

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Song syntax and singing behavior of Cassin's Vireo (*Vireo cassinii*)

**Permalink**

<https://escholarship.org/uc/item/9kq5n3ks>

**Author**

Hedley, Richard William

**Publication Date**

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Song syntax and singing behavior of Cassin's Vireo (*Vireo cassinii*)

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Biology

by

Richard William Hedley

2016

© Copyright by

Richard William Hedley

2016

## ABSTRACT OF THE DISSERTATION

Song syntax and singing behavior of Cassin's Vireo (*Vireo cassinii*)

by

Richard William Hedley

Doctor of Philosophy in Biology

University of California, Los Angeles, 2016

Professor Charles E. Taylor, Chair

Songbirds display immense variability in the complexity of their songs, ranging from simple, repeated trills to intricate sequences. Explaining the evolutionary origins of song complexity has proven challenging because the role of song complexity in communication is poorly understood. This may, in part, arise from a disproportionate focus on the role of repertoires in communication, while ignoring other characteristics of song. In this thesis, I study the songs of Cassin's Vireos, taking a broader view of song organization that encompasses both repertoires and the syntax governing the organization of repertoires into sequences. Chapter 1 describes the repertoire sizes and patterns of sequential organization in males at my study site. Repertoires are comprised of about 50 phrase types delivered in sequences that show evidence of temporal structure and repeated patterns. These sequences contain small-world structure, as has been demonstrated in the songs of other bird species. Chapter 2 more formally examines the syntax governing these sequences. I tested the fit of three competing syntactic models of varying

complexity on the song sequences of Cassin's Vireos. Analysis revealed that song sequences were too complex to be described by a zero- or first-order Markov model, meaning that the true complexity is higher, being well described by a model intermediate in complexity between a first- and second-order model. Sequences proved to be predictable with >55% accuracy, and syntax appears to change little, if at all, during adult life. In Chapter 3, I conducted playback experiments to examine the role of repertoires and syntax during counter-singing interactions. Birds presented with playback of a phrase type often responded with the phrase type(s) that would typically follow the stimulus phrase type in their normal sequences, as though they had delivered the stimulus phrase themselves. This implies that both repertoires and syntax play a role in counter-singing exchanges, suggesting that a complete understanding of vocal interactions will require an understanding of syntax in addition to accurate designation of repertoires. This thesis helps elucidate the nature of song complexity in Cassin's Vireo and its role in communication, and provides methodological and theoretical advances that may be applicable to other species and the study of animal communication more generally.

The dissertation of Richard William Hedley is approved.

Abeer A. H. Alwan

Pamela J. Yeh

Charles E. Taylor, Committee Chair

University of California, Los Angeles

2016

TABLE OF CONTENTS

**Abstract of the Dissertation ..... ii**

**Acknowledgments ..... ix**

**Vita ..... xi**

**1 Introduction.....1**

    1.1 Complexity of Bird Song.....1

        1.1.1 Complexity of the Acoustic Structure of Individual Vocal Elements .....2

        1.1.2 Repertoire Size.....8

        1.1.3 Song Syntax .....18

    1.2 Songs of Birds in the Genus *Vireo*.....22

    References.....25

**2 Composition and Sequential Organization of Song Repertoires in Cassin’s Vireo (*Vireo cassinii*).....36**

    2.1 Abstract.....36

    2.2 Introduction.....36

    2.3 Materials and Methods.....39

        2.3.1 Study Site .....39

        2.3.2 Field Methods .....39

        2.3.3 Terminology.....40

        2.3.4 Phrase Identification and Annotation.....41

        2.3.5 Network Analysis.....43

    2.4 Results.....44

        2.4.1 Repertoires and Sharing.....44

2.4.2	Classification Results.....	45
2.4.3	Phrase Type Use .....	46
2.4.4	Sequential Arrangement of Phrases .....	46
2.5	Discussion.....	47
2.5.1	Sequential Arrangement of Phrases .....	47
2.5.2	Repertoire Size.....	49
2.5.3	Song Sharing.....	51
	List of Figures .....	53
	Figures.....	56
	References.....	62
<b>3</b>	<b>Complexity, Predictability, and Time Homogeneity of Syntax in the Songs of Cassin’s</b>	
	<b><i>Vireo (Vireo cassinii)</i>.....</b>	<b>65</b>
3.1	Abstract.....	65
3.2	Introduction.....	66
3.3	Results.....	68
3.3.1	Song Characteristics.....	68
3.3.2	Simulations .....	68
3.3.3	Fisher’s Exact Tests for Higher-order Dependencies .....	70
3.3.4	Model Likelihood.....	71
3.3.5	Predictability of Sequences.....	72
3.3.6	Time Homogeneity .....	74
3.4	Discussion.....	75
3.4.1	Syntactic Complexity.....	75
3.4.2	Predictability of Sequences.....	78



3.4.3	Non-Markovian Properties.....	79
3.5	Materials and Methods.....	82
3.5.1	Recording Techniques and Individual Identification.....	82
3.5.2	Recording Annotation.....	83
3.5.3	Simulations .....	84
3.5.4	Fisher’s Tests for Higher-order Dependencies .....	88
3.5.5	Model Likelihood.....	89
3.5.6	Predictability of Sequences.....	91
3.5.7	Time Homogeneity .....	92
	List of Figures .....	94
	List of Tables .....	97
	Figures and Tables .....	98
	References.....	106
<b>4</b>	<b>Song Mirroring in Birds: Evidence from a Songbird with a Complex Song .....</b>	<b>110</b>
4.1	Abstract.....	110
4.2	Introduction.....	110
4.3	Results.....	113
4.3.1	Literature Review.....	113
4.3.2	Playback Experiments.....	114
4.3.3	Physical Response to Playbacks .....	114
4.3.4	Matching or Advancing Above Chance Levels .....	115
4.3.5	Treatment Effects.....	117
4.3.6	Effects of Distance on Song Mirroring.....	119

4.4	Discussion.....	119
4.4.1	Evidence for the Song Mirroring Hypothesis .....	119
4.4.2	Song Mirroring as a Signal .....	120
4.4.3	A Potential Neural Mechanism.....	122
4.4.4	Future Research .....	123
4.5	Materials and Methods.....	125
4.5.1	Study Site and Species .....	125
4.5.2	Playback Design.....	125
4.5.3	Playback Field Methods.....	126
4.5.4	Motivation for Statistical Analyses.....	127
4.5.5	Statistical Analyses .....	128
	List of Figures .....	131
	List of Tables .....	134
	Figures and Tables .....	135
	Appendix A.....	141
	References.....	150
<b>5</b>	<b>General Discussion.....</b>	<b>154</b>
5.1	Repertoires and Syntax in Cassin’s Vireo Songs.....	154
5.2	The Importance of Syntax of Understanding Bird Songs.....	156
5.3	Functions of Song Syntax .....	159
5.4	Directions of Future Research .....	160
	References.....	163

## ACKNOWLEDGMENTS

First, I am indebted to my advisor Charles Taylor. Throughout my time at UCLA, he has provided me with an ideal combination of support and freedom to pursue my own ideas. I could always rely on him for different perspectives on my findings, writing, and analyses, which always improved the quality of the work and helped me develop as a scientist. I am equally indebted to Martin and Daryl Ann Cody for accommodating me at their house in Amador County for many weeks in the field each summer. Martin's superior knowledge of the ecology and behavior of birds was an invaluable resource throughout the process, and his insights helped me with the design of the studies included here.

Next I am thankful to the other members of my PhD committee, Abeer Alwan, Pamela Yeh and Barney Schlinger. On the occasions that all of us managed to gather in the same room, I always came away with useful insights and advice that has helped me along the way.

I am grateful to Zac Harlow for mentoring me through the first two years of my degree. George Kossan and Carolee Caffrey were great companions in the lab. I thank Kaleda Denton for help with various aspects of Chapter 4, and Rob Weiss for invaluable statistical advice. I am grateful to Jiawei Zhang for his enthusiasm during our collaborations and for hosting me in Beijing during my trip to China, and to Tim Huang for continuing where Jiawei left off in our effort to further the potential of recording systems. I thank Julio Arriaga, Hector Sanchez, Keisuke Daimon, Edgar Vallejo, and Ed Stabler for constructive conversations regarding my research, and Lee Tan for the use of her classification algorithms. Thank you to Ryan Dudragne for his assistance and companionship on a trip to Mexico to collect data for an offshoot of this project.

Other friends and family have helped make my time in LA memorable. Foremost among these is Anna Taylor, who has talked me through countless obstacles, and provided unending support and motivation throughout the past few years. I thank my parents for their support from afar during my degree, and for their decades of support prior to this undertaking.

Funding for this project came from several sources. Much of the work was supported with NSF grant 1125423 and from four small grants from the LA Audubon Society, the Pasadena Audubon Society, the Santa Monica Audubon Society and the Sea and Sage Audubon Society. Additional funding came from the Ecology and Evolutionary Biology department at UCLA and from the La Kretz Center for Conservation Science.

Chapter 2 is a reproduction of a published paper – Hedley, R.W. 2016. Composition and sequential organization of song repertoires in Cassin’s Vireo (*Vireo cassinii*). Journal of Ornithology 157(1): 13-22. Chapter 3 is also a reproduction of a published paper – Hedley, R.W. 2016. Complexity, predictability, and time homogeneity of syntax in the songs of Cassin’s Vireo (*Vireo cassinii*). PLoS One 11:e0150822. Both papers are open access under a CC-BY license, allowing free reproduction without permission. Chapter 4 was co-authored with Kaleda K. Denton (second author) and Robert E. Weiss (third author), and it is reproduced here with their permission.

## VITA

- 2010                    B.S. Biology  
University of Victoria  
Victoria, British Columbia, Canada
- 2010-2012            Ornithology Researcher and Field Crew Leader  
Bird Studies Canada  
Saskatoon, Saskatchewan, Canada
- 2011                    Ornithology Research Assistant  
Tulane University/Smithsonian Museum  
Jamaica
- 2010                    Grizzly Bear Research Assistant  
University of Alberta  
Edmonton, Alberta, Canada
- 2008                    Field Biology Technician  
Keystone Wildlife Research  
Vancouver, British Columbia, Canada
- 2007                    Herpetology Technician  
University of California, Los Angeles  
Bilsa Biological Reserve, Ecuador

## PUBLICATIONS AND PRESENTATIONS

- Hedley, R.W. 2016. Complexity, predictability, and time homogeneity of syntax in the songs of Cassin's Vireo (*Vireo cassinii*). PLoS One 11:e0150822.
- Hedley, R.W. 2016. Composition and sequential organization of song repertoires in Cassin's Vireo (*Vireo cassinii*). Journal of Ornithology 157(1): 13-22.
- Arriaga, J.G., Sanchez, H., Vallejo, E.E., Hedley, R.W., & Taylor, C.E. 2016. Identification of Cassin's Vireo (*Vireo cassinii*) individuals from their songs using an ensemble of learners. Neurocomputing 175: 966-979.
- Drake, K.L., Frey, M., Hogan, D., & Hedley, R.W. 2016. Using digital recordings and sonogram analysis to obtain counts of Yellow Rails. Wildlife Society Bulletin, doi: 10.1002/wsb.658
- Arriaga, J.G., Hedley, R.W., Vallejo, E.E., & Taylor, C.E. Learning Cassin's Vireo (*Vireo cassinii*) syntax through grammatical inference. Accepted to ALIFE XV: The

- International Conference on the Simulation and Synthesis of Living Systems. July 4-8, 2016. Cancun, Mexico.
- Hedley, R.W. 2015. Complexity in the songs of Cassin's Vireo: Markov properties and evidence of higher-order dependencies. American Ornithologists' Union and Cooper Ornithological Society Joint Meeting, July 28-August 2, 2015. Norman, Oklahoma, USA.
- Suzuki, R., Hedley, R.W., & Cody, M.L. Exploring temporal sound-space partitioning in bird communities emerging from inter- and intra-specific variations in behavioral plasticity using a microphone array. American Ornithologists' Union and Cooper Ornithological Society Joint Meeting, July 28-August 2, 2015. Norman, Oklahoma, USA.
- Jongsma, G.F.M., Hedley, R.W., Duraes, R., & J. Karubian. 2014. Amphibian diversity and species composition in relation to habitat type and alteration in the Mache-Chindul Reserve, northwest Ecuador. *Herpetologica* 70(1): 34-46.
- Zhang, J., Kossan, G., Hedley, R.W., Hudson, R.E., Taylor, C.E., Yao, K., & M. Bao. 2014. Fast 3D AML-based bird song estimation. *Unmanned Systems* 2(3): 249-259.
- Zhang, J., Kossan, G., Hedley, R.W., Hudson, R.E., Taylor, C.E., Yao, K., & M. Bao. 2014. Single node 3D DOA estimation of various bird song sources. 21<sup>st</sup> International Congress on Sound and Vibration, July 13-17, 2014. Beijing, China.
- Arriaga, J.G., Sanchez, H., Hedley, R.W., Vallejo, E.E., & C.E. Taylor. 2014. Using song to identify Cassin's Vireo individuals. A comparative study of pattern recognition algorithms. 6<sup>th</sup> Mexican Conference on Pattern Recognition, June 25-28, 2014. Cancún, Mexico.
- Drake, K., Hedley, R.W., Francis, C., Wilson, S., Conkin, J., & Dudragne, R. 2012. Using digital recordings to estimate occupancy and detection of marsh birds in prairie and parkland biomes. 5<sup>th</sup> North American Ornithological Conference, August 14-18, 2012. Vancouver, BC, Canada.

# CHAPTER 1

## Introduction

Many species of birds deliver repertoires of multiple song types in complex sequences that often rank among the most elaborate vocal outputs from non-human animals. Explaining this complexity has proven challenging, and while bird song has been studied intensively for more than fifty years, the evolutionary pressures driving and maintaining song elaboration remain a topic of debate (Byers and Kroodsma 2009). Progress in this regard is likely to come from a combination of in-depth research on individual species investigating the nature and functions of song complexity in a single species, and comparative studies seeking to identify the ecological correlates of song elaboration on longer evolutionary timescales. This thesis takes the former approach by examining the songs of a population of Cassin's Vireos (*Vireo cassinii*), first outlining the singing behaviors of individual birds, then building upon findings at the individual-level to identify the nuances that emerge as these songs are exchanged between simultaneously singing individuals. Chapter 2 describes in detail the songs of this species and the patterns that emerge in long sequences of songs, Chapter 3 compares generative models of syntax to assess the complexity of the intrinsic rules governing these song sequences, and Chapter 4 aims to identify the ways in which these songs are used in male-male interactions during territorial disputes. Bird song provides a unique window into the inner workings of the brains of birds (Bolhuis and Everaert 2013); this thesis aims to peer through this window by examining the structure and organization of the song system of one species of bird, out of hopes the findings will add to current knowledge regarding bird song, specifically, and animal communication more broadly.

### 1.1 Complexity of Bird Song

Bird song in most species takes the form of sequences of vocalizations that are hierarchically comprised of smaller subunits (Berwick et al. 2011). These sequences can vary in complexity along at least three primary axes, leading to great differences between species. These three axes are 1) the complexity of the acoustic structure of individual vocal elements, 2) the number of distinct vocal elements delivered by an individual (i.e. repertoire size), and 3) the complexity of the rules governing the arrangement of these vocal elements into sequences. Each of these three axes has been studied to varying extents, and all appear to show a great deal of variation between species. I will discuss each of these three components of song complexity in turn, to summarize current hypotheses regarding the causes and consequences of variations in vocal complexity in birds.

### **1.1.1 Complexity of the Acoustic Structure of Individual Vocal Elements**

The first component of song complexity – the complexity of individual vocal elements – has been studied in greatest depth in species that lack vocal repertoires, and therefore lack the other two aforementioned aspects of song complexity. Species with just a single type of stereotyped song are estimated to comprise some 25% of the more than 4000 species of songbirds in existence (Macdougall-Shackleton 1997). Despite the uniformity of repertoire sizes within these species, song complexity can still vary, both among individuals within a species and between species. For example, a single song of the White-crowned Sparrow is comprised of a musical set of whistles, buzzes and trills (Marler and Tamura 1962); this multifaceted series is certainly more complex than the song of its relative, the Chipping Sparrow, whose song comprises a single, monotonous trill (Marler and Isaac 1960; Liu and Kroodsma 2006).

Various factors have been demonstrated to affect the acoustic structure of individual songs, and are therefore presumed to drive differences between species over evolutionary



timescales. Females of various species, for instance, pay close attention to the structure of songs. Song characteristics allow females to locate males of their own species and to distinguish them from those of other species (e.g. Benites et al. 2014). This level of discrimination, however, requires attention to only the coarsest species-level characteristics of song; females are known to exceed this minimal level of attentiveness, perceiving and showing differential responses to often remarkably minute differences in song structure.

In many species, females perceive differences between regional dialects, and they have been shown to give more copulation solicitation displays in response to songs from their natal dialect than to songs from foreign dialects (e.g. White-crowned Sparrow, Baker 1983; Yellowhammer, Baker et al. 1987; Red-winged Blackbird, Searcy 1990). This suggests a role for inter-sexual selection on the maintenance of local song characteristics, and a potential role for female choice for familiar songs to act in the enforcement of reproductive isolation during speciation (Slabbekoorn and Smith 2002a). Female Swamp Sparrows show a similar preference, preferring local over foreign dialects (Balaban 1988), preferring more typical local songs over versions with anomalous acoustic structure (Lachlan et al. 2014), and preferring songs that are physically difficult to produce over their less taxing counterparts (Ballentine et al. 2004). Song consistency, too, appears to be important, as males whose songs are less variable from one rendition to the next achieve higher rates of extra-pair paternity in some species (Byers 2007; Cramer et al. 2011; Taff et al. 2012).

The hypothesized reasons for female preferences for familiar songs, and in particular high-quality renditions of familiar songs, are various (Reviewed in Henry et al. 2015). One idea that has received support is that song learning is metabolically costly, allowing song quality in adult life to act as an honest indicator of physical fitness and individual history (Byers et al.

2010). In support of this, Nowicki et al. (2002) found that birds subjected to nutritional stress early in life produced songs that were less similar to the song exemplars on which they were tutored. Furthermore, lower song quality coincided with smaller sizes of various brain structures, suggesting that song characteristics may honestly reflect internal physiological states. Similar results from other species (e.g. Song Sparrow, MacDonald et al. 2006; Zebra Finch, Holveck et al. 2008) support the notion that the occurrence of song learning early in life, when developmental stress is highest, makes the adult song a particularly suitable target for sexual selection during mate choice. Females, as a result, have evolved sensory systems equipped to discriminate subtle differences between the songs of different males, in order to select the best mate from among a suite of potential suitors.

