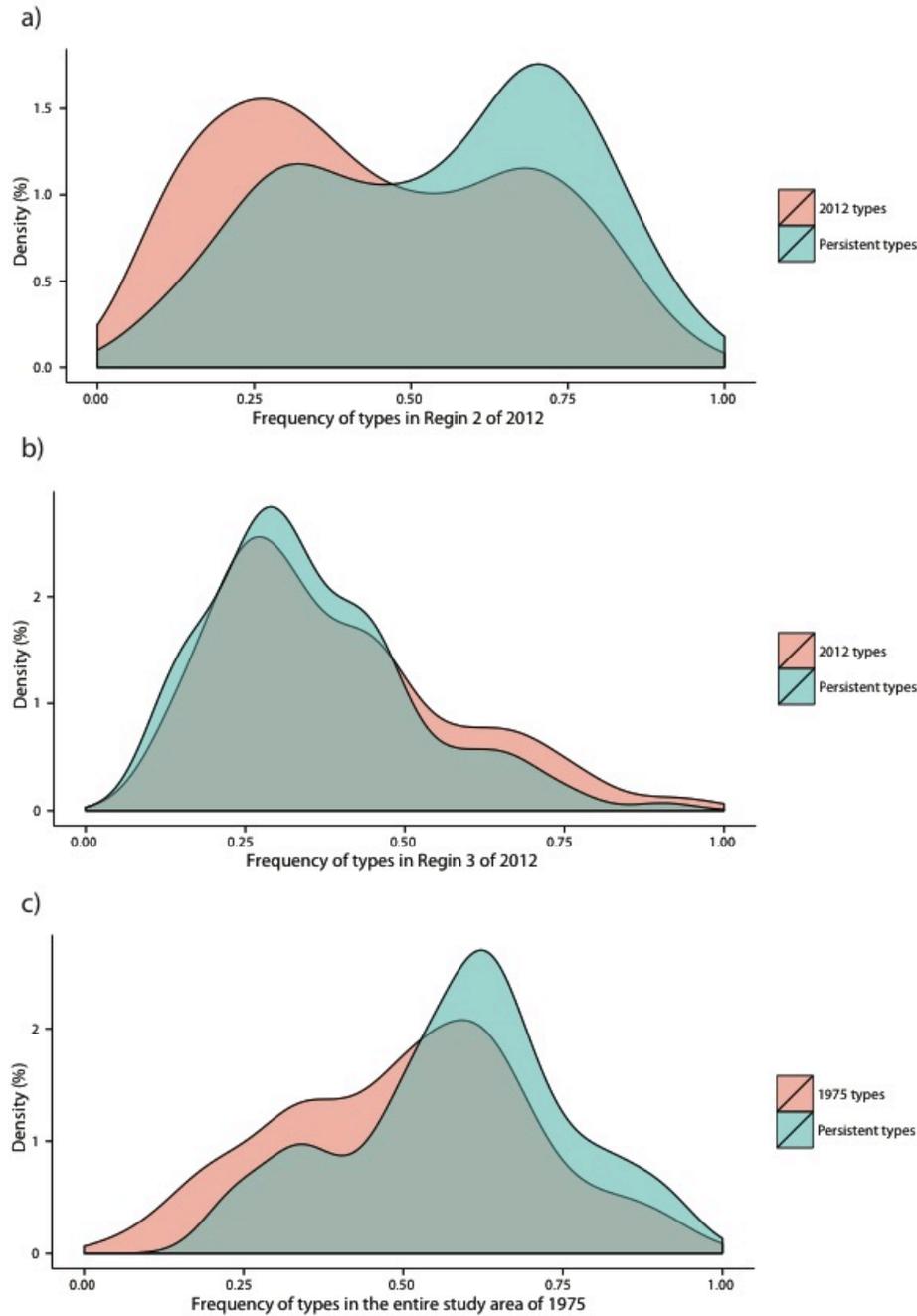


Figure 3.12. Distribution of frequency of syllable types shared by individuals in different regions and different years

a) Frequency of types in region 2 of 2012 (Number of individuals = 37). **b)** Frequency of types in region 3 of 2012 (Number of individuals = 53). **c)** Frequency of types in 1975 (Number of individuals = 94). Persistent syllable types across both sample years still retained high sharing between individuals in region 2, while most types were shared only by a small proportion of birds in region 3.