The acoustic structure of individual song renditions also has consequences for male-male interactions. Males, like females, perceive differences between local and foreign dialects, and their aggressive response during playback experiments varies according to the geographic source of the song stimulus (e.g. Balaban 1988; Morton et al. 2006; Pitocchelli 2014). A well-documented component of these discrimination abilities of males is known as the ‘dear-enemy effect’, wherein males learn to identify the particular structure of their neighbors’ songs, and respond less aggressively to these songs than to those of unfamiliar interlopers that are likely to represent a more imminent threat. Species that show this effect include those with repertoires of single song types (Baker et al. 1981), and those with larger repertoires (Brooks and Falls 1975; Stoddard et al. 1992; Molles and Vehrencamp 2001b; Moser-Purdy and Mennill 2016). The prevalence of the dear-enemy effect across songbirds implies a selective pressure acting on song production and enforced through male-male aggression. Males singing consistent songs would be

expected to be more accurately classifiable than those with excessive variation in their songs, with benefits conferred through the establishment of long-term relationships between neighbors.

In addition, males, like females, readily distinguish high-performance song from low-performance song on the basis of characteristics such as trill rate or consistency (de Kort et al. 2009; Lachlan et al. 2014). Song therefore serves to broadcast the characteristics of the singer, and males in the vicinity can weigh the information gleaned from the perceived song against their own physical characteristics to modulate the aggressiveness of their subsequent response (Moseley et al. 2013).

The above evidence suggests that social interactions exert strong selective pressures on the performance of individual song types. On the one hand, males must deliver high-performance renditions of local song types to improve the chances of pairing and copulating with females. These performance characteristics are salient to other males as well, and the threat of increased aggression from neighboring males exerts pressure for performance consistency to facilitate individual recognition from one day to the next. These selective pressures appear to be consistently conservative, such that songs deviating too much from a population norm tend to be disproportionately punished, either through aggression from other males or through reduced success in pairing with females. Longitudinal data from an island population of Savannah Sparrows, however, illustrates that this conservatism in male and female preferences is not universal, and that social pressures can drive turnover in song characteristics on short timescales. Over thirty years of recording, Williams et al. (2013) documented a shortening of trills coupled with shifts to lower frequencies, the elimination of high-frequency song elements altogether, and an increase in the prevalence of click sounds. These shifts were associated with increased reproductive success on the part of the males delivering them, suggesting an adaptive function,

though it was not clear if benefits were accrued through increased success in mate attraction or in male-male conflicts.

An intriguing theoretical parallel may exist between the innovation of songs and the rise and spread of genetic mutations in a population (Lynch 1996). Most genetic mutations are known to be either neutral or deleterious in their fitness effects and are therefore likely to be rapidly removed from a population (Eyre-Walker and Keightley 2007). Rare beneficial mutations can arise that are favored in a population, and can spread rapidly. The results from Williams et al. (2013), in combination with the generally conservative preferences of males and females, support a similar perspective for bird songs: while most song innovations may be deemed unattractive to females or abnormal to males, some variants may confer benefits in interactions with males, females, or both. It is the spread of these variants, which may signal male quality or simply capitalize on pre-existing sensory preferences in females (Owren et al. 2010), that is likely to lead to the gross differences in song structure found between geographically isolated populations and between species (Podos and Warren 2007; Cardoso and Atwell 2011).

In addition to the social factors described above, environmental characteristics also influence the acoustic structure of individual songs. Songs must transmit through environments that are often thick with vegetation and noisy with biotic or abiotic sounds. Distances between a singer and the intended recipient can often be sufficiently large that sound degradation and attenuation affect the transmission of signals (Holland et al. 1998). However, not all sounds transmit equally well in all environments. In general, low frequency sounds are less subject to attenuation and degradation in densely vegetated habitats than are high frequency sounds (Martens 1980). These physical properties of dense forests have led to selection for lower frequency sounds with fewer frequency modulations in forests, and higher frequency songs with

increased frequency modulation in open habitats such as grasslands (Morton 1975). In addition to the absolute frequency of songs, the frequency bandwidth of signals also influences sound transmission; Slabbekoorn et al. (2002) showed that habitat-induced reverberations can benefit transmission of sounds with minimal frequency modulations, suggesting that habitat characteristics may lead birds to alter the complexity of songs to allow for improved song transmission. The acoustic characteristics of songs have been shown to mirror habitat characteristics between populations within a species (Slabbekoorn and Smith 2002b), and a meta-analysis has also shown a significant, albeit weak, effect of habitat structure on various song characteristics across species (Boncoraglio and Saino 2007).

In a few cases, changes in song characteristics due to environmental pressures have been observable on very short timescales, particularly in urban bird populations. Faced with heightened levels of low-frequency traffic noise, White-crowned Sparrows in San Francisco have increased the frequency of their songs by several hundred Hz in the past fifty years (Luther and Derryberry 2012). Similar frequency shifts have been documented in other species as anthropogenic sources of noise have come to dominate the soundscape in many areas (e.g. Great Tit, Slabbekoorn and Peet 2003; Song Sparrow, Wood and Yezerinac 2006; Dark-eyed Junco, Slabbekoorn et al. 2007), illustrating the importance of sound transmission for efficient communication.

Collectively, the above results illustrate the selective pressures shaping the acoustic structure of individual song types. Whether selection is social or environmental, the fundamental driver in all cases appears to be the sensory system of the recipient. Social factors dictate the information content of songs and influence the behavior of receivers, determining the evolutionary outcomes for the sender by affecting their success in competition for mates and

territories. Environmental factors interfere with signal transmission itself, leading to selection for signals that are best suited to the particular environment in which they are found. The few longitudinal studies that have examined changes in bird songs over multiple decades have revealed that bird song characteristics are not static, but undergo constant alterations, with innovation introducing new song variants into a population at the same time that selection acts to remove others. Over time, this process is expected to drive the differences between species that we see today, even when the other aspects of song complexity – repertoire size and syntax – remain unchanged or vary independently.

### **1.1.2 Repertoire Size**

The second axis of song complexity is repertoire size, an aspect of song complexity that has received substantial research attention. Repertoire sizes often vary considerably within species (Peters et al. 2000) and between closely related species (Lemon et al. 2000). Definitions of repertoires, too, vary between species, necessitated by the large differences in singing styles employed by different species. In some, repertoires are comprised of song types, with each song type being an acoustically intricate, yet highly stereotyped string of sounds (e.g. Marsh Wren, Verner 1975; Song Sparrow, Peters et al. 2000). In others, repertoires are comprised of short syllable types, which are strung together into variable songs that are not readily classifiable into song types (e.g. Sedge Warbler, Catchpole 1976; Canary, Markowitz et al. 2013). Though the lack of a clear and unambiguous definition of repertoire size between species can make interpretations of interspecific comparisons difficult (Macdougall-Shackleton 1997), repertoire size continues to be the preferred measure of song complexity due to the large amount of variation evident in the trait, and the ease with this variation can be quantified.

The causes and consequences of variations in repertoire size remain under debate. The primary hypothesized drivers of repertoires can be sorted into two primary groups (Macdougall-Shackleton 1997). First, larger song repertoires may indicate increased overall quality of the singer, thereby being used as a useful characteristic during mate choice. Second, song repertoires may allow for more effective communication by facilitating the use of increasingly flexible signaling strategies.

For song repertoire size to be a useful cue for mate choice, it must accurately and honestly reflect the quality of the singer, and differences in repertoire sizes must be perceptible to the opposite sex. Evidence supporting a relationship between song repertoire size and male quality has come from a large number of unrelated species. In some, individuals with larger repertoires live longer (Hiebert et al. 1989), are more attentive parents (Reid et al. 2005), and perform better in certain cognitive tasks (Boogert et al. 2008; but see Boogert et al. 2011). Furthermore, nutritional deficits early in life constrain the development of song control regions in the brain (MacDonald et al. 2006), leading to subsequent impairment of song repertoires (Spencer et al. 2003; Spencer et al. 2004). This relationship between food provisioning early in life and repertoire development has led to the idea that song repertoires may provide a permanent record of individual history available to prospective mates during reproduction (Nowicki et al. 1998; Pfaff et al. 2007).

Females in many species appear able to discriminate smaller repertoire sizes from larger ones, and results have generally confirmed preferences for the latter. When implanted with estradiol to promote sexual receptivity, females of several species give more copulation solicitation displays in response to playback of larger song repertoires (e.g. Sedge Warbler, Catchpole et al. 1984; Song Sparrow, Searcy 1984; Yellowhammer, Baker et al. 1987; Brown-

headed Cowbird, Hosoi et al. 2005; reviewed in Searcy 1992), and in others are more stimulated to begin nesting in response to large rather than small repertoires (Kroodsma 1976). This preference appears to extend beyond the laboratory in some species, as free-living males with larger song repertoires have been observed to pair earlier (Catchpole 1980; Horn et al. 1993), achieve higher rates of extra-pair copulations (Hasselquist et al. 1996), and, in polygynous species, attract more females into their harem (Yasukawa et al. 1980).

Collectively, the above results appear to suggest direct links between song repertoire size and female preferences in many species. More specifically, the results agree with the “good genes” model of sexual selection (Catchpole 1996), which proposes that females select mates with larger repertoires because these males are likely to possess desirable traits that will allow the female to produce superior offspring. The relationship between song and female choice, however, has not received universal support. In particular, Byers and Kroodsma (2009) criticized the role of female choice in the elaboration of song repertoires, suggesting that the importance of this effect has been overstated. They proposed that the literature suffers from a publication bias, a view that has since been confirmed by the analyses of Soma and Garamszegi (2011), though controlling for this bias did not eliminate the relationship between female choice and repertoire size. Moreover, they point out various aspects of song development and song delivery that contradict the predictions of the female choice model of song elaboration. For starters, many species of birds undergo selective attrition of song types following learning, so their repertoires diminish, rather than expand, prior to adulthood (Marler and Peters 1982). The song types that remain are often delivered in ways that would not be expected if the goal is to impress females with their diversity. Many, if not most, songbird species deliver their songs with eventual variety, meaning each song type is repeated many times prior to switching to another song type (Price



2013). Further, some song types are common while others are only extremely rarely delivered (Wunderle Jr. et al. 1992). These characteristics of song are difficult to reconcile with the claim that a primary function of song repertoires is to impress females with vocal versatility. If size were the most important feature of repertoires, the expectation would be that males would sing largely improvised song repertoires, and to deliver all song types approximately equally, and in a cyclical fashion without repetitions; these predictions are not met in most species.

As an alternative to the female choice paradigm of repertoire evolution where song is presumed to act as an advertisement signal, Byers and Kroodsma (2009) advocated for an increased focus on song as a communication system. In their view, repertoires may have arisen to allow for more flexible exchange of information in ways that would not be possible with simpler repertoires. There are two primary ways that song repertoires could accomplish communication exchange: through time-invariant mappings between song types and information, and through transient relationships between song and context.

Time-invariant relationships between song and information are those in which birds consistently deliver particular song types in particular contexts, such that a naïve listener may acquire information about a signaler's behavior or motivational state by hearing their songs. Examples of this are common in bird calls, which can encode information about predator type (Suzuki 2014), the urgency of threats (Baker and Becker 2002), or the presence of food (Evans and Evans 2007), but examples of this type of semantic content in songs are less well understood. Smith et al. (1978) found that Yellow-throated Vireos often deliver different song types according to their proximity to their nest. Trillo and Vehrencamp (2005) documented a similar pattern in Banded Wrens, which altered their use of different song types during counter-singing, when near a female, and at different times of day; similar results have been presented for

Song Sparrows (Kramer and Lemon 1983). Wood-warblers in the family Parulidae commonly partition their repertoires into two subsets of song types that appear to be delivered in different social contexts (Reviewed in Spector 1992). Some wood-warbler species changing their song patterns in accordance with their mating status (Morse 1966; Kroodsma et al. 1989; Demko et al. 2013), while others vary their choice of song type according to their location on their territory (Lein 1978), as a function of their current behavior (i.e. whether foraging or actively defending their territory, Morse 1967) or depending on the sex of the intended receiver (Morse 1966; Morse 1967; Morrison and Hardy 1983; Demko et al. 2013). In these species, the two repertoire subsets may themselves contain one to several song types which differ consistently in either their acoustic structure (Kroodsma et al. 1989) or their mode of delivery (Demko et al. 2013). These differences presumably allow the two song categories to send different messages, even when the listener has limited or no prior familiarity with the singer.

Despite the breadth of evidence for consistent relationships between contextual variables and the use of particular song types, this aspect of song is often considered to be a peculiarity of the songs of a few species, rather than a widespread feature of the songs of many species (Macdougall-Shackleton 1997). This conclusion may result from the tendency for correlations between context and vocal output to be imperfect, meaning semantic content is probabilistic, rather than exact, and inferences based on the semantic content of song must be similarly approximate. Models of communication, however, can account for this uncertainty: rather than relying on precise, word-like meaning, signals are instead evaluated based on the extent to which they increase a receiver's ability to make correct decisions in the face of uncertainty about their surroundings (Bradbury and Vehrencamp 2000). More accurate signals are likely to be more valuable to receivers, and are expected to warrant attention if the information they provide is

sufficiently valuable that it outweighs the cost associated with processing and deciphering the signal. Signal accuracy, however, is constrained by discordance between the interests of the sender and receiver, such that it is not always favorable for signalers to freely deliver information to conspecifics who may be competitors for food, mates, or other resources (Krebs and Dawkins 1984). The resulting signals are often intermediate with respect to the accuracy with which they correspond to external events, but this does not mean that semantic content is altogether lacking. Further field studies are surely warranted to evaluate the extent to which songs serve to convey this type of information between senders and receivers, and to clarify the nature of the information that is conveyed.

The second way that song repertoires can be used to convey information is through context-dependent use of song. In this case, a song type heard in isolation does not reveal information about a singer's behavior or internal state – the meaning of the signal is inextricably linked to the social context in which the song was delivered. One example of this mode of signaling is song type matching, where a bird repeats the song type of a rival in rapid succession. Hinde (1958) was the first to describe this behavior. When he broadcast a particular song type to a captive Chaffinch, he noted that the bird would often reply with the same song type. Song matching has since been described in a variety of distantly-related species (e.g. Carolina Wren, Simpson 1985; Song Sparrow, Stoddard et al. 1992; Rufous Bristlebird, Rogers 2004), suggesting that it is an ancient and evolutionarily conserved behavior that may be fundamental to the signaling behavior of many species of songbirds. Variants on this behavior are also known. Song sparrows, for example, engage in repertoire matching, where instead of responding with the same song type as their rival, they respond with another song type in their rival's repertoire (Beecher et al. 1996; Beecher et al. 2000). This species also engages in partial song matching,

where a bird replies with a song type that matches certain aspects of the stimulus, but is not of the same type (Burt et al. 2002). In contrast to the time-invariant semantic content of discussed previously, the information conveyed in song matching interactions is specific to the context in which it occurs. It is not the song type, per se, that conveys information, but the relationship of that song type to the individual identity of the receiver (repertoire matching) or to the most recent song type delivered by the rival (song type matching).

Considerable efforts have been made to understand the function of song type matching. Hypotheses include the idea that song matching may allow females to better compare the singing behavior of two males, by allowing an eavesdropping female to compare two identical, rather than two different, songs during counter-singing (Logue and Forstmeier 2008). This idea originates from the theory put forth by Zahavi (1980) that ritualized signals (in this case, shared song types) provide more useful information to an observer than non-ritualized signals. By engaging in matched counter-singing, differences in performance of a shared song type between the two participants become maximally detectable, providing a benefit to the superior male.

This hypothesis has yet to be rigorously tested, but the plausibility of the hypothesis is supported by several studies demonstrating that eavesdropping is a common, if little understood, behavior in birds. Bartsch et al. (2014) used speakers to simulate matched counter-singing between two rivals, and found that both males and females responded asymmetrically to the two speakers, attending more closely to the “leader” in the interaction (that is, the speaker being matched). At a minimum, this result shows that song matching interactions are not ignored by nearby birds, but it does not resolve the precise consequences of matched counter-singing for either the bird engaging in matching or the bird being matched.

Though not involving song matching specifically, Otter et al. (1999) conducted playback to overlap the songs of some Great Tit males, signifying a poor performance by that male in a counter-singing duel, and independently conducted alternating playback with a neighbor, signifying a well-performed singing interaction. They found that females mated to the poor performing males preferentially intruded into the neighboring territories, presumably because they had learned, through eavesdropping on the counter-singing interaction, that their current mate was lower quality than the neighbor into whose territory they ventured. Female Black-capped Chickadees showed a similar pattern: females made to believe their high-ranking mate had fared poorly in a counter-singing exchange were more likely to seek extra-pair copulations (Mennill 2002), showing a direct consequence of counter-singing performance on reproductive output. Despite evidence for eavesdropping in nature, it is clear that a more complete understanding of its consequences will be required to evaluate the hypothesis that song matching is directed at eavesdroppers.

A hypothesis regarding song matching that has received more attention and empirical support is the notion that song matching serves to convey aggressive intentions from sender to receiver. Support for this idea came from the observation that matching correlates with various aspects of the strength of the physical response, implying that a singing bird whose song is matched may be able to infer that an aggressive encounter will soon follow. Increased levels of matching are related to decreased latency to approach a playback speaker and the tendency to closely approach the speaker (Krebs et al. 1981), as well as increased length of songs given in response to the playback (Vehrencamp et al. 2007), all of which are presumed to indicate heightened levels of aggression. Furthermore, birds are more likely to engage in song matching early in the breeding season, when competition for mates and territories is fiercest and aggressive

signaling is expected to be most frequent (Beecher et al. 2000). The aggressive content of matched singing is perceptible to receivers, leading birds whose songs are interactively matched to respond more aggressively than those whose songs are not matched (Molles and Vehrencamp 2001a; Burt et al. 2001).

An early theoretical problem with the aggressive signaling view of song matching is the problem of signal honesty. Signaling theory requires mechanisms by which the honesty of signals can be maintained (Krebs and Dawkins 1984; Bradbury and Vehrencamp 2011). That is, the production of signals must be costly, or cheaters would gain a fitness advantage by always signaling the highest level of aggression at all times, thereby eliminating the information content of the signal. In gazelles, for instance, the honesty of stotting displays is maintained by the physical demands required for performance of the display, meaning that weak or injured individuals will be unable to perform the display or will perform it poorly (FitzGibbon and Fanshawe 1988). Song matching, however, does not have obvious physiological costs that are absent from non-matching song, so the factors maintaining signal honesty were initially unclear. Addressing this issue, Vehrencamp (2001) proposed that song matching may be a conventional signal, defined as a signal whose honesty is maintained through the threat of retaliation: individuals that falsely advertise their willingness to fight through song matching suffer repercussions in the form of increased aggression from neighbors.

Evidence that song matching is a conventional signal has been mixed. In support of this notion, Vehrencamp (2001) showed that in song sparrows, individuals that matched a playback behaved more aggressively than non-matchers, and those whose songs were matched also responded more aggressively. She reasoned that individuals engaging in song matching must therefore be willing to escalate an aggressive encounter, or will incur a retaliation cost. Searcy et

al. (2006) criticized these and other experiments for their lack of realistic design; they argued that the lack of a taxidermic mount accompanying the playback speaker in these experiments might be confusing for the subject and makes inference on aggressive behaviors ambiguous. Their experiments, which included a taxidermic model of a song sparrow placed near the speaker, found that song matching did not predict physical attack. These and other negative results led Searcy et al. (2009) to argue that song matching is not a conventional signal, and may not convey aggression at all. More recently, Akcay et al. (2013) proposed that the negative results could themselves be attributed to playback design issues: if the stimulus is placed too far inside a territory boundary or otherwise represents an unnaturally strong threat, the regular series of signals may be skipped altogether in favor of immediate attack. Their experiment instead simulated a progressively escalating encounter, moving from outside a territory across the boundary and into the territory center. In contrast to the previous results, song matching reliably predicted subsequent attack on a taxidermic model, supporting the conventional signaling hypothesis.

Though the precise role of song matching in counter-singing dynamics remains a topic of active research, it is clear that matching is a common behavior in various bird species, and that it is a behavior that directly relies upon the existence of repertoires of song types that are shared between neighbors. In addition to song matching, repertoires facilitate a number of other, potentially salient, signaling behaviors. For example, the act of switching from one song type to another sends a different signal than remaining on the same song type in some species. In some species, switching appears to relate to increased levels of escalation during singing contests (Kramer et al. 1985; Peake et al. 2005), while in others, switching rates are reduced during intense interactions (Molles 2006).

Collectively, the evidence supports the perspective of Byers and Kroodsma (2009) that song repertoires are more than simple indices of male quality. If song repertoires were “the acoustic equivalent of a peacock’s tail”, as has been claimed (Catchpole 1987), we would expect males to cycle through their repertoires in an effort to show off their size, and to deliver songs that are not dependent on the songs of other males or on current social contexts. The results reviewed above reveal a more nuanced view: that song repertoires provide plasticity of signaling, allowing birds to send and receive signals in ways that would be impossible with a more limited suite of signals. However, even in well-studied species such as Song Sparrows, debate persists surrounding even the most fundamental aspects of the rules governing counter-singing interactions and their consequences for the participants. Resolution of these debates will likely come from more sophisticated experimental designs and analyses, and from studies on a larger array of species.

### **1.1.3 Song Syntax**

The third axis of song complexity is song syntax. Syntax, in this thesis, refers to the unvoiced rules that govern the arrangement of vocalizations (Bolhuis et al. 2010). Syntax ties together the other two aspects of song complexity by arranging the elements of song repertoires into sequential outputs. Despite its seemingly important role in structuring the song outputs of birds, relatively little is known about syntax in comparison to the other aspects of song complexity discussed above.

The term syntax is borrowed from the study of human language, but differences exist between the syntax of human language and the syntax of animal vocalizations. In human languages, information is conveyed in two primary ways: first in the combination of phonemes to produce words (i.e. phonology), and second through the arrangement of words into sentences



(i.e. grammars, including syntax). Birds appear to possess a type of phonology, often composing salient units such as song types from smaller units with apparently arbitrary structure (Podos et al. 1992; Marler 1998; Samuels 2015). The discussion of song repertoires in the previous section illustrates that the resulting song types can convey information in meaningful ways. However, evidence that vocalizations can be combined to generate new meaning, as words are combined to form sentences in human language, has only rarely been reported in non-human animals.

A few examples of compositional syntax have been reported from non-human primates. Campbell's monkeys combine a limited set of 'boom', 'krak', and 'hok' calls into sequences associated with particular contexts with apparently high specificity, communicating the presence of two different predators, and having a role in the coordination of group movements (Ouattara et al. 2009). Putty-nosed monkeys combine 'pyow' and 'hack' calls to variably signal the presence of eagles or leopards, or to initiate group movements (Arnold and Zuberbühler 2006; Arnold and Zuberbühler 2008). In birds, however, compositional syntax has been clearly demonstrated in only one species, the Japanese Great Tit (Suzuki et al. 2016). In this species, birds deliver two different call sequences, sequence 'ABC' and 'D', which cause recipients to scan for danger and approach the caller, respectively. When combined into a sequence ABCD, birds respond by both scanning and approaching the caller, but only rarely give either response when the order of calls is reversed to DABC.

The compositional syntax of Japanese Great Tits does not involve songs, but calls, which differ from song both in their simpler acoustic structure and in their tendency to be more obviously related to external events (Marler 2004). Furthermore, the example above involves combination of just two elements, resulting in the formation of very simple sequences. In stark contrast, the syntax of bird song can involve arrangement of dozens, or even hundreds, of vocal

units into remarkably intricate sequences. The apparent complexity of these sequences raises two primary questions: first, what is the nature of song syntax, and second, what is its function?

The study of song syntax dates back to the 1970's when Chatfield and Lemon (1970) published a tutorial on the use of Markov models to study sequences of song. Markov models are probabilistic models that comprise a set of states, typically corresponding to the song types in a bird's repertoire, and a set of transition probabilities describing the probability of transitioning from any given song type to any other. The simplest application of Markov models is to determine whether transitions between song types occur randomly, or whether certain transitions are preferred over others. In every species that studied so far, transitions occur non-randomly (e.g. Northern Cardinal, Lemon and Chatfield 1971; Rose-breasted Grosbeak, Lemon and Chatfield 1973; Eurasian Blackbird, Morgan 1976; Willow Warbler, Gil and Slater 2000). Indeed, given that song sequences are outputs of bird brains, it would be truly surprising if the rules governing these sequences were completely random.

More recently, attempts have been made to more accurately characterize the syntax of various species. This research has been particularly timely because of the suggestion by various authors that syntactic complexity is one of the defining features distinguishing human language from the communication systems of non-human animals (Hauser et al. 2002; Berwick et al. 2011; ten Cate and Okanoya 2012; Berwick et al. 2012). In particular, these authors have proposed that bird song sequences, and the vocalizations of non-human animals more broadly, can be modeled using finite-state automata, similar to the Markov models discussed above. Human languages, in contrast, are presumed to require more complex computational models for their description, implying more sophisticated neural mechanisms underpinning their generation and comprehension (Chomsky 1956). Empirical evidence for this distinction between human and

non-human syntactic abilities, however, has derived from just a few species of birds. Some of these have demonstrated that Markov models are not sufficiently complex to describe sequences of bird songs, though Markov models with relatively simple modifications have generally performed well (Jin and Kozhevnikov 2011; Kershenbaum et al. 2014). Others have identified long-range statistical dependencies between syllables occurring several seconds apart in song sequences, and while these dependencies can be modeled with high-order Markov models, they imply an unexpected level of sophistication in the sequencing rules employed by birds (Markowitz et al. 2013). The diversity of results obtained from just a small sample of songbird species demonstrate the need for additional research before conclusive statements can be made regarding the syntactic abilities of birds in general.

Research to understand the functional aspects of bird song syntax for birds has been similarly scarce. The few studies that have experimentally manipulated syntax have shown that altering the ordering of temporally structured song sequences can alter the responses of listeners. Skylarks, for example, deliver their elaborate songs in sequences that contain short sequences that can be shared between neighboring birds (Briefer et al. 2013). Permuting the order of these shared sequences leads to more aggressive responses from neighboring males. The authors proposed that the syntax of songs, in addition to their acoustic structure, plays a role in the dear-enemy effect described above, where males living in close proximity establish long-term relationships characterized by reduced levels of aggression. These results have been supported by other studies that have shown a key role for song syntax in the dear-enemy effect (Balaban 1988) and individual identification more generally (Gentner 2008), as well as in the selection of mates by females (Balaban 1988; Nowicki et al. 2001). In contrast to the above results, Holland

et al. (2000) found that broadcasts of the songs with disrupted syntax diminished the responses of male Eurasian Wrens.

The above results illustrate that, at a minimum, syntax encodes information in some species. Whether this information is restricted to individual or species identity, or whether syntax conveys additional information or plays some other role in communication, has received little attention. Furthermore, there is a lack of information regarding the costs associated with complex syntax. It seems likely that the evolution of additional syntactic complexity, as with increased repertoire sizes discussed in the previous section, requires the diversion of nutritional resources towards the development of the song system in the brain. A neural model for syntax generation proposed by Katahira et al. (2007) explicitly implies such a cost, but laboratory research has yet to confirm the precise neural mechanisms underpinning variable sequencing rules in bird songs or the costs associated with their development (Jin 2013). A coherent picture of song syntax therefore remains elusive, with progress dependent upon additional studies from a greater diversity of species that employ complex and variable syntax in their songs.

## **1.2 Songs of Birds in the Genus *Vireo***

The Vireo genus comprises approximately 30 species of small songbirds distributed widely within the new world (Slager et al. 2014). Within this group, a wide variety of singing behaviors are known. Considered in light of the three axes of song complexity discussed above, the songs of Hutton's Vireo are likely to rank among the simplest of any songbird. They possess a small repertoire of approximately three phrase types per individual (Mountjoy and Leger 2001), and each phrase type is an acoustically simple buzz. Furthermore, their syntax does not appear complex, as their sequences of song are characterized by monotonous repetitions of the same phrase type, often for minutes at a time, prior to switching to another phrase type (Davis

1995). Warbling Vireos, in contrast, rearrange a set of short syllable types to produce an apparently limitless variety of song outputs. The research conducted so far on this species has been unable to enumerate the upper limit to either the syllable repertoire or the song type repertoire of this species, or to determine if one exists (Howes-Jones 1985). Though larger sample sizes and more detailed analyses of the songs of both of these species would be desirable, it is nevertheless clear that the complexity of the songs of Warbling Vireos far exceeds that of Hutton's Vireos.

In general, the factors driving divergence of song within this genus are poorly characterized. Evolutionary history has an apparently strong effect on the structure of songs, evidenced by the similar singing styles employed by close relatives. For example, the focal species of this thesis, Cassin's Vireo, sings songs that are only distinguishable from its closest relatives, the Plumbeous and Blue-headed Vireos, on the basis of subtle characteristics such as the extent of frequency modulations in the phrases (James 1981). Another clade, comprising the Yucatan Vireo, Yellow-green Vireo, Black-whiskered Vireo, and Red-eyed Vireo, all deliver songs comprised of short two to three syllable phrases, separated by intervals of silence (Borrer 1981; Chase et al. 2002).

One set of closely-related vireos has eschewed this pattern of song similarity. The Philadelphia Vireo, a migratory species that breeds in the boreal forests of North America, has a song very different from its closest relatives, the Warbling Vireo and the similar-sounding Brown-capped Vireo. In stark contrast to the other two, the Philadelphia Vireo possesses a song that, to human ears, very closely resembles the songs of the Red-eyed Vireo (Moskoff and Robinson 2011), being composed of short phrases separated by a second or two of silence. These two species often co-occur and compete for the same resources, engaging in regular physical

confrontations (Robinson 1981). Playback of song revealed an asymmetry in the discrimination abilities of the two species: Philadelphia Vireos responded less frequently to Red-eyed Vireo song than to conspecific song, and their physical responses to the two stimuli differed in various ways (Rice 1978). Red-eyed Vireos, in contrast, responded indistinguishably to the two song playbacks. The conclusion drawn was that Philadelphia Vireos, the smaller of the two species, benefit by deceiving the larger Red-eyed Vireos, allowing the former species to exclude the latter from territories more effectively. It seems likely that this deception has been facilitated by rapid evolutionary change in the song characteristics of Philadelphia Vireos to more closely resemble the Red-eyed Vireo songs. This type of interspecific deception does not appear common in published examples of bird song, but represents yet another factor that may shape the acoustic structure of bird songs.

Song development in the Vireos, as in other songbirds, appears to occur through learning. Evidence for this has not been conclusively demonstrated, but support has come from a variety of sources. Bradley (1980) reported that the repertoires of hatchling White-eyed Vireos are derived from a combination of their father's repertoire and the repertoire of neighboring males. In the same species, Adkisson and Conner (1978) showed evidence of inter-specific mimicry, suggesting that songs are sometimes influenced by environmental, as well as conspecific, sounds. Examples are known of rare and apparently erroneous inter-specific learning from other vireo species, including a Red-eyed Vireo that seemed to have learned the songs of an Olive-sided Flycatcher, a Warbling Vireo whose song resembled a Red-eyed Vireo (James 1976), and a Yellow-throated Vireo that shared phrase types with nearby Blue-headed Vireos (James 1984). These lines of evidence are indirect, but point towards a role for learning in the development of song repertoires.

Vireos, therefore, appear similar to many other songbird species in that their songs appear to develop early in life through social learning. The factors outlined previously that shape the structure and complexity of songs in other species may apply similarly to vireos, but much remains unknown. In particular, the use of repertoires during counter-singing interactions and the syntax of song in this group has not been studied in depth. This thesis will focus on the songs of Cassin's Vireo (*Vireo cassinii*), a species that is widespread and common in the montane forests of western North America. The primary goal of the thesis will be to characterize the complexity of song in this species, investigating both the repertoire composition and syntax, then to determine how this complexity is employed by the birds. By examining the use of repertoires and the patterns of syntax, I aim to work towards an understanding of the evolutionary pressures that drive the evolution of song complexity in birds, and in animals more broadly.

### References

- Adkisson C, Conner R (1978) Interspecific vocal imitation in White-eyed Vireos. *Auk* 95:602–606.
- Akçay Ç, Tom ME, Campbell SE, Beecher MD (2013) Song type matching is an honest early threat signal in a hierarchical animal communication system. *Proc R Soc B* 280:20122517. doi: 10.1098/rspb.2012.2517
- Arnold K, Zuberbühler K (2006) Language evolution: semantic combinations in primate calls. *Nature* 441:303. doi: 10.1038/441303a
- Arnold K, Zuberbühler K (2008) Meaningful call combinations in a non-human primate. *Curr Biol* 18:R202–203. doi: 10.1016/j.cub.2008.01.040
- Baker MC (1983) The behavioral response of female Nuttall's White-crowned Sparrows to male song of natal and alien dialects. *Behav Ecol Sociobiol* 12:309–315. doi: 10.1007/BF00302898
- Baker MC, Becker AM (2002) Mobbing calls of Black-capped Chickadees: effects of urgency on call production. *Wilson Bull* 114:510–516. doi: 10.1676/0043-5643(2002)114[0510:MCOBCC]2.0.CO;2
- Baker MC, Bjerke TK, Lampe HU, Espmark YO (1987) Sexual response of female Yellowhammers to differences in regional song dialects and repertoire sizes. *Anim Behav*

35:395–401. doi: 10.1016/S0003-3472(87)80263-4

- Baker MC, Thompson DB, Sherman GL (1981) Neighbor/stranger song discrimination in White-crowned Sparrows. *Condor* 83:265–267.
- Balaban E (1988) Bird song syntax: learned intraspecific variation is meaningful. *Proc Natl Acad Sci U S A* 85:3657–60.
- Ballentine B, Hyman J, Nowicki S (2004) Vocal performance influences female response to male bird song: An experimental test. *Behav Ecol* 15:163–168. doi: 10.1093/beheco/arg090
- Bartsch C, Wenchel R, Kaiser A, Kipper S (2014) Singing onstage: female and male Common Nightingales eavesdrop on song type matching. *Behav Ecol Sociobiol*. doi: 10.1007/s00265-014-1727-6
- Beecher MD, Campbell SE, Burt JM, et al. (2000) Song-type matching between neighbouring Song Sparrows. *Anim Behav* 59:21–27. doi: 10.1006/anbe.1999.1276
- Beecher MD, Stoddard PK, Campbell SE, Horning CL (1996) Repertoire matching between neighbouring Song Sparrows. *Anim Behav* 51:917–923. doi: 10.1006/anbe.1996.0095
- Benites P, Campagna L, Tubaro PL (2014) Song-based species discrimination in a rapid Neotropical radiation of grassland seedeaters. *J Avian Biol* 45:001-008. doi: 10.1111/jav.00447
- Berwick RC, Beckers GJL, Okanoya K, Bolhuis JJ (2012) A bird's eye view of human language evolution. *Front Evol Neurosci* 4:5. doi: 10.3389/fnevo.2012.00005
- Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ (2011) Songs to syntax: the linguistics of birdsong. *Trends Cogn Sci* 15:113–121. doi: 10.1016/j.tics.2011.01.002
- Bolhuis JJ, Everaert M (2013) *Birdsong, speech, and language: exploring the evolution of mind and brain*. MIT Press, Cambridge, MA
- Bolhuis JJ, Okanoya K, Scharff C (2010) Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci* 11:747–759. doi: 10.1038/nrn2931
- Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. *Funct Ecol* 21:134–142. doi: 10.1111/j.1365-2435.2006.01207.x
- Boogert NJ, Anderson RC, Peters S, et al. (2011) Song repertoire size in male Song Sparrows correlates with detour reaching, but not with other cognitive measures. *Anim Behav* 81:1209–1216. doi: 10.1016/j.anbehav.2011.03.004
- Boogert NJ, Giraldeau LA, Lefebvre L (2008) Song complexity correlates with learning ability in Zebra Finch males. *Anim Behav* 76:1735–1741. doi: 10.1016/j.anbehav.2008.08.009



- Borror DJ (1981) The songs and singing behavior of the Red-eyed Vireo. *Condor* 83:217–228.
- Bradbury JW, Vehrencamp SL (2000) Economic models of animal communication. *Anim Behav* 59:259–268. doi: 10.1006/anbe.1999.1330\ranbe.1999.1330 [pii]
- Bradbury JW, Vehrencamp SL (2011) *Principles of Animal Communication*, 2nd ed. Sunderland, MA
- Bradley RA (1980) Vocal and territorial behavior in the White-eyed Vireo. *Wilson Bull* 92:302–311.
- Briefer EF, Rybak F, Aubin T (2013) Does true syntax or simple auditory object support the role of skylark song dialect? *Anim Behav* 86:1131–1137. doi: 10.1016/j.anbehav.2013.09.019
- Brooks RJ, Falls JB (1975) Individual recognition by song in White-throated Sparrows. I. Discrimination of songs of neighbors and strangers. *Can J Zool* 53:879–888.
- Burt JM, Bard SC, Campbell SE, Beecher MD (2002) Alternative forms of song matching in Song Sparrows. *Anim Behav* 63:1143–1151. doi: 10.1006/anbe.2002.3011
- Burt JM, Campbell SE, Beecher MD (2001) Song type matching as threat: a test using interactive playback. *Anim Behav* 62:1163–1170. doi: 10.1006/anbe.2001.1847
- Byers BE (2007) Extrapair paternity in Chestnut-sided Warblers is correlated with consistent vocal performance. *Behav Ecol* 18:130–136. doi: 10.1093/beheco/arl058
- Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. *Anim Behav* 77:13–22. doi: 10.1016/j.anbehav.2008.10.003
- Byers J, Hebets E, Podos J (2010) Female mate choice based upon male motor performance. *Anim Behav* 79:771–778. doi: 10.1016/j.anbehav.2010.01.009
- Cardoso GC, Atwell JW (2011) Directional cultural change by modification and replacement of memes. *Evolution* 65:295–300. doi: 10.1111/j.1558-5646.2010.01102.x
- Catchpole C (1976) Temporal and sequential organisation of song in the sedge warbler (*Acrocephalus schoenobaenus*). *Behaviour* 59:226–246.
- Catchpole C (1980) Sexual selection and the evolution of complex songs among European warblers of the genus *Acrocephalus*. *Behaviour* 74:149–166.
- Catchpole C, Dittami J, Leisler B (1984) Differential responses to male song repertoires in female songbirds implanted with oestradiol. *Nature* 312:563–564.
- Catchpole CK (1987) Bird song, sexual selection and female choice. *Trends Ecol Evol* 2:94–7. doi: 10.1016/0169-5347(87)90165-0
- Catchpole CK (1996) Song and female choice: good genes and big brains? *Ecol Evol Acoust Commun Birds* 11:358–360.