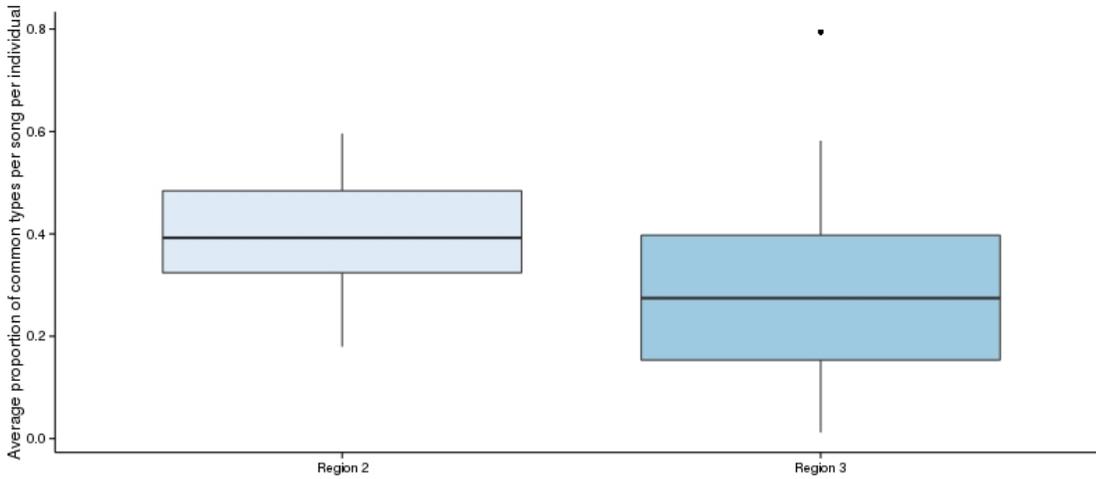


Figure 3.13. Average proportion of common syllable types per song per individual in two different regions of 2012

Significantly higher proportion of persistent syllable types in each song per individual in region 2 is observed compared to that in region 3

Table 3.3. Summary of song comparison between 1975 and 2012

Sample size	1975	2012
Song sharing between individuals (70% similarity)	94 individuals (1042 songs)	96 individuals (981 songs)
Song similarity within individuals	132 pairs	No sharing at all
Temporal lability in song structure	Higher within song similarity in 1975	
Number of syllables	No persistent song type	
Average proportion of syllables	238	287
Syllable sharing against distance	0.51	0.39
Song consistency	Higher syllable sharing in 1975 in every class of distance	
Song complexity (song repertoire)	Higher song consistency in 1975	
Song complexity (syllable repertoire)	3.3	3.1
Acoustic features	80.3	70.6

References:

Able, K. P., & Belthoff, J. R. (1998). Rapid 'evolution' of migratory behaviour in the introduced house finch of eastern North America. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1410), 2063-2071.

Aldrich, J. W., & Weske, J. S. (1978). Origin and evolution of the eastern House Finch population. *The Auk*, 95, 528-536.

Anderson, D. (2008). *Model Based Inference in the Life Sciences: A Primer on Evidence*: Springer.

Anderson, M. J. (2006). Distance-Based Tests for Homogeneity of Multivariate Dispersions. *Biometrics*, 62(1), 245-253.

Badyaev, A. V., & Hill, G. E. (2000). Evolution of sexual dichromatism: contribution of carotenoid-versus melanin-based coloration. *Biological Journal of the Linnean Society*, 69(2), 153-172.

Badyaev, A. V., & Hill, G. E. (2003). Avian sexual dichromatism in relation to phylogeny and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 34, 27-49.