- Chase JF, Woodworth BL, Cruz A (2002) Black-whiskered Vireo (*Vireo altiloquus*). Birds North Am. Online
- Chatfield C, Lemon RE (1970) Analysing sequences of behavioural events. J Theor Biol 29:427–445. doi: 10.1016/0022-5193(70)90107-4
- Chomsky N (1956) Three models for the description of language. IRE Trans Inf Theory 2:113–124. doi: 10.1109/TIT.1956.1056813
- Cramer ERA, Hall ML, de Kort SR, et al. (2011) Infrequent extra-pair paternity in the Banded Wren, a synchronously breeding tropical passerine. Condor 113:637–645. doi: 10.1525/cond.2011.100233
- Davis JN (1995) Hutton’s Vireo (*Vireo Huttoni*). Birds North Am Online. doi: doi:10.2173/bna.189
- de Kort SR, Eldermire ERB, Valderrama S, et al. (2009) Trill consistency is an age-related assessment signal in Banded Wrens. Proc Biol Sci 276:2315–21. doi: 10.1098/rspb.2009.0127
- Demko AD, Reitsma LR, Staicer CA (2013) Two song categories in the Canada Warbler (*Cardellina canadensis*). Auk 130:609–616. doi: 10.1525/auk.2013.13059
- Evans CS, Evans L (2007) Representational signalling in birds. Biol Lett 3:8–11. doi: 10.1098/rsbl.2006.0561
- Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nat Rev Genet 8:610–618. doi: 10.1038/nrg2146
- FitzGibbon CD, Fanshawe JH (1988) Stotting in Thomson’s gazelles: an honest signal of condition. Behav Ecol Sociobiol 23:69–74. doi: 10.1007/BF00299889
- Gentner TQ (2008) Temporal scales of auditory objects underlying birdsong vocal recognition. J Acoust Soc Am 124:1350–1359. doi: 10.1121/1.2945705
- Gil D, Slater PJB (2000) Song organisation and singing patterns of the Willow Warbler, *Phylloscopus trochilus*. Behaviour 137:759–782.
- Hasselquist D, Bensch S, von Schantz T (1996) Correlation between male song repertoire, extra-pair paternity and offspring survival in the Great Reed Warbler. Nature 381:229–232.
- Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–79. doi: 10.1126/science.298.5598.1569
- Henry L, Barbu S, Lemasson A, Hausberger M (2015) Dialects in animals: evidence, development and potential functions. Anim Behav Cogn 2:132–155. doi: 10.12966/abc.05.03.2015

- Hiebert SM, Stoddard PK, Arcese P (1989) Repertoire size, territory acquisition and reproductive success in the Song Sparrow. *Anim Behav* 37:266–273. doi: 10.1016/0003-3472(89)90115-2
- Hinde R (1958) Alternative motor patterns in chaffinch song. *Anim Behav* 6:211–218.
- Holland J, Dabelsteen T, Paris A (2000) Coding in the song of the wren: importance of rhythmicity, syntax and element structure. *Anim Behav* 60:463–470. doi: 10.1006/anbe.2000.1529
- Holland J, Dabelsteen T, Pedersen SB, Larsen ON (1998) Degradation of wren *Troglodytes troglodytes* song: Implications for information transfer and ranging. *J Acoust Soc Am* 103:2154–2166. doi: 10.1121/1.421361
- Holveck MJ, Vieira De Castro AC, Lachlan RF, et al. (2008) Accuracy of song syntax learning and singing consistency signal early condition in Zebra Finches. *Behav Ecol* 19:1267–1281. doi: 10.1093/beheco/arn078
- Horn A, Dickinson T, Falls J (1993) Male quality and song repertoires in western meadowlarks (*Sturnella neglecta*). *Can J Zool* 71:1059–1061.
- Hosoi SA, Rothstein SI, O’Loughlen AL (2005) Sexual preferences of female Brown-Headed Cowbirds (*Molothrus ater*) for perched song repertoires. *Auk* 122:82–93. doi: 10.1642/0004-8038(2005)122
- Howes-Jones D (1985) The complex song of the Warbling Vireo. *Can J Zool* 63:2756–2766.
- James RD (1981) Factors affecting variation in the primary song of North American Solitary Vireos (Aves: Vireonidae). *Can J Zool* 59:2001–2009.
- James RD (1976) Unusual songs with comments on song learning among vireos. *Can J Zool* 54:1223–1226. doi: 10.1139/z76-138
- James RD (1984) Structure, frequency of usage, and apparent learning in the primary song of the Yellow-throated Vireo, with comparative notes on Solitary Vireos (Aves: Vireonidae). *Can J Zool* 62:468–472. doi: 10.1139/z84-071
- Jin DZ (2013) The neural basis of birdsong syntax. *Prog. Cogn. Sci. from Cell. Mech. to Comput. Theor.* 1951:
- Jin DZ, Kozhevnikov AA (2011) A compact statistical model of the song syntax in Bengalese finch. *PLoS Comput Biol* 7:e1001108. doi: 10.1371/journal.pcbi.1001108
- Katahira K, Okanoya K, Okada M (2007) A neural network model for generating complex birdsong syntax. *Biol Cybern* 97:441–448. doi: 10.1007/s00422-007-0184-y
- Kershenbaum A, Bowles AE, Freeberg TM, et al. (2014) Animal vocal sequences: not the Markov chains we thought they were. *Proc R Soc B* 281:20141370.

- Kramer H, Lemon R, Morris M (1985) Song switching and agonistic stimulation in the Song Sparrow (*Melospiza melodia*): five tests. *Anim Behav* 33:135–149. doi: 10.2173/bna.704
- Kramer HG, Lemon RE (1983) Dynamics of territorial singing between neighboring Song Sparrows (*Melospiza melodia*). *Behaviour* 85:198–223. doi: 10.1017/CBO9781107415324.004
- Krebs J, Ashcroft R, Orsdol K (1981) Song matching in the Great Tit *Parus major* L. *Anim Behav* 29:918–923.
- Krebs JR, Dawkins R (1984) Animal signals: mind-reading and manipulation. *Behav. Ecol. An Evol. Approach*. Oxford, UK, pp 380–403
- Kroodsma DE (1976) Reproductive development in a female songbird: differential stimulation by quality of male song. *Science* (80- ) 192:574–5. doi: 10.1126/science.192.4239.574
- Kroodsma DE, Bereson RC, Byers BE, Minear E (1989) Use of song types by the Chestnut-sided Warbler: evidence for both intra- and inter-sexual functions. *Can J Zool* 67:447–456. doi: 10.1139/z89-065
- Lachlan R, Anderson R, Peters S, et al. (2014) Typical versions of learned Swamp Sparrow song types are more effective signals than are less typical versions. *Proc R Soc B Biol Sci* 281:1–7.
- Lein MR (1978) Song variation in a population of chestnut-sided warblers (*Dendroica pensylvanica*): its nature and suggested significance. *Can J Zool* 56:1266–1283. doi: 10.1139/z78-182
- Lemon R, Falls J, Dickinson T, et al. (2000) Song clustering by meadowlarks: Is it related to repertoire size? *Behaviour* 137:75–92.
- Lemon RE, Chatfield C (1973) Organization of song of Rose-breasted Grosbeaks. *Anim Behav* 21:28–44.
- Lemon RE, Chatfield C (1971) Organization of song in cardinals. *Anim Behav* 19:1–17. doi: 10.1016/S0003-3472(71)80129-X
- Liu W, Kroodsma DE (2006) Song learning by Chipping Sparrows: when, where, and from whom. *Condor* 108:509–517.
- Logue DM, Forstmeier W (2008) Constrained performance in a communication network: implications for the function of song-type matching and for the evolution of multiple ornaments. *Am Nat* 172:34–41. doi: 10.1086/587849
- Luther DA, Derryberry EP (2012) Birdsongs keep pace with city life: Changes in song over time in an urban songbird affects communication. *Anim Behav* 83:1059–1066. doi: 10.1016/j.anbehav.2012.01.034

- Lynch A (1996) The population memetics of birdsong. *Ecol Evol Acoust Commun birds* 181–197.
- MacDonald IF, Kempster B, Zann L, MacDougall-Shackleton SA (2006) Early nutritional stress impairs development of a song-control brain region in both male and female juvenile Song Sparrows (*Melospiza melodia*) at the onset of song learning. *Proc R Soc London B* 273:2559–2564. doi: 10.1098/rspb.2006.3547
- Macdougall-Shackleton SA (1997) Sexual selection and the evolution of song repertoires. *Curr Ornithol* 14:81–124.
- Markowitz JE, Ivie E, Kligler L, Gardner TJ (2013) Long-range order in canary song. *PLoS Comput Biol* 9:e1003052. doi: 10.1371/journal.pcbi.1003052
- Marler P (1998) Animal Communication and Human Language. In: Leviton A, Martin K (eds) *Orig. Diversif. Lang.* California Academy of Sciences, San Francisco, CA, pp 1–20
- Marler P (2004) Bird calls: a cornucopia for communication. *Nature's Music Sci. Birdsong.* Elsevier, San Diego, CA, pp 135–180
- Marler P, Isaac D (1960) Physical analysis of a simple bird song as exemplified by the Chipping Sparrow. *Condor* 62:124–135.
- Marler P, Peters S (1982) Developmental overproduction and selective attrition: new processes in the epigenesis of birdsong. *Dev Psychobiol* 15:369–378. doi: 10.1002/dev.420150409
- Marler P, Tamura M (1962) Song “dialects” in three populations of White-crowned Sparrows. *Condor* 64:368–377.
- Martens MJM (1980) Foliage as a low-pass filter: Experiments with model forests in an anechoic chamber. *J Acoust Soc Am* 67:66. doi: 10.1121/1.383790
- Mennill DJ (2002) Female eavesdropping on male song contests in songbirds. *Science* 296:873–873. doi: 10.1126/science.296.5569.873
- Molles L (2006) Singing complexity of the Banded Wren (*Thryothorus pleurostictus*): Do switching rate and song-type diversity send different messages? *Auk* 123:991–1003.
- Molles LE, Vehrencamp SL (2001a) Songbird cheaters pay a retaliation cost: evidence for auditory conventional signals. *Proc Biol Sci* 268:2013–2019. doi: 10.1098/rspb.2001.1757
- Molles LE, Vehrencamp SL (2001b) Neighbour recognition by resident males in the Banded Wren, *Thryothorus pleurostictus*, a tropical songbird with high song type sharing. *Anim Behav* 61:119–127. doi: 10.1006/anbe.2000.1561
- Morgan BJT (1976) Markov properties of sequences of behaviours. *J R Stat Soc* 25:31–36.
- Morrison ML, Hardy JW (1983) Vocalizations of the Black-throated Gray Warbler. *Wilson Bull*

95:640–643.

- Morse DH (1966) The context of songs in the Yellow Warbler. *Wilson Bull* 78:444–455.
- Morse DH (1967) The contexts of songs in Black-throated Green and Blackburnian Warblers. *Wilson Bull* 79:64–74.
- Morton ES (1975) Ecological sources of selection on avian sounds. *Am Nat* 109:17–34.
- Morton ES, Howlett J, Kopysh NC, Chiver I (2006) Song ranging by incubating male Blue-headed Vireos: the importance of song representation in repertoires and implications for song delivery patterns and local/foreign dialect discrimination. *J F Ornithol* 77:291–301. doi: 10.1111/j.1557-9263.2006.00055.x
- Moseley D, Lahti D, Podos J (2013) Responses to song playback vary with the vocal performance of both signal senders and receivers. *Proc. R. Soc. B* 280:
- Moser-Purdy C, Mennill DJ (2016) Large vocal repertoires do not constrain the dear enemy effect: a playback experiment and comparative study of songbirds. *Anim Behav* 118:55–64. doi: 10.1016/j.anbehav.2016.05.011
- Moskoff W, Robinson SK (2011) Philadelphia Vireo (*Vireo philadelphicus*). *Birds North Am. Online*
- Mountjoy DJ, Leger DW (2001) Vireo song repertoires and migratory distance: three sexual selection hypotheses fail to explain the correlation. *Behav Ecol* 12:98–102. doi: 10.1093/oxfordjournals.beheco.a000385
- Nowicki S, Peters S, Podos J (1998) Song learning, early nutrition and sexual selection in songbirds. *Am Zool* 38:179–190.
- Nowicki S, Searcy WA, Hughes M, Podos J (2001) The evolution of bird song: male and female response to song innovation in Swamp Sparrows. *Anim Behav* 62:1189–1195. doi: 10.1006/anbe.2001.1854
- Nowicki S, Searcy WA, Peters S (2002) Brain development, song learning and mate choice in birds: A review and experimental test of the “nutritional stress hypothesis.” *J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol* 188:1003–1014. doi: 10.1007/s00359-002-0361-3
- Otter K, McGregor PK, Terry AMR, et al. (1999) Do female great tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proc R Soc B Biol Sci* 266:1305–1309. doi: 10.1098/rspb.1999.0779
- Ouattara K, Lemasson A, Zuberbuhler K (2009) Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *Proc Natl Acad Sci* 106:22026–22031.
- Owren MJ, Rendall D, Ryan MJ (2010) Redefining animal signaling: Influence versus

- information in communication. *Biol Philos* 25:755–780. doi: 10.1007/s10539-010-9224-4
- Peake TM, Matessi G, Mcgregor PK, Dabelsteen T (2005) Song type matching, song type switching and eavesdropping in male great tits. *Anim Behav* 69:1063–1068. doi: 10.1016/j.anbehav.2004.08.009
- Peters S, Searcy WA, Beecher MD, Nowicki S (2000) Geographic variation in the organization of Song Sparrow repertoires. *Auk* 117:936–942.
- Pfaff JA, Zann L, MacDougall-Shackleton SA, MacDougall-Shackleton EA (2007) Song repertoire size varies with HVC volume and is indicative of male quality in Song Sparrows (*Melospiza melodia*). *Proc Biol Sci* 274:2035–2040. doi: 10.1098/rspb.2007.0170
- Pitocchelli J (2014) Song discrimination by male Mourning Warblers (*Geothlypis philadelphia*) and implications for population divergence across the breeding range. *Can F Nat* 128:408–415.
- Podos J, Peters S, Rudnicki T, et al. (1992) The organization of song repertoires in Song Sparrows - themes and variations. *Ethology* 90:89–106. doi: Article
- Podos J, Warren PS (2007) The evolution of geographic variation in birdsong. *Adv Study Behav* 37:403–458. doi: 10.1016/S0065-3454(07)37009-5
- Price J (2013) Why is birdsong so repetitive? Signal detection and the evolution of avian singing modes. *Behaviour* 150:1–19. doi: 10.1163/1568539X-00003051
- Reid JM, Arcese P, Cassidy ALE V, et al. (2005) Fitness correlates of song repertoire size in free-living song sparrows (*Melospiza melodia*). *Am Nat* 165:299–310. doi: 10.1086/428299
- Rice J (1978) Behavioural interactions of interspecifically territorial vireos. I. Song discrimination and natural interactions. *Anim Behav* 26:527–549.
- Robinson S (1981) Ecological relations and social interactions of Philadelphia and Red-eyed Vireos. *Condor* 83:16–26.
- Rogers D (2004) Repertoire size, song sharing and type matching in the Rufous Bristlebird (*Dasyornis broadbenti*). *Emu* 104:7–13.
- Samuels BD (2015) Can a bird brain do phonology? *Front Psychol* 6:1–10. doi: 10.3389/fpsyg.2015.01082
- Searcy W (1984) Song repertoire size and female preferences in Song Sparrows. *Behav Ecol Sociobiol* 14:281–286.
- Searcy W (1992) Song repertoire and mate choice in birds. *Am Zool* 32:71–80.
- Searcy W, Anderson R, Nowicki S (2006) Bird song as a signal of aggressive intent. *Behav Ecol Sociobiol* 60:234–241. doi: 10.1007/S00265-006-0161-9

- Searcy WA (1990) Species recognition of song by female Red-winged Blackbirds. *Anim Behav* 40:1119–1127. doi: 10.1016/S0003-3472(05)80178-2
- Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. *Anim Behav* 78:1281–1292. doi: 10.1016/j.anbehav.2009.08.011
- Simpson B (1985) Effects of location in territory and distance from neighbours on the use of song repertoires by Carolina wrens. *Anim Behav* 33:793–804.
- Slabbekoorn H, Eilers J, Smith TB (2002) Birdsong and sound transmission: the benefits of reverberations. *Condor* 104:564–573.
- Slabbekoorn H, Peet M (2003) Ecology: Birds sing at a higher pitch in urban noise. *Nature* 424:267. doi: 10.1038/424267a
- Slabbekoorn H, Smith TB (2002a) Bird song, ecology and speciation. *Philos Trans R Soc Lond B Biol Sci* 357:493–503. doi: 10.1098/rstb.2001.1056
- Slabbekoorn H, Smith TB (2002b) Habitat-dependent song divergence in the Little Greenbul: an analysis of environmental selection pressures on acoustic signals. *Evolution* 56:1849–58.
- Slabbekoorn H, Yeh P, Hunt K (2007) Sound transmission and song divergence: a comparison of urban and forest acoustics. *Condor* 67–78.
- Slager DL, Battey CJ, Bryson RW, et al. (2014) A multilocus phylogeny of a major New World avian radiation: The Vireonidae. *Mol Phylogenet Evol.* doi: 10.1016/j.ympev.2014.07.021
- Smith WJ, Pawlukiewicz J, Smith ST (1978) Kinds of activities correlated with singing patterns of the Yellow-throated Vireo. *Anim Behav* 26:862–884. doi: 10.1016/0003-3472(78)90153-7
- Soma M, Garamszegi LZ (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. *Behav Ecol* 22:363–371. doi: 10.1093/beheco/arq219
- Spector D (1992) Wood-warbler song systems: a review of paruline singing behaviors. *Curr Ornithol* 11:199–238.
- Spencer K a, Buchanan KL, Goldsmith a R, Catchpole CK (2004) Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). *Proc R Soc B Biol Sci* 271 Suppl :S121–S123. doi: 10.1098/rsbl.2003.0122
- Spencer KA, Buchanan KL, Goldsmith AR, Catchpole CK (2003) Song as an honest signal of developmental stress in the zebra finch (*Taeniopygia guttata*). *Horm Behav* 44:132–139. doi: 10.1016/S0018-506X(03)00124-7
- Stoddard PK, Beecher MD, Campbell SE, Horning CL (1992) Song-type matching in the Song Sparrow. *Can J Zool* 70:1440–1444.



- Suzuki TN (2014) Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. *Anim Behav* 87:59–65. doi: 10.1016/j.anbehav.2013.10.009
- Suzuki TN, Wheatcroft D, Griesser M (2016) Experimental evidence for compositional syntax in bird calls. *Nat Commun* 7:10986. doi: 10.1038/ncomms10986
- Taff CC, Steinberger D, Clark C, et al. (2012) Multimodal sexual selection in a warbler: plumage and song are related to different fitness components. *Anim Behav* 84:813–821. doi: 10.1016/j.anbehav.2012.07.002
- ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philos Trans R Soc B Biol Sci* 367:1984–1994. doi: 10.1098/rstb.2012.0055
- Trillo PA, Vehrencamp SL (2005) Song types and their structural features are associated with specific contexts in the Banded Wren. *Anim Behav* 70:921–935. doi: 10.1016/j.anbehav.2005.02.004
- Vehrencamp SL (2001) Is song-type matching a conventional signal of aggressive intentions? *Proc Biol Sci* 268:1637–42. doi: 10.1098/rspb.2001.1714
- Vehrencamp SL, Hall ML, Bohman ER, et al. (2007) Song matching, overlapping, and switching in the Banded Wren: the sender's perspective. *Behav Ecol* 18:849–859. doi: 10.1093/beheco/arm054
- Verner J (1975) Complex song repertoire of male long-billed marsh wrens in eastern Washington. *Living Bird* 14:263–300.
- Williams H, Levin II, Norris DR, et al. (2013) Three decades of cultural evolution in Savannah sparrow songs. *Anim Behav* 85:213–223. doi: 10.1016/j.anbehav.2012.10.028
- Wood WE, Yezerinac SM (2006) Song sparrow (*Melospiza melodia*) song varies with urban noise. *Auk* 123:650–659. doi: 10.1642/0004-8038(2006)123{[]650:SSMMSV]2.0.CO;2
- Wunderle Jr. J, Cortes R, Carronero W (1992) Song characteristics and variation in a population of Bananaquits on Puerto Rico. *Condor* 94:680–691.
- Yasukawa K, Blank J, Patterson C (1980) Song repertoires and sexual selection in the red-winged blackbird. *Behav Ecol ...* 7:233–238.
- Zahavi A (1980) Ritualization and the evolution of movement signals. *Behaviour* 72:77–80.

## CHAPTER 2

### Composition and Sequential Organization of Song Repertoires in Cassin's Vireo (*Vireo cassinii*)

#### 2.1 Abstract

The rules governing bird song sequences vary considerably across the avian phylogeny and modifications to these rules represent one of the many ways in which bird song varies interspecifically. Cassin's Vireo (*Vireo cassinii*) is one species which shows a highly structured syntax, with clearly non-random patterns of sequential organization in their songs. Here I present a description of Cassin's Vireo song sequences from the Sierra Nevada Mountains in California and employ network analysis to quantify transition patterns within the songs. Repertoire sizes varied between 44 and 60 phrase-types per bird for the thirteen birds analyzed here. The repertoire was subdivided into 'themes' containing between two and seven phrase-types. The birds sang the phrase-types in a given theme for a time before eventually introducing a new theme; in this manner the repertoire was revealed relatively slowly over time. Theme composition within a bird's repertoire did not change within or between singing bouts throughout the season. The tendency to sing in themes was corroborated by network analysis, which revealed small-world structure in the songs. Phrase-types were widely shared within the population. I discuss these findings as they compare with the singing styles of other species, both closely- and distantly-related.

#### 2.2 Introduction

Complex syntax is a defining feature of the songs of many species of birds. Given its prevalence in bird songs, it is likely that syntax plays a role in the conveyance of information from sender to receiver (Kershenbaum et al. 2014). The syntax of most species of birds appear to follow sets of rules, the nature and complexity of which vary greatly between species (Bolhuis

and Everaert 2013), and which are sufficiently variable that the particular structure of sequences could conceivably contain information about the sender or their surroundings. To fully understand the information contained within the songs of any species, it is first necessary to evaluate in detail the nature of their song repertoire and syntax.

Though most songbird species utilize song repertoires comprising multiple vocal units, and are therefore potential subjects of studies of syntax, research to date has overwhelmingly focused on variation in repertoire size and composition between species (e.g. Byers and Kroodsma 2009) while largely ignoring the question of how the vocal units are combined into sequences. To do so overlooks the possibility that the ordering of songs may convey biologically relevant information about individual fitness (Leitner et al. 2006) or external factors such as predators (Baker and Becker 2002). Furthermore, in some species investigations into song syntax have provided more insight into the nature of songs than subjective estimates of repertoire size alone. For example, re-arrangement of relatively few vocal units according to probabilistic rules gives rise to considerable variability in the songs of many species (e.g. Sedge Warbler, Catchpole 1976; Willow Warbler, Gil and Slater 2000; Brown-throated Wren, Sosa-López and Mennill 2013; Bengalese Finch, Jin and Kozhevnikov 2011).

The songs of birds are subject to selection based on various ecological variables, including habitat characteristics (Morton 1975; Anderson and Conner 1985; Slabbekoorn and Smith 2002; Boncoraglio and Saino 2007), mating systems (Catchpole 1987, but see Soma and Garamszegi 2011) and migratory distance (Read and Weary 1992; Mountjoy and Leger 2001). Detailed and accurate descriptions of the songs of a given species can greatly facilitate future efforts to understand the influence of these factors on the evolution of bird songs. In this study, I

will lay such a foundation by describing the songs of Cassin's Vireo (*Vireo cassinii*) with a focus on the syntactical patterns that emerge over the course of long bouts of song.

Cassin's Vireo is a migratory songbird that breeds in the forests of western North America. Past work analyzing the songs of the Solitary Vireo complex included only brief analysis of the songs of Cassin's Vireo (James 1981), but the recent elevation of Cassin's Vireo to full species status warrants a revisitation of past efforts (American Ornithologists' Union 1997). Cassin's Vireo belongs to the Vireonidae, a family whose species display a striking variety of singing styles. Songs vary from the simple repeated phrases of Hutton's Vireo, a species that possesses small repertoires of approximately three phrase-types (Mountjoy and Leger 2001), to the highly variable songs of the Warbling Vireo, which combines subunits to produce a heretofore unquantified variety of songs (Howes-Jones 1985). Intermediate between these two extremes are a few close relatives of Cassin's Vireo: Blue-headed Vireo repertoires have been estimated to contain twelve phrase-types (James 1981), while Yellow-throated and White-eyed Vireos possess repertoires of five (James 1984) and twelve songs (Borror 1987), respectively. The intermediate repertoire sizes of these species appear to be representative of the genus: a review of 28 *Vireo* species by Mountjoy and Leger (2001) showed a mean repertoire size of twelve.

This paper aims to identify the repertoire sizes of Cassin's Vireo individuals and outline the patterns of repertoire delivery that became apparent over the course of extensive recordings on a population of thirteen individuals. I begin with an assessment of the repertoire sizes of the males in the focal population, using classification algorithms to confirm judgments of discrete phrase-types. Then I illustrate patterns of song sharing amongst males. Finally, I describe the

sequential arrangement of the songs and discuss the findings with respect to previously published work on the syntax of the songs of Vireos and other songbirds.

## **2.3 Materials and Methods**

### **2.3.1 Study Site**

Fieldwork was conducted on private land five kilometers north of the town of Volcano in Amador County, California, USA, in the foothills of the Sierra Nevada Mountains at an elevation of 750m (10 S 706584 4262742, datum WGS 84). The forest is characterized by mixed evergreen and deciduous trees; the tree community is dominated by Incense Cedar (*Calocedrus decurrens*), Douglas Fir (*Pseudotsuga menziesii*), Live Oak (*Quercus chrysolepis*), Black Oak (*Quercus kelloggii*) and Ponderosa Pine (*Pinus ponderosa*). Fieldwork took place between April 25 and June 28, 2013 and between May 5 and June 25, 2014.

### **2.3.2 Field Methods**

I made recordings using a Marantz PMD661 solid state digital recorder and a Sennheiser MKH20-P48 microphone with a Telinga parabolic reflector. Recordings were saved to wav files using a sampling rate of 44kHz. Recordings were made opportunistically: I first located an individual aurally, then approached and recorded it for as long as possible, stopping only when the bird either ceased singing for a considerable period of time or moved too far away to permit a sufficiently high quality recording. Due to the unpredictable nature of the birds' behaviors, recordings varied in the number of phrases they contained. Since the goal of the study was to investigate patterns of sequential delivery, I discarded recordings containing fewer than 50 phrases. The resulting dataset contained 216 recordings, varying in length from less than two minutes to more than two hours in length (mean=0:25:02, median=16:44) and containing between 50 and 1525 phrases per recording (mean=252, median=173).

Seven males were recorded in 2013. Five of these individuals returned to breed in the same area in 2014 and were thus recorded in both years; a further six males were recorded in 2014, for a total of thirteen males in the focal population. During the 2013 season, individuals were identified based on their consistent association with non-overlapping territories, and identifications were later confirmed based on the phrase-types in the recording, justified by the observation that each male possesses an individually distinct repertoire. To identify males between years, I again capitalized on the song characteristics, and identified individuals using an ensemble of classifiers which have been shown to identify individuals with near perfect accuracy (J. Arriaga, unpublished). In the five instances where the algorithm identified the same male at the study site in both years, the male established a territory with roughly the same boundaries, providing further evidence that the identification was accurate (Fig. 2.1). In early May of 2014, all males were banded with unique color combinations for subsequent visual identification. Individuals are here referred to by a unique letter combination corresponding to the colors of their band combination, with the exception of the two birds that were only recorded in 2013, which are referred to as ‘Meadow’ and ‘Gate’ individuals. The number of phrases recorded from each individual ranged from 1335 to 10336 (mean=4193). The aggregate number of phrases annotated from all thirteen birds was 54515. All recordings and annotations used in this analysis are available online in a birdsong database at <http://taylor0.biology.ucla.edu/birdDBQuery>.

I plotted individual territories onto a map of the study area throughout each season by observing singing bouts and counter-singing interactions between neighbors (Fig 2.1). Some territories abutted one another, while others were separated by up to 700 meters.

### **2.3.3 Terminology**

Assessments of repertoire sizes and interpretations of song complexity can vary with the level of analysis; thus, it is important to explicitly define the elements of songs as they pertain to this study. *Phrases* are short bursts of sound less than 0.7 seconds in duration (mean=0.34s) and usually separated from preceding and succeeding phrases by at least one second of silence (Fig. 2.2). *Phrase-types* are categories of phrases that are considered equivalent based on shared spectrographic characteristics. A *song bout* was defined as a long string of phrases delivered in sequence.

#### **2.3.4 Phrase Identification and Annotation**

Cassin's Vireos deliver songs in two audibly distinct ways: their *typical song* consists of phrases separated by one or more seconds of silence, while their *chatter vocalizations* are sequences of phrases and call notes delivered rapidly and continuously without pauses. Previous authors have referred to these latter vocalizations as 'pre-copulatory songs' (James 1978). For the purposes of this study I focused exclusively on the typical song, which accounted for more than 99% of the total vocal output of the birds. I have observed that phrase-types found in the chatter vocalizations are often distinct from those in the typical song; individuals that do not deliver a certain phrase-type in their typical repertoire may include it in their chatter songs. The structure and function of chatter songs remain as topics for future study.

I used the linguistic program Praat (Boersma and Weenink 2014) to annotate recordings. By visually inspecting spectrograms from 216 recordings of thirteen individuals, I created a catalogue containing the spectrograms of phrase-types against which subsequent phrases could be compared. Each phrase-type was assigned a unique two letter code and was subjectively identifiable based on spectrographic characteristics and acoustic structure.

I marked the start and end times of phrases in each recording with approximately +/- 0.01s accuracy, then identified the phrase-types by visually comparing their spectrograms with the reference file. Phrases were highly stereotyped, such that there were few visible differences between any two exemplars of the same phrase-type and phrase-type classification was unambiguous. Portions of some recordings were not annotated because of uncertainty regarding the identity of the bird. In such cases where identity was in doubt or recording quality suffered, I annotated only the recording either before or after the unusable portion.

The classification of phrase-types by visually inspecting spectrograms is inherently subjective, but in my experience has proven to be the most noise-robust technique currently available for this species. Since my ultimate goal was to study the sequential arrangements of phrase-types in the songs of Cassin's Vireos, I could not sacrifice accuracy or noise robustness for the speed permitted by automated classification algorithms. Still, to ensure that my subjective classifications represented objectively identifiable and discrete phrase-types, I verified a subset of my annotations using a supervised machine learning algorithm implemented by Tan et al. (2013); this is a sparse-representation algorithm that uses spectrographic features to perform classification. The algorithm was designed specifically for use on Cassin's Vireo phrases, and has been shown to perform at a high level on recordings with minimal background noise (Tan et al. 2012; Tan et al. 2013). I ran two tests with the classification algorithm. In both cases, I used only clean recordings for both training and testing, meaning I first manually inspected the annotated sound files and removed all instances of phrases with low signal to noise ratios. For the first test, I trained the algorithm using four recordings from four different individuals and tested on phrases from seven individuals at the study site. This tested the ability of the algorithm to generalize across individuals, and tested whether my evaluation of the extent of phrase-type



sharing between males was supported. In the second test, I trained the algorithm using a single recording from seven individuals, and tested on 47 files from those same individuals. The algorithm used five exemplars of each phrase-type in the training set, ignoring phrase-types that occurred less than five times in the training files. As a result, the algorithm was not trained and tested on all phrase-types used by the focal birds. In the first test, the algorithm was tested on 6863 examples of 83 phrase-types, and in the second it was tested on 7145 examples of 90 phrase-types.

### **2.3.5 Network Analysis**

I calculated the small-world coefficient for each of the 216 recordings. Networks that meet the criteria for small-worldness are characterized by high local clustering and short path lengths between vertices; small-world structure has been identified in social networks, power grids, and neural networks (Watts and Strogatz 1998). Measures of small-worldness applied to bird song measure the tendency for phrase-types to be clustered together within a bird's repertoire, such that certain phrase-types consistently appear near each other in sequences more than would be expected by chance. The methods I employed were identical to those of Sasahara et al. (2012), who identified small-world structure in the songs of California Thrashers. For more detail on small-world algorithms, see Watts and Strogatz (1998) and Humphries and Gurney (2008).

I first converted each recording into a transition graph, with each phrase-type represented as a vertex and each observed transition between phrase-types represented as an edge. Self-loops (in which a phrase-type was repeated more than once) were removed from the graph and the directionality was removed from the transitions. The small-world coefficient was calculated using the following formula:

$$S = \left( \frac{C}{C_{rand}} \right) / \left( \frac{L}{L_{rand}} \right)$$

where  $C$  is the average clustering coefficient of the transition graph from a given recording (a measure of the average number of connected edges with respect to the maximum number of edges possible) averaged for each phrase-type in the recording.  $L$  is the characteristic path length (the average number of steps along the shortest paths between all vertices in the graph).  $C_{rand}$  and  $L_{rand}$  are the same measures calculated from randomly generated graphs with the same number of edges and vertices as the observed graph. For each recording 5000 random graphs were generated and the  $C$  and  $L$  values for these graphs were averaged to arrive at estimates for  $C_{rand}$  and  $L_{rand}$ . The small-world coefficient,  $S$ , thus compares the clustering of phrase-types from observed song sequences to randomly generated networks. A value of  $S$  greater than 1 indicates a tendency to organize phrase-types into themes with more connections between phrase-types within the theme than to phrase-types outside of the theme (Humphries and Gurney 2008). Network analysis was conducted using the PajaroLoco package (Sanchez et al. 2015) in Mathematica version 9.0 (Wolfram Research Inc. 2012).

## 2.4 Results

### 2.4.1 Repertoires and Sharing

In total, I identified 122 phrase-types in the focal population, and individual repertoire sizes for the thirteen birds varied between 44 and 60 phrase-types (mean=51.07). Most phrase-types appeared as continuous traces on a spectrogram. Nineteen of the phrase-types, however, were ‘compound’, in that they were comprised of two subunits separated by a very short interval of silence. Some of the subunits of compound phrase-types were themselves independent phrase-types (Fig. 2.2), while other subunits were only delivered as part of a compound phrase-type. Compound phrase-types were considered single phrases because: a) the two subunits were

closely and consistently associated with one another with only very short periods of silence (~0.1s) between subunits; and b) some individuals sang compound phrase-types without ever singing either of the component subunits independently.

Phrase-type sharing was common, and pairwise comparisons between birds showed that males shared on average  $25.4 \pm 4.9$  phrase-types with any other male. A Mantel test conducted with the 'ade4' package in R version 3.1.2 (R Core Team 2014) comparing the pairwise repertoire similarity – calculated as  $C/(\sqrt{A}*\sqrt{B})$ , where  $C$  is the number of shared phrase-types between two individuals, and  $A$  and  $B$  are their respective repertoire sizes – with the distance between territories showed no significant relationship between the two variables in either year (2013:  $r=-0.06$ ,  $p=0.61$ ,  $n=7$  individuals; 2014:  $r=0.1033$ ,  $p=0.26$ ,  $n=11$  individuals; 9999 permutations).