Badyaev, A. V., Hill, G. E., & Weckworth, B. V. (2002). Species divergence in sexually selected traits: increase in song elaboration is related to decrease in plumage ornamentation in finches. *Evolution*, 56(2), 412-419.

Badyaev, A. V., Young, R. L., Oh, K. P., & Addison, C. (2008). Evolution on a local scale: developmental, functional, and genetic bases of divergence in bill form and associated changes in song structure between adjacent habitats. *Evolution*, 62(8), 1951-1964.

Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1547), 1443-1450.

Bitterbaum, E., & Baptista, L. F. (1979). Geographical variation in songs of California House Finches (*Carpodacus mexicanus*). *The Auk*, 462-474.

Byers, B. E. (2014). Migration and song elaboration in wood-warblers (*Geothlypis*). *The Auk*, 132(1), 167-179.

Byers, B. E., Belinsky, K. L., & Bentley, R. A. (2010). Independent Cultural Evolution of Two Song Traditions in the Chestnut-Sided Warbler. *The American Naturalist*, 176(4), 476-489.

Cardoso, G. C., & Atwell, J. W. (2011). Directional cultural change by modification and replacement of memes. *Evolution*, 65(1), 295-300.

Catchpole, C. K., & Rowell, A. (1993). Song sharing and local dialects in a population of the European wren *Troglodytes troglodytes*. *Behaviour*, 125(1), 67-78.

Catchpole, C. K., & Slater, P. (2008). *Bird Song: Biological Themes and Variations*: Cambridge University Press.

- Collins, S. A., de Kort, S. R., Pérez-Tris, J., & Tellería, J. L. (2009). Migration strategy and divergent sexual selection on bird song. *Proceedings of the Royal Society B: Biological Sciences*, 276(1656), 585-590.
- Derryberry, E. P. (2009). Ecology shapes birdsong evolution: variation in morphology and habitat explains variation in white-crowned sparrow song. *The American Naturalist*, 174(1), 24-33.
- Ellers, J., & Slabbekoorn, H. (2003). Song divergence and male dispersal among bird populations: a spatially explicit model testing the role of vocal learning. *Animal Behaviour*, 65(4), 671-681.
- Elliott, J. J., & Arbib Jr, R. S. (1953). Origin and status of the house finch in the eastern United States. *The Auk*, 70, 31-37.
- Fayet, A. L., Tobias, J. A., Hintzen, R. E., & Seddon, N. (2014). Immigration and dispersal are key determinants of cultural diversity in a songbird population. *Behavioral Ecology*, 25(4), 744-753.
- Fee, M. S., & Scharff, C. (2010). The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR Journal*, 51(4), 362-377.
- Hernandez, A. M., & MacDougall-Shackleton, S. A. (2004). Effects of early song experience on song preferences and song control and auditory brain regions in female house finches (*Carpodacus mexicanus*). *Journal of Neurobiology*, 59(2), 247-258.
- Hill, G. E. (1993). Male mate choice and the evolution of female plumage coloration in the house finch. *Evolution*, 47, 1515-1525.
- Hill, G. E. (2002). *A Red Bird in A Brown Bag: the Function and Evolution of Colorful Plumage in the House Finch*: Oxford University Press.
- Horn, A., & Falls, J. B. (1988). Structure of western meadowlark (*Sturnella neglecta*) song repertoires. *Canadian Journal of Zoology*, 66(2), 284-288.
- ILipkind, D., & Tchernichovski, O. (2011). Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. *Proceedings of the National Academy of Sciences*, 108(Supplement 3), 15572-15579.
- Janes, S. W., & Ryker, L. (2013). Rapid change in a Type I song dialect of hermit warblers (*Setophaga occidentalis*). *The Auk*, 130(1), 30-35.
- Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), 1-11.
- Krebs, C. (1989). *Ecological Methodology*. Har-per Collins Publishers.
- Kroodsma, D. E., Liu, W.-C., Goodwin, E., & Bedell, P. A. (1999). The ecology of song improvisation as illustrated by North American Sedge Wrens. *The Auk*, 116, 373-386.
- Lachlan, R., & Feldman, M. (2003). Evolution of cultural communication systems: the coevolution of cultural signals and genes encoding learning preferences. *Journal of Evolutionary Biology*, 16(6), 1084-1095.