Of the 122 phrase-types encountered in the study, 107 were shared between at least two of the thirteen focal individuals. Two phrase-types were identified in the repertoires of all thirteen individuals, while on average each phrase-type was identified in the repertoires of 5.4 of thirteen individuals. The total number of phrase-types encountered in the population showed signs of leveling off as individuals were added to the sample (Fig. 2.3), an indication that few of the phrase-types were unique to any individual.

#### **2.4.2 Classification Results**

In the first experiment, in which I trained the algorithm using exemplars from four individuals and tested the algorithm's ability to generalize across individuals, the sparse representation algorithm agreed with my spectrogram-inspection method on 99.4% of the 6863 phrases tested. The second experiment, in which I trained the algorithm on a single sound file from each of seven individuals, showed similar results (99.4% agreement on 7145 phrases). I

conclude that the phrase-type categories, while subjectively determined by a human at first, did provide a consistent and objective basis for analyzing songs of the species.

### **2.4.3 Phrase-type Use**

Phrase-types were not used equally within an individual's repertoire (Chi-square goodness of fit test:  $p < 0.0001$  for each of thirteen individuals; fig. 2.4). The most common phrase-type in each individual's repertoire comprised an average of 9.5% (range= 6.1-13.4%) of the total phrases delivered. The rarest phrase-types in each bird's repertoire comprised an average of only 0.07% (range=0.0009-0.5%) of the total phrases.

### **2.4.4 Sequential Arrangement of Phrases**

Consecutive phrases comprised distinct phrase-types 93.4% of the time, so song organization strongly tended towards immediate, rather than eventual variety. The average small-world coefficient from all recordings was 3.73, reflecting a strong tendency towards grouping phrase-types into themes. Sequences from only one of 216 recordings failed to meet the threshold for small-worldness, and this was likely a result of the short length of the recording which contained just 58 phrases. Supporting this idea, there was a significant positive correlation between the number of phrases in a recording and its small-world coefficient (Pearson:  $r^2 = 0.177$ ,  $N = 216$ ,  $p < 0.0001$ ).

A representative sequence of phrases is illustrated in figure 2.5, and shows the typical singing style for the species. In this figure, phrase-types were assigned numbers according to the order in which they were introduced. The male delivered phrase-types in themes, such that groups of phrase-types were consistently delivered in close succession. Each theme contained between 2 and 7 phrase-types, and themes are clearly evident as clusters of points within figure 2.5. The particular order of phrase-types was variable within each theme, although the magnitude

and nature of this variability remains to be investigated. The identity of the phrase-types comprising the themes differed between individuals; within the songs of a given male, however, theme composition was consistent over the course of the season. Temporal conservation of theme composition is evident over a short timeframe in figure 2.5, where the bird repeated previously-delivered themes after approximately 120 phrases had been delivered in the sequence; in each case, the phrase-types in the themes were unchanged from previous instances of the same theme. After delivering each phrase-type in a given theme one or a few times (mean=3.17, median=2), another theme was introduced with minimal overlap between themes. Themes were generally only revisited after a considerable portion of the repertoire had been delivered. All thirteen individuals showed similar tendencies with regards to their arrangements of songs, organizing their phrase-types into discrete themes that were consistently delivered together in sequence and were conserved over the course of the season.

A notable consequence of the use of themes is that once a given phrase-type has been delivered, it is likely to recur again within the next few phrases because transitions between phrase-types within a theme occur more frequently than transitions to other phrase-types in the repertoire. A useful way of illustrating this is by calculating the recurrence interval, defined as the number of intervening phrases between two examples of a given phrase-type. For all individuals, recurrence intervals up to eight were reasonably common, while intervals longer than eight were increasingly rare (Fig. 2.6). All individuals showed a similar trend, with peak recurrence intervals of one, two or three.

## **2.5 Discussion**

### **2.5.1 Sequential Arrangement of Phrases**

The patterns of sequential phrase delivery described here illustrate the inadequacy of the categorization of birds as singing with either immediate or eventual variety. Within the broad category of ‘immediate variety’, there are many possibilities for species-specific syntax; typically these have been overlooked in discussions of the evolutionary significance of song repertoires. As an example, some aspects of Cassin’s Vireo syntax run counter to the idea that large song repertoires are maintained by intersexual selection. First of all, the observation that some phrase-types are exceptionally rare, being encountered only a few times in samples containing as many as several thousand phrases per individual, would not be expected under a female-choice model (Fig. 2.4). Under such a model – commonly invoked to explain large repertoires and immediate variety – males should deliver all phrase-types approximately equally to increase their effective repertoire size (Byers and Kroodsma 2009). In the songs of Cassin’s Vireos, however, many phrase-types in a male’s repertoire were very rarely delivered. Furthermore, under a female choice model, males should arrange their songs cyclically to advertise their repertoire size in an expedient manner. The tendency of Cassin’s Vireos to arrange their songs into themes which are often delivered at length before additional themes are introduced contradicts the patterns expected under an intersexually selected song repertoire. In the case of Cassin’s Vireo, at least, alternate explanations seem required to explain the large repertoires and complex syntax of the species.

The organization of phrase-types into small-world themes is not unique to Cassin’s Vireos. The same finding has been identified in California Thrashers (Sasahara et al. 2012), Southern House Wrens (Deslandes et al. 2014) and Nightingales (Weiss et al. 2014). A similar hierarchical song structure has been qualitatively described in Rock Wrens (Kroodsma 1975), though the author did not employ the same network-based approach to quantify the strength of

the pattern. Common to all of these species are large repertoires of song-types, phrase-types or syllable-types. This may indicate a broad convergence to small-world structure amongst birds with complex song repertoires – at the very least it shows widespread non-random arrangements of song sequences.

Even within the family Vireonidae, the style of singing described here may not be unique. Black-capped Vireos (*Vireo atricapilla*) tend to group their songs into groups of three to seven phrase-types, switching between groups in a fashion comparable to that described here for Cassin's Vireo (Grzybowski 1995). Bell's Vireos (*Vireo bellii*) tend to alternate between groups of two or three song-types, before abruptly switching to another group of song-types (Kus et al. 2010). Similarly, Thick-billed Vireos (*Vireo crassirostris*) alternate between two song-types for a period of time before switching to another pair of song-types (M.R. Walker, unpublished). Though these patterns have not yet been formally analyzed, these singing styles appear to indicate a phylogenetically conserved tendency to group vocal units into themes. Syntax may therefore be influenced in part by the complexity of the repertoire, while also being strongly influenced by evolutionary history.

### **2.5.2 Repertoire Size**

The average repertoire size of 51 phrase-types per individual is 325% larger than the previously published repertoire size for Cassin's Vireo. James (1981) recorded five individuals from a population in British Columbia and identified an average of twelve phrase-types per individual, with a range of between six and fourteen phrase-types. Though he recorded a total of only 251 phrases, or approximately 50 phrases per bird, he maintained that his methodology approached the full repertoire size of the birds in his study population. Subsequent studies on Blue-headed Vireos (*Vireo solitarius*) have identified comparable repertoire sizes of

approximately 15 types, lending some credibility to the reproducibility of James' estimates (Martindale 1980).

The reasons for such a significant difference in repertoire size estimates are unclear. Incomplete sampling on the part of James (1981) most likely played a role; in my study population, new phrase-types were still being encountered after thousands of phrases had been recorded from a single individual (Fig. 2.7). The singing style of the species, in which the bird occasionally delivers a single theme for an extended period of time, may have exacerbated this issue, by giving previous authors a false sense that the repertoire had been completely sampled. The possibility of geographic variation in repertoire sizes could also account for this, and has been identified in other species (McGregor et al. 1981; Kroodsma 1983; Baker 1996; Irwin 2000; Peters et al. 2000).

Another potential cause of the repertoire size differences is methodological: distinguishing between phrase-types by inspecting spectrograms is inherently subjective. Martindale (1980) borrowed terminology from systematics when he framed the issue as one of 'lumpers' versus 'splitters', wherein lumpers tend to categorize similar phrase-types as being the same and splitters would allow even slight differences to characterize distinct types, so long as those differences are consistent. My approach, justified by the remarkable stereotypy in the songs of Cassin's Vireo, tended towards the 'splitter' approach. I believe this approach is generally vindicated by objective methods of supervised learning, discussed above.

The ultimate goal of either approach should be to accurately assess the phrase-types as employed by the birds themselves. In the case of Cassin's Vireos, the sequential arrangements of phrase-types provide hints about individual- and population-level treatment of song types. In figure 2.8, for example, which depicts spectrographically similar phrases from the 'Gate' and



‘Meadow’ individuals from my study site, each phrase-type has clearly distinct acoustic characteristics, and these differences are conserved between individuals. Inspection of figure 2.5, a depiction of a long sequence of the Meadow individual’s song, shows that these three phrase-types were treated as distinct. Phrase-type *ae*, labeled as phrase number 14 in the figure, was delivered twice: both times it was followed by the theme containing phrases numbered 15 to 19. In contrast, phrase-types *au* and *en*, labeled as phrase numbers 6 and 28, respectively, were consistently delivered together in a theme with numbers 4, 5, and 7. Moreover, some individuals sang only one or two of the three types, but never strayed from the stereotyped features of the phrase-types (Fig. 2.8). Distinctions between phrase-types were confirmed by the sparse-representation algorithm, which showed almost complete agreement with my classifications of phrase-types in the population. Taken together, my findings suggest that the repertoires of Cassin’s Vireos are much larger than has been previously appreciated.

### **2.5.3 Song Sharing**

Rates of phrase-type sharing between the focal individuals were high, with pairs of individuals sharing approximately half of their repertoires on average. Even more striking was the fact that the vast majority of phrase-types were shared between at least two individuals. Only fifteen of the 122 phrase-types in the sample were unshared amongst individuals, and subsequent opportunistic recording efforts from elsewhere in the Sierra Nevada Mountains have identified a further ten of these phrase-types in the repertoires of other males (R. Hedley, unpublished data). This raises questions about whether the species is capable of improvising phrase-types, or whether individuals require tutoring for repertoire development. The near complete absence of unshared phrase-types suggests a role for imitation, though the lack of a significant relationship between inter-territory distance and sharing rates leaves open the question of when and from

whom they might acquire their songs. A further question, especially worthwhile for a species with such a distinct syntax is, do individuals share common rules for the sequential delivery of shared phrase-types? If not, are such differences in syntax meaningful to the birds? These questions and questions regarding the functional role of syntax will provide ample directions for future research.

## List of Figures

**Fig. 2.1** Map of territories held by the thirteen male Cassin's Vireos recorded during the study. Individuals recorded in both years held overlapping territories that are depicted with the same color on the map. Territories drawn with thin borders were breeding territories in 2013, while those with thick borders were held in 2014.

**Fig. 2.2** A spectrogram illustrating the typical singing style of Cassin's Vireo. Four phrases of three phrase-types are shown, forming the sequence *aq*, *ar*, *dl*, *aq*. Comparison of the two exemplars of *aq* demonstrates the strongly stereotyped delivery of phrases, with very little within-type variation. The third phrase in the sequence, *dl*, is a compound phrase-type, comprised of two subunits which themselves are occasionally delivered independently as the phrase-types *dr* and *en*.

**Fig. 2.3** Phrase-type accumulation curve showing the number of unique phrase-types in a sample as individuals are added. The cumulative number of phrase-types encountered (solid line) within a population increased as a function of the number of individuals sampled. The number of new phrase-types encountered (dotted line) declined as more individuals were added to the sample, a result of the high levels of phrase-type sharing in the population.

**Fig. 2.4** Relative use of phrase-types from recordings of 'AOBu' individual. Phrase-types were arranged by rank from most common to least common. Analysis of all individuals yielded

qualitatively similar results, showing strong biases favoring the delivery of some phrase-types over others.

**Fig. 2.5** Graphical representation of the sequential arrangement of a song bout from a single recording of the ‘Meadow’ individual. Phrase numbers were assigned based solely on the order of first appearance of each phrase-type within the recording. Themes are clearly identifiable as clusters of phrase-types regularly associated with one another. Phrase-types *au*, *ae*, and *en* are denoted with horizontal lines, illustrating the distinctness of these three spectrographically similar phrase-types in the bird’s repertoire. The style of singing illustrated here is typical of all individuals analyzed.

**Fig. 2.6** Distribution of recurrence intervals for all thirteen individuals in the population, showing a strong tendency towards low recurrence intervals. Horizontal dashed lines show the mean for each value of recurrence interval, and the solid line connects the mean values to illustrate the overall trend. Intervals greater than 8 occurred at low rates, but were not included here.

**Fig. 2.7** Cumulative numbers of phrase-types encountered in the songs of thirteen focal individuals over the course of the study. Most phrase-types were encountered within the first few hundred recorded phrases, but a more substantial recording effort was required to confidently estimate total repertoire sizes. Numbers in parentheses represent the final repertoire size estimate for each bird followed by the total number of phrases recorded from that bird.

**Fig. 2.8** Spectrograms of three similar phrase-types from the repertoires of two different birds. Examples a, c and e are from the ‘Meadow’ individual, while examples b, d and f were recorded from the ‘Gate’ individual. Examples a and b show phrase-type *ae*, c and d show phrase-type *au*, and e and f show phrase-type *en*. Despite obvious similarities between these three phrase-types, variation in spectral characteristics between phrase-types consistently exceeds within-type variation, even when exemplars from different individuals are considered.

## Figures

Figure 2.1

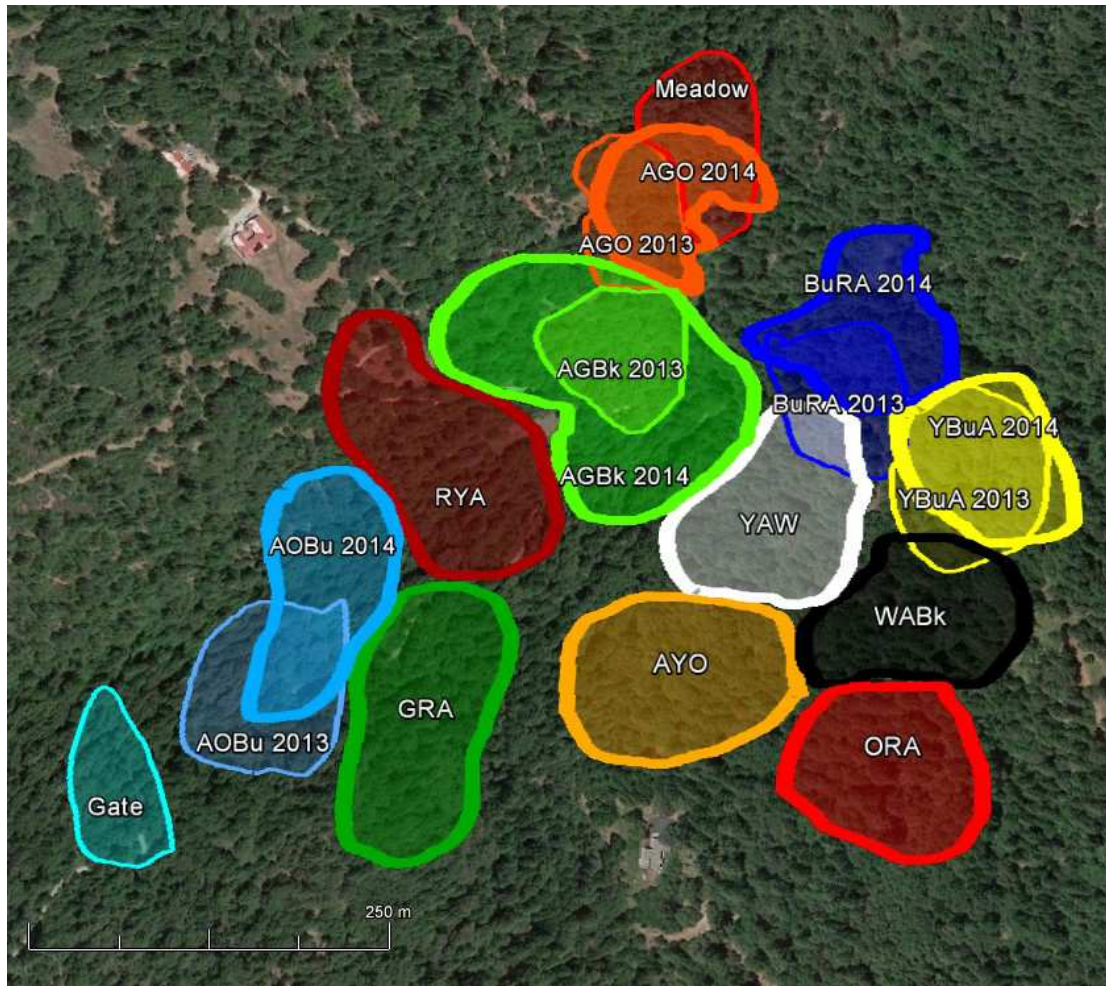


Figure 2.2

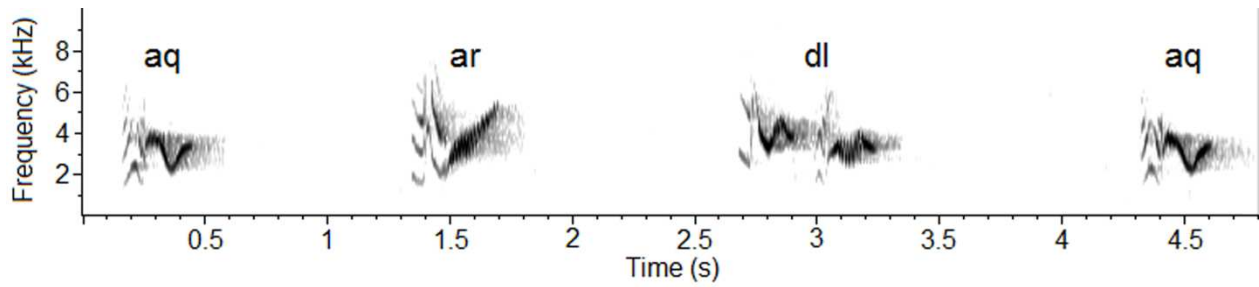


Figure 2.3

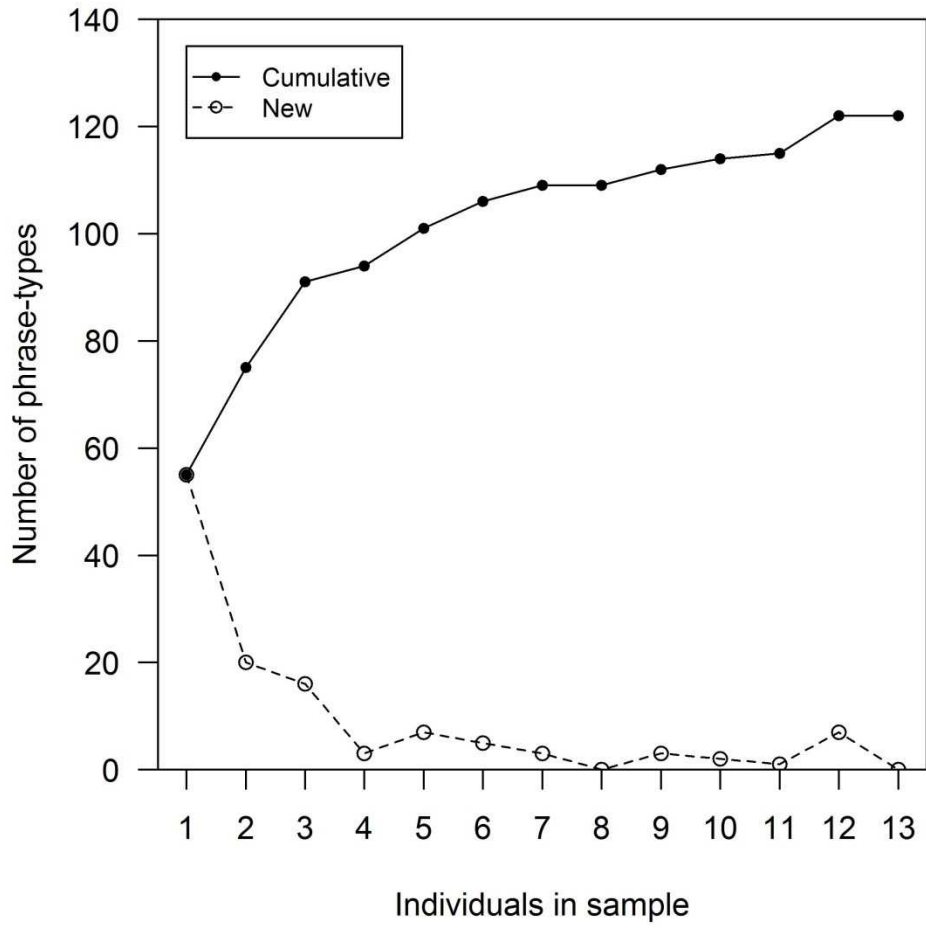


Figure 2.4

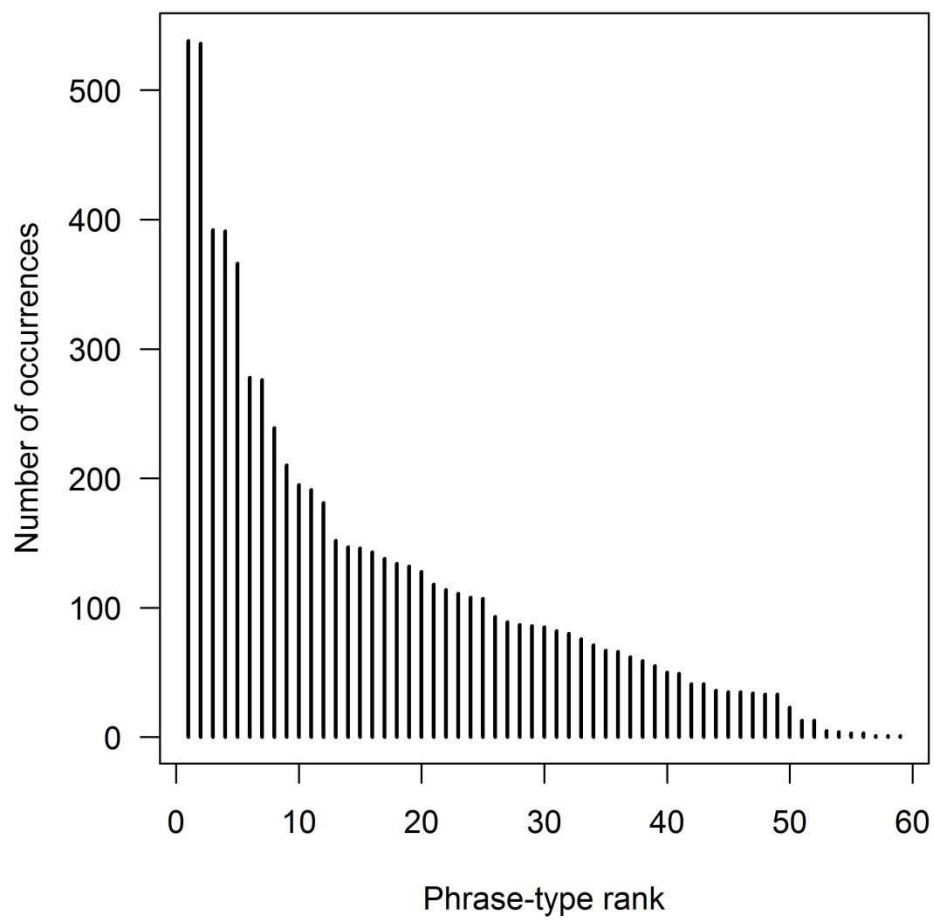




Figure 2.5

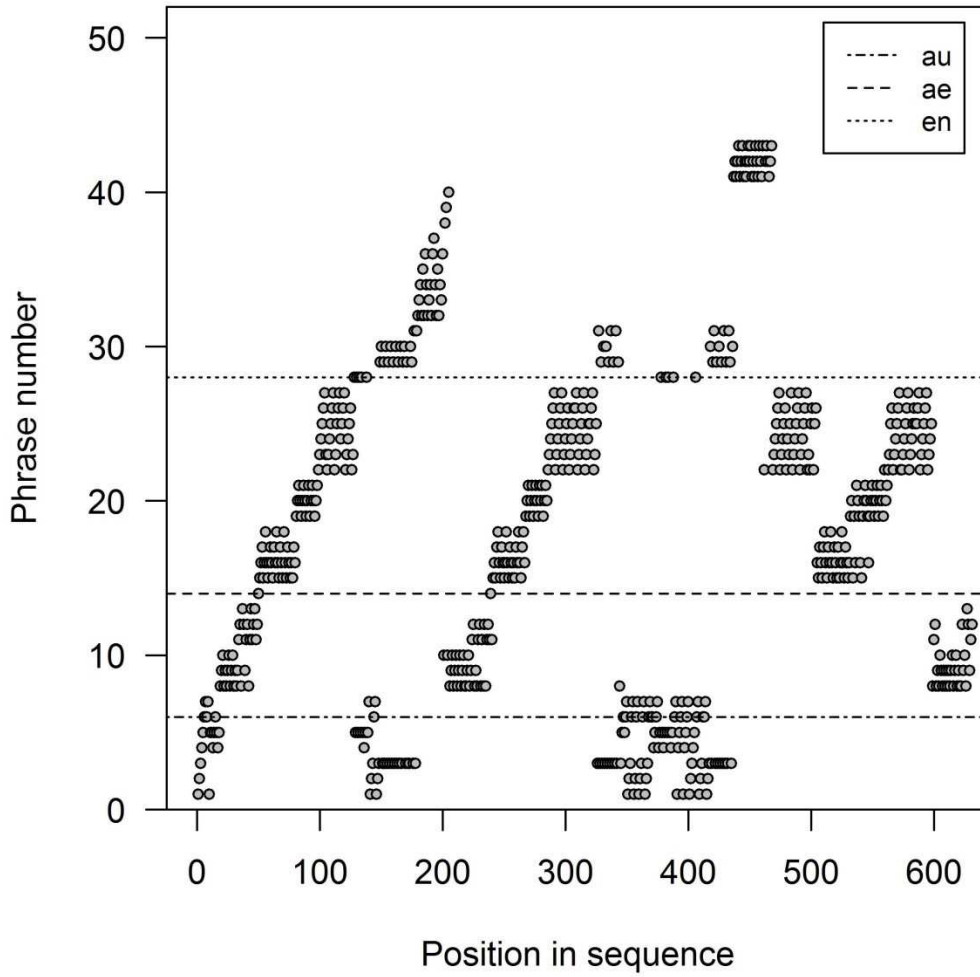


Figure 2.6

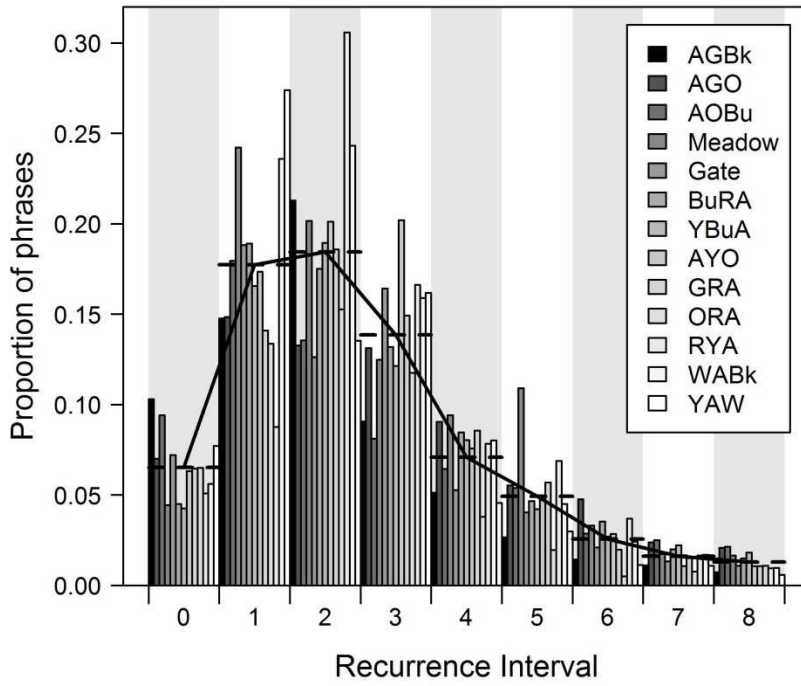


Figure 2.7

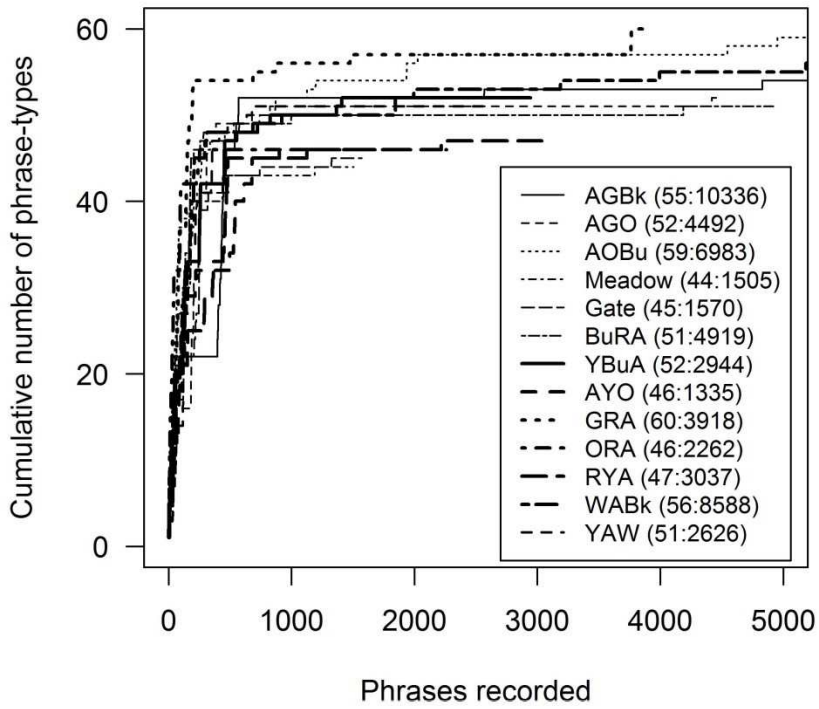
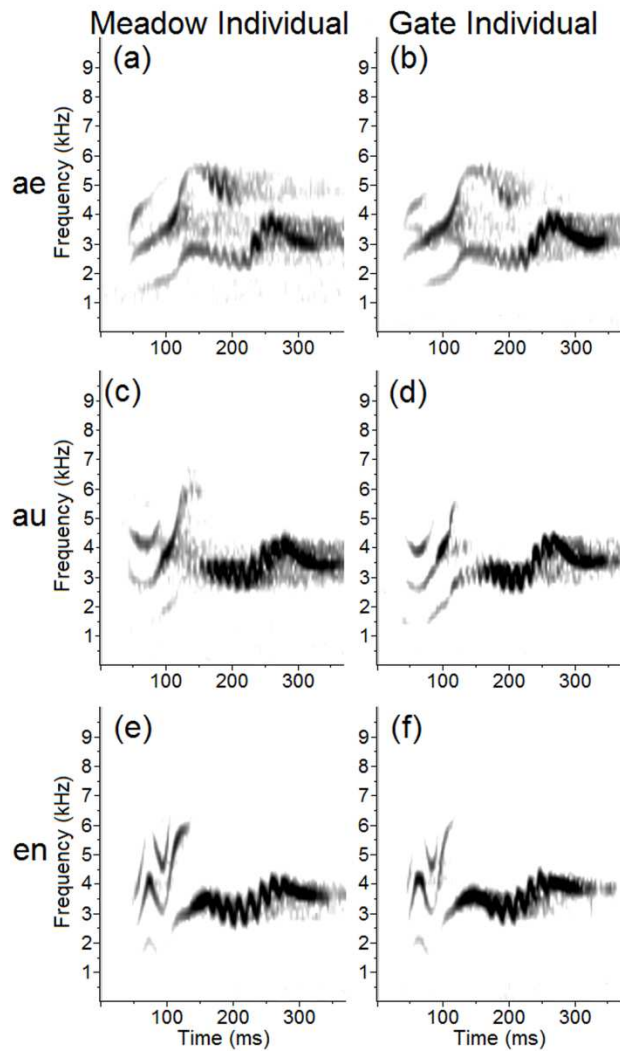


Figure 2.8



## References

- American Ornithologists' Union (1997) Forty-First Supplement to the American Ornithologists' Union Check-List of North American Birds. *Auk* 114:542–552.
- Anderson ME, Conner RN (1985) Northern Cardinal song in three forest habitats in eastern Texas. *Wilson Bull* 97:436–449.
- Baker MC (1996) Depauperate meme pool of vocal signals in an island population of Singing Honeyeaters. *Anim Behav* 51:853–858.
- Baker MC, Becker AM (2002) Mobbing calls of Black-capped Chickadees: effects of urgency on call production. *Wilson Bull* 114:510–516.
- Boersma P, Weenink D (2014) Praat: doing phonetics by computer. Version 5.3.82. <http://www.praat.org/>
- Bolhuis JJ, Everaert M (2013) *Birdsong, speech, and language: exploring the evolution of mind and brain*. MIT Press, Cambridge, MA
- Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. *Funct Ecol* 21:134–142.
- Borror DJ (1987) Song in the white-eyed vireo. *Wilson Bull* 99:377–397.
- Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. *Anim Behav* 77:13–22. doi: 10.1016/j.anbehav.2008.10.003
- Catchpole CK (1976) Temporal and sequential organisation of song in the Sedge Warbler (*Acrocephalus schoenobaenus*). *Behaviour* 59:226–246.
- Catchpole CK (1987) Bird song, sexual selection and female choice. *Trends Ecol Evol* 2:94–7. doi: 10.1016/0169-5347(87)90165-0
- Deslandes V, Faria LRR, Borges ME, Pie MR (2014) The structure of an avian syllable syntax network. *Behav Processes* 106:53–59. doi: 10.1016/j.beproc.2014.04.010
- Gil D, Slater PJB (2000) Song organisation and singing patterns of the Willow Warbler, *Phylloscopus trochilus*. *Behaviour* 137:759–782.
- Grzybowski JA (1995) Black-capped Vireo (*Vireo atricapilla*). In: Poole A (ed) *The birds of North America online*. Cornell lab of ornithology, Ithaca. doi:10.2173/bna.181
- Howes-Jones D (1985) The complex song of the Warbling Vireo. *Can J Zool* 63:2756–2766.
- Humphries MD, Gurney K (2008) Network “small-world-ness”: a quantitative method for determining canonical network equivalence. *PLoS One* 3:e0002051. doi: 10.1371/journal.pone.0002051

- Irwin DE (2000) Song variation in an avian ring species. *Evolution* 54:998–1010.
- James RD (1978) Pairing and nest site selection in Solitary and Yellow-throated Vireos with a description of a ritualized nest building display. *Can J Zool* 56:1163–1169.
- James RD (1981) Factors affecting variation in the primary song of North American Solitary Vireos (Aves: Vireonidae). *Can J Zool* 59:2001–2009.
- James RD (1984) Structure, frequency of usage, and apparent learning in the primary song of the yellow-throated vireo, with comparative notes on solitary vireos (Aves: Vireonidae). *Can J Zool* 62:468–472. doi: 10.1139/z84-071
- Jin DZ, Kozhevnikov AA (2011) A compact statistical model of the song syntax in Bengalese Finch. *PLoS Comput Biol* 7:e1001108. doi: 10.1371/journal.pcbi.1001108
- Kershenbaum A, Blumstein DT, Roch MA, et al. (2014) Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol Rev* doi: 10.1111/brv.12160
- Kroodsma DE (1975) Song patterning in the Rock Wren. *Condor* 77:294–303.
- Kroodsma DE (1983) The ecology of avian vocal learning. *Bioscience* 33:165–171.
- Kus B, Hopp SL, Johnson RR, Brown BT (2010) Bell's Vireo (*Vireo bellii*). In: Poole A (ed) *The birds of North America online*. Cornell lab of ornithology, Ithaca. doi:10.2173/bna.35
- Leitner S, Marshall RC, Leisler B, Catchpole CK (2006) Male song quality, egg size and offspring sex in captive canaries (*Serinus canaria*). *Ethology* 112:554–563. doi: 10.1111/j.1439-0310.2006.01195.x
- Martindale S (1980) A numerical approach to the analysis of Solitary Vireo songs. *Condor* 82:199–211.
- McGregor PK, Krebs JR, Perrins CM (1981) Song repertoires and lifetime reproductive success in the Great Tit (*Parus major*). *Am Nat* 118:149–159.
- Morton ES (1975) Ecological sources of selection on avian sounds. *Am Nat* 109:17–34.
- Mountjoy DJ, Leger DW (2001) Vireo song repertoires and migratory distance: three sexual selection hypotheses fail to explain the correlation. *Behav Ecol* 12:98–102. doi: 10.1093/oxfordjournals.beheco.a000385
- Peters S, Searcy WA, Beecher MD, Nowicki S (2000) Geographic variation in the organization of Song Sparrow repertoires. *Auk* 117:936–942.
- R Core Team (2014) R: a language and environment for statistical computing.
- Read AF, Weary DM (1992) The evolution of bird song: comparative analyses. *Philos Trans R Soc B Biol Sci* 338:165–187. doi: 10.1098/rstb.1992.0137

- Sanchez HM, Vallejo EE, Taylor CE (2015) PajaroLoco: a suite of programs to study the grammatical structure of bird songs. Proceedings of the 20th International Symposium on Artificial Life and Robotics, January 2015.
- Sasahara K, Cody ML, Cohen D, Taylor CE (2012) Structural design principles of complex bird songs: a network-based approach. PLoS One 7:e44436. doi: 10.1371/journal.pone.0044436
- Slabbekoorn H, Smith TB (2002) Habitat-dependent song divergence in the little greenbul: an analysis of environmental selection pressures on acoustic signals. Evolution 56:1849–58.
- Soma M, Garamszegi LZ (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. Behav Ecol 22:363–371. doi: 10.1093/beheco/arq219
- Sosa-López JR, Mennill DJ (2013) The vocal behavior of the Brown-throated Wren (*Troglodytes brunneicollis*): song structure, repertoires, sharing, syntax, and diel variation. J Ornithol 155:435–446. doi: 10.1007/s10336-013-1024-6
- Tan LN, Kaewtip K, Cody ML, Taylor CE, Alwan, A (2012) Evaluation of a sparse representation-based classifier for bird phrase classification under limited data conditions. Interspeech 2012:3–6.
- Tan LN, Alwan A, Kossan G, Cody ML, Taylor CE (2013) A sparse representation-based classifier for in-set bird phrase verification and classification with limited training data. ICASSP 2013:763–767.
- Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. Nature 393:440–2.
- Weiss M, Hultsch H, Adam I, Scharff C, Kipper, S (2014) The use of network analysis to study complex animal communication systems: a study on nightingale song. Proc R Soc B 281: 20140460.
- Wolfram Research Inc. (2012) Mathematica, Version 9.0. Champaign, IL.