- Lachlan, R., & Servedio, M. (2004). Song learning accelerates allopatric speciation. *Evolution*, 58(9), 2049-2063.
- Lachlan, R., & Slater, P. J. (1999). The maintenance of vocal learning by gene–culture interaction: the cultural trap hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1420), 701-706.
- Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., . . . Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, 24(9), 487-496.
- Lahti, D. C., Moseley, D. L., & Podos, J. (2011). A tradeoff between performance and accuracy in bird song learning. *Ethology*, 117(9), 802-811.
- Lang, A. L., & Barlow, J. C. (1997). Cultural evolution in the Eurasian tree sparrow: divergence between introduced and ancestral populations. *Condor*, 99, 413-423.
- Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*, 24(5), 719-720.
- Leadbeater, E., & Chittka, L. (2007). Social learning in insects—from miniature brains to consensus building. *Current Biology*, 17(16), R703-R713.
- Legendre, P., & Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, 80(2), 107-138.
- Lynch, A., & Baker, A. J. (1993). A population memetics approach to cultural evolution in chaffinch song: meme diversity within populations. *American Naturalist*, 141, 597-620.
- Lynch, A., & Baker, A. J. (1994). A population memetics approach to cultural evolution in chaffinch song: differentiation among populations. *Evolution*, 48, 351-359.
- Marler, P., & Tamura, M. (1962). Song "dialects" in three populations of White-crowned Sparrows. *Condor*, 64, 368-377.
- Marler, P., & Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science*, 146(3650), 1483-1486.
- Mcgregor, P. K., & Krebs, J. R. (1982). Song types in a population of great tits (*Parus major*): their distribution, abundance and acquisition by individuals. *Behaviour*, 79(2), 126-152.
- Mundinger, P. (1975). Song dialects and colonization in the house finch, *Carpodacus mexicanus*, on the east coast. *Condor*, 77, 407-422.
- Mundinger, P. (1982). Microgeographic and macrogeographic variation in the acquired vocalizations of birds. *Acoustic Communication in Birds*, 2, 147-208.
- Mundinger, P. C., & Lahti, D. C. (2014). Quantitative integration of genetic factors in the learning and production of canary song. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1781), 20132631.
- Nelson, D., Marler, P., & Morton, M. (1996). Overproduction in song development: an evolutionary correlate with migration. *Animal Behaviour*, 51(5), 1127-1140.

- Nelson, D. A. (1998). Geographic variation in song of Gambel's white-crowned sparrow. *Behaviour*, 135(3), 321-342.
- Nelson, D. A., Hallberg, K. I., & Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology*, 110(11), 879-908.
- Nelson, D. A., & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences*, 91(22), 10498-10501.
- Nolan, P. M., & Hill, G. E. (2004). Female choice for song characteristics in the house finch. *Animal Behaviour*, 67(3), 403-410.
- Nowicki, S., Searcy, W., & Peters, S. (2002). Brain development, song learning and mate choice in birds: a review and experimental test of the "nutritional stress hypothesis". *Journal of Comparative Physiology A*, 188(11-12), 1003-1014.
- Paradis, E., Baillie, S. R., Sutherland, W. J., & Gregory, R. D. (1998). Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, 67(4), 518-536.
- Payne, R. B. (1996). Song traditions in indigo buntings: origin, improvisation, dispersal, and extinction in cultural evolution. *Ecology and Evolution of Acoustic Communication in Birds*, 198-220.
- Payne, R. B., Payne, L. L., & Doehlert, S. M. (1988). Biological and cultural success of song memes in indigo buntings. *Ecology*, 69, 104-117.
- Podos, J. (2001). Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*, 409(6817), 185-188.
- Podos, J. (2007). Discrimination of geographical song variants by Darwin's finches. *Animal Behaviour*, 73(5), 833-844.
- Podos, J., Huber, S. K., & Taft, B. (2004). Bird song: the interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics*, 55-87.
- Podos, J., Peters, S., Rudnicki, T., Marler, P., & Nowicki, S. (1992). The organization of song repertoires in song sparrows: themes and variations. *Ethology*, 90(2), 89-106.
- Podos, J., & Warren, P. S. (2007). The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, 37, 403-458.
- Potvin, D. A. (2013). Larger body size on islands affects silvereye *Zosterops lateralis* song and call frequency. *Journal of Avian Biology*, 44(3), 221-225.
- Pytte, C. L. (1997). Song organization of house finches at the edge of an expanding range. *Condor*, 99, 942-954.
- Rapaport, L. G., & Brown, G. R. (2008). Social influences on foraging behavior in young nonhuman primates: learning what, where, and how to eat. *Evolutionary Anthropology: Issues, News, and Reviews*, 17(4), 189-201.
- Richerson, P. J., & Boyd, R. (2008). *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press.

- Ristad, E. S., & Yianilos, P. N. (1998). Learning string-edit distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(5), 522-532.
- Rivera-Gutierrez, H. F., Matthysen, E., Adriaensen, F., & Slabbekoorn, H. (2010). Repertoire sharing and song similarity between great tit males decline with distance between forest fragments. *Ethology*, 116(10), 951-960.
- Rothstein, S. I., & Fleischer, R. C. (1987). Vocal dialects and their possible relation to honest status signalling in the brown-headed cowbird. *Condor*, 89, 1-23.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.
- Searcy, W. A., & Yasukawa, K. (1996). Song and female choice. *Ecology and Evolution of Acoustic Communication in Birds*, 2.
- Slabbekoorn, H., Jesse, A., & Bell, D. A. (2003). Microgeographic song variation in island populations of the white-crowned sparrow (*Zonotrichia leucophrys nutalli*): innovation through recombination. *Behaviour*, 140(7), 947-963.
- Slater, P., Ince, S., & Colgan, P. (1980). Chaffinch song types: their frequencies in the population and distribution between repertoires of different individuals. *Behaviour*, 75(3), 207-218.
- Sokal, R. R. (1958). A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull*, 38, 1409-1438.
- Spencer, K., Buchanan, K., Goldsmith, A., & Catchpole, C. (2004). Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(Suppl 3), S121-S123.
- Taff, C. C., Steinberger, D., Clark, C., Belinsky, K., Sacks, H., Freeman-Gallant, C. R., . . . Whittingham, L. A. (2012). Multimodal sexual selection in a warbler: plumage and song are related to different fitness components. *Animal Behaviour*, 84(4), 813-821.
- Tchernichovski, O., Lints, T., Deregnaucourt, S., Cimenser, A., & Mitra, P. (2004). Studying the song development process: rationale and methods. *Annals of the New York Academy of Sciences*, 1016(1), 348-363.
- Thompson, W. L. (1960). Agonistic behavior in the House Finch. Part I: Annual cycle and display patterns. *Condor*, 62, 245-271.
- Tracy, T. T., & Baker, M. C. (1999). Geographic variation in syllables of House Finch songs. *The Auk*, 116, 666-676.
- Tubaro, P. L. (1991). Can Troglodytes Aedon in Argentina "Mimic" the Songs of Thryomanes bewickii? *Condor*, 93, 443-445.
- van Schaik, C. P. (2010). Social learning and culture in animals. *Animal behaviour: Evolution and Mechanisms* (pp. 623-653): Springer.
- Wang, Z., Baker, A. J., Hill, G. E., & Edwards, S. V. (2003). Reconciling actual and inferred population histories in the house finch (*Carpodacus mexicanus*) by AFLP analysis. *Evolution*, 57(12), 2852-2864.

Wilkins, M. R., Seddon, N., & Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution*, 28(3), 156-166.

Williams, H., Levin, I. I., Norris, D. R., Newman, A. E., & Wheelwright, N. T. (2013). Three decades of cultural evolution in Savannah sparrow songs. *Animal Behaviour*, 85(1), 213-223.