## CHAPTER 3

### Complexity, Predictability, and Time Homogeneity of Syntax in the Songs of Cassin's Vireo

(*Vireo cassinii*)

#### 3.1 Abstract

Many species of animals deliver vocalizations in sequences presumed to be governed by internal rules, though the nature and complexity of these syntactical rules have been investigated in relatively few species. Here I present an investigation into the song syntax of fourteen male Cassin's Vireos (*Vireo cassinii*), a species whose song sequences are highly temporally structured. I compare their song sequences to three candidate models of varying levels of complexity – zero-order, first-order and second-order Markov models – and employ novel methods to interpolate between these three models. A variety of analyses, including sequence simulations, Fisher's exact tests, and model likelihood analyses, showed that the songs of this species are too complex to be described by a zero-order or first-order Markov model. The model that best fit the data was intermediate in complexity between a first- and second-order model, though I also present evidence that some transition probabilities are conditioned on up to three preceding phrases. In addition, sequences were shown to be predictable with more than 54% accuracy overall, and predictability was positively correlated with the rate of song delivery. An assessment of the time homogeneity of syntax showed that transition probabilities between phrase types are largely stable over time, but that there was some evidence for modest changes in syntax within and between breeding seasons, a finding that I interpret to represent changes in breeding stage and social context rather than irreversible, secular shifts in syntax over time. These findings constitute a valuable addition to our understanding of bird song syntax in free-

living birds, and will contribute to future attempts to understand the evolutionary importance of bird song syntax in avian communication.

### **3.2 Introduction**

Bird song ranges from very simple to highly variable and complex (Bolhuis and Everaert 2013). The preferred measure of song complexity has traditionally been repertoire size (Byers and Kroodsma 2009; Soma and Garamszegi 2011), but this is just one way to consider the topic. Many species, for example, appear to deliver their songs according to highly structured sequencing rules, or syntax, effectively increasing their apparent complexity while maintaining modest repertoire sizes (Lemon and Chatfield 1973; Howes-Jones 1985; Gil and Slater 2000). Though various studies have presented evidence for this type of non-random syntactic structure in the songs of numerous bird species (Chatfield and Lemon 1970; Lemon and Chatfield 1971; Lemon and Chatfield 1973; Falls and Krebs 1975; Dobson and Lemon 1977; Dobson and Lemon 1979; Borror 1981; Martin 1990; Lemon et al. 1993; Gil and Slater 2000), only a few have attempted to rigorously classify the statistical complexity of songbird syntax (Katahira et al. 2011; Jin and Kozhevnikov 2011; Markowitz et al. 2013), and these have been restricted to captive birds. The paucity of empirical results regarding the nature of bird song syntax restricts our understanding of the functional importance of syntax in communication and the extent of variation possible in this trait. In particular, various authors have suggested that the syntactic rules employed by non-human animals appear to be constrained in their sophistication (Hauser et al. 2002; Berwick et al. 2011; ten Cate and Okanoya 2012), a hypothesis that can only be evaluated with additional analyses from a broader array of species. This paper assesses the syntactic complexity of the songs of a population of free-living Cassin's Vireos, contributing to a



growing understanding of bird song syntax and providing a foundation for future studies of the potential role of syntax in the conveyance of information in this species.

Studies of bird song syntax have typically relied upon Markov models to evaluate their complexity, with the complexity of syntax relating to the complexity of the model that best fits the observed sequences (Chatfield and Lemon 1970; Lemon and Chatfield 1973; Falls and Krebs 1975; Dobson and Lemon 1977; Dobson and Lemon 1979; Lemon et al. 1993; Katahira et al. 2011; Jin and Kozhevnikov 2011). This approach shares similarities with the concept of algorithmic complexity in computer science, where the complexity of a sequence of symbols – in this case, a sequence of bird songs – is analogous to the length of the shortest computer program that can describe it (Li and Vitanyi 1993). Algorithmic complexity is probably too theoretical to be of use for applications in biology, however, because even simple biological systems with uncomplicated rules can produce phenomena whose properties are immensely complex (Wolfram 2002). For this paper, I shall adopt a definition from statistical learning theory, which relates model complexity to the number of independent parameters in the model (James et al. 2013). In the case of Markov models, I consider the complexity of the model to be equivalent to the number of states in the model (Bell et al. 1990). Simple syntax, therefore, produces sequences that can be described by a zero-order or first-order Markov model (Lemon and Chatfield 1971), while sequences generated by more complex syntax can be described by either higher-order Markov models (Chatfield and Lemon 1970; Lemon and Chatfield 1973; Falls and Krebs 1975) or variants thereof, such as Prediction Suffix Trees (Markowitz et al. 2013), hidden Markov models (Katahira et al. 2011), or partially-observable Markov models (Jin and Kozhevnikov 2011). The latter models are characterized by larger numbers of model states

allowing them to model non-adjacent dependencies in sequences, where the identity of upcoming vocalizations depends on more than one preceding vocalization.

Cassin's Vireos sing elaborate sequences of phrases that, when recorded over long periods, show extensive evidence of repeated syntactic patterns (Hedley 2016). The species also displays high levels of interannual site fidelity, non-overlapping territories, and prolific vocal output, making them an ideal species in which to study the sequential complexity of songs. Here I present analyses fitting Markov models to the songs of a wild population of Cassin's Vireos from California, to assess the complexity of the song syntax and to investigate the existence of non-adjacent dependencies in the songs of the species.

### **3.3 Results**

#### **3.3.1 Song Characteristics**

I recorded the songs of fourteen male Cassin's Vireos over two breeding seasons in 2013 and 2014 (see Materials and Methods). Males in this species sing short phrases at varying rates throughout the day. Each phrase can be classified as a phrase type based on its acoustic features (Fig. 3.1), and each individual possesses a repertoire containing an average of 51 phrase types; repertoire sizes and summary statistics of the recording corpora for each individual can be found in table 3.1. Singing bouts are not well defined in this species: although their modal song rate is approximately one phrase every two seconds, they often sing at lower, yet steady, rates. Accordingly, each recording was analyzed as a single sequence, regardless of the durations of silence contained therein. A more detailed description of this species' singing behavior can be found in Hedley (2016).

#### **3.3.2 Simulations**

Sequence simulations are a common technique used to assess the syntactic rules of bird song, since they are a simple method for generating sequences from models with known complexity which can then be compared with the observed sequences derived from a bird's internal song system with unknown complexity (Dobson and Lemon 1979; Kershenbaum and Garland 2015). The methods used here were broadly similar to those of Jin and Kozhevnikov (2011), in that I divided each bird's song corpus into a training set for parametrization and a testing set held out for comparison. I used a zero-order, first-order and second-order Markov model to simulate song sequences, and the resulting sequences and the training set were compared to the testing set, by comparing N-gram distributions and recurrence interval distributions with those of the testing set using  $L_1$ -distances (see Materials and Methods). If the differences, reflected in the  $L_1$ -distances, between the simulated data and the testing set were comparable in magnitude to the differences between the training set and the testing set, this would indicate that the model used to generate the simulations was similar to the syntax employed by the bird itself, or at least that their output shared similar properties.

The simulations suggested that the second-order Markov model closely approximated the singing style for all individuals (Fig. 3.2). At  $N=1$ , the N-gram distribution for all models did not differ from the  $L_1$ -distances for the training set (one-tailed paired t-tests vs training set: zero-order,  $p=0.13$ ; first-order,  $p=0.27$ ; second-order,  $p=0.25$ ). At  $N \geq 2$ , the  $L_1$ -distances for the zero-order model diverged from the expected distribution (one-tailed paired t-tests vs training set,  $p < 0.001$  at  $N = 2$  thru  $N = 7$ ), illustrating that the sequences generated by the zero-order model differed drastically from the sequences upon which they were parametrized. As the N-gram increased, the  $L_1$ -distances for the first-order model diverged from the values for the training set (one-tailed paired t-tests vs training set,  $p=0.28$  at  $N=2$ ,  $p < 0.001$  at  $N = 3$  to  $N = 7$ ), while the

values for the second-order model remained similar to those of the training set until  $N=6$ , only showing a significant difference at  $N=7$  (Fig. 3.2; one-tailed paired t-tests vs training set,  $p=0.27$  at  $N=2$ ,  $p=0.23$  at  $N=3$ ,  $p=0.54$  at  $N=4$ ,  $p=0.20$  at  $N=5$ ,  $p=0.081$  at  $N=6$ ,  $p=0.003$  at  $N=7$ ). The first- and second-order models both produced a similar recurrence interval distribution to that described by Hedley (2016), biased towards low recurrence intervals. Unlike the distributions derived from the zero-order model, these distributions did not differ significantly from the distributions observed in the training sets (one-tailed paired t-tests vs training set: zero-order,  $p<0.001$ ; first-order  $p=0.66$ ; second-order  $p=0.99$ ), suggesting that even models that condition transition probabilities on a short preceding sequence can recover apparent longer-term structure in song sequences (Fig. 3.2).

### **3.3.3 Fisher's Exact Tests for Higher-order Dependencies**

The crucial difference between a first-order Markov process and a second-order (or higher) Markov process is that the second-order Markov process can contain non-adjacent dependencies between phrases (Chatfield and Lemon 1970). In a sequence generated by a first-order process, observed transition probabilities depend solely upon the identity of the ultimate phrase type, while in a sequence generated by a second-order process, transition probabilities may be influenced by the identity of the penultimate phrase type.

To investigate the presence of higher-order dependencies, I used Fisher's exact tests to assess whether the probability of upcoming phrase types was influenced by the identity of the phrase type two or three phrases prior (Fig. 3.3; see Materials and Methods). I found evidence of at least second-order dependencies in the songs of every individual (Table 3.1). On average, 46 Fisher's tests were conducted to investigate second-order dependencies in each individual's song sequences, and an average of 7.6 (16%) of these comparisons met the Bonferroni-corrected

threshold for significance (Table 3.1). Each individual's songs were subjected to a larger number of Fisher's tests for third-order dependencies (mean=160 comparisons). On average, 1.2 (1%) of these comparisons were significant at the Bonferroni-corrected significance threshold, though six individuals showed no evidence of third-order dependencies (Table 3.1). To control for the possibility that higher-order dependencies may arise spuriously, I simulated each bird's recording corpus using a first- and second-order Markov model (see Materials and Methods). Sequences simulated using a first-order Markov model did not show evidence of second-order dependencies at the Bonferroni-corrected significance level, nor did sequences generated using a second-order Markov model show evidence of third-order dependencies, indicating that the significant results for the observed sequence reflect a more complex underlying syntax and could not have arisen spuriously. There was a positive correlation between the proportion of Fisher's tests for third-order dependencies that were significant and the total number of phrases in a bird's song corpus (Pearson's test,  $r=0.63$ ,  $p=0.017$ ), suggesting that insufficient sample sizes may have impeded the detection of higher-order dependencies in the songs of some individuals.

### **3.3.4 Model Likelihood**

I assessed the likelihood that each model could have generated observed sequences, using backoff smoothing and Witten-Bell discounting to account for events that were not observed in the training set (Witten and Bell 1991; Jurafsky and Martin 2000). The best model was that for which the probability of generating the observed sequence was highest, reflected in a low negative log-likelihood value (see Materials and Methods). In addition to the three models used previously, I used forward selection to interpolate between the models (see Materials and Methods), allowing assessments of models with a combination of zero-order, first-order and second-order properties. From these models of intermediate complexity, a model was selected

(hereafter referred to as the ‘interpolated model’) that had the lowest negative log-likelihood on the testing set. This model was then assessed alongside the three original models using both a train-test paradigm and a Leave-one-out cross-validation (LOOCV) paradigm (see Materials and Methods). The inclusion of the interpolated model was based on the fact that the three original models differed significantly in their complexity: the zero-order model contained a single state; the first-order model contained a few dozen; and the second-order contained several hundred observed states. It seems likely that the syntax of any species will not abide strictly to any one of these models, and that the true complexity will often be intermediate between them. The use of forward selection to identify an interpolated model permitted evaluation of hundreds of candidate models with varying numbers of states in a straightforward and principled manner.

For all individuals, the interpolated model best fit the data under the train-test paradigm, though for two birds, the first-order model performed equally well (Table 3.2). Under the LOOCV paradigm, however, the interpolated model showed the best fit for all individuals (Table 3.2). The structure of the interpolated model differed between individuals, but always included a mixture of first- and second-order states (Fig. 3.4). On average, 8% (range: 0-12%) of phrase types were grouped together within the zero-order category, suggesting that they did not show distinct first-order properties, or at least that their properties were not very different from one other. Fifty-two percent (range: 4-76%) of phrase types showed first-order properties, and 40% (range: 17-96%) showed evidence of second-order relationships. The interpolated Markov model included an average of 83 states, 66% more than the number of states in the first-order model, yet 67% fewer than the number of states in the second-order model.

### **3.3.5 Predictability of Sequences**

I examined the predictability of sequences using the four models used above. At each point in a sequence, the model returned the most likely phrase type based on the immediately preceding sequence (see Materials and methods). The zero-order model performed poorly for all individuals under the train-test paradigm, predicting upcoming phrase identity with an accuracy of 8.9%, while the first-order, second-order and interpolated models gave prediction accuracies of 54.2%, 56.1%, and 56.8%, respectively (Table 3.2). Under the LOOCV paradigm, the zero-order, first-order, second-order and interpolated models predicted upcoming phrase types with 8.5%, 55.3%, 57.9% and 58.2% accuracy, respectively (Table 3.2). A repeated measures ANOVA showed significant differences in the accuracy of the latter three models under the train-test paradigm ( $F(2,26)=11.74$ ,  $p<0.001$ ) and the LOOCV paradigm ( $F(2,26)=14.57$ ,  $p<0.001$ ). Under both paradigms, pairwise t-tests with Bonferroni-adjusted p-values determined that the first-order model gave significantly lower prediction accuracy than either the second-order or interpolated model (Train-test: first order vs second order  $p=0.0496$ ; first order vs interpolated  $p=0.003$ . LOOCV: first order vs second order  $p=0.034$ ; first order vs interpolated  $p<0.001$ ), but that the latter two models did not differ significantly in their prediction accuracies (both paradigms:  $p=0.08$ ). Though the magnitudes of these differences are small, this is because the first-order, second-order and interpolated models made identical predictions for 76% of phrases tested. Considering only instances where at least one model differed from the other two under the LOOCV paradigm ( $n=13492$  instances), the first-order, second-order and interpolated models had accuracies of 28.0%, 39.1%, and 40.3%, respectively, accounting for the differences in overall accuracy between the three models. Prediction accuracy depended strongly upon the length of the time interval across which the prediction was taking place: predictions across

intervals shorter than two seconds, representing a high song output, were more than 65% accurate, declining to less than 25% for intervals longer than ten seconds (Fig. 3.5).

### 3.3.6 Time Homogeneity

An assumption central to Markov models is that the process is time homogeneous, meaning that the transition probabilities do not change over time. This requirement is also critical to any attempt to model bird song syntax, as an explicit model of any bird's syntax would be difficult to ascertain if the nature of the syntax were to change from one moment to the next, or from one month to the next. These data were particularly appropriate for investigating the time homogeneity of the songs because recordings were collected throughout the breeding season and, for some individuals, over multiple years.

First, I observed that phrase-type use and transition probabilities appeared to be largely conserved between the training sets and testing sets used above. Phrase types that were common in a training set were also common in the corresponding testing set for each individual (Fig. 3.6a, Pearson's test,  $r(710)=0.80$ ,  $p<0.001$ ). The same held true for bigrams (Fig. 3.6b, Pearson's test,  $r(4374)=0.83$ ,  $p<0.001$ ), suggesting that this may hold for simple syntactic patterns as well as individual phrase types. Furthermore, first-order transition probabilities showed a strong correlation when considering only probabilities that were conditioned on a large number of observations ( $>50$ ) in both the training and testing sets (Fig. 3.6c, Pearson's test,  $r(1534)=0.96$ ,  $p<0.001$ ). Correlations between the training sets and testing sets were not perfect, however, raising the question of whether the disparity between the two sets was a result of gradual drift in repertoire use or syntax over time, or simply represented an artifact of the finite sample sizes of the two datasets. To assess this, I simulated recording corpora for each bird using second-order Markov models, which are, by definition, time homogeneous, and which above were shown to



largely capture the syntactic structure of the songs. I then compared the  $L_1$ -distances of the  $N$ -gram distributions for the 100 simulated training and testing sets to the values for the observed data (see Materials and Methods). The  $L_1$ -distances are equivalent to the sum of the residuals around the line of equality ( $y=x$ ) in Fig. 3.6a ( $N=1$ ) and 3.6b ( $N=2$ ).

I found some evidence that the songs of the species were not completely time homogeneous, although the magnitude of the change in syntax did not appear great. At  $N=1$  and  $N=2$ , the observed  $L_1$ -distances were not different from those expected under a time-homogeneous second-order Markov process (Fig. 3.7;  $p=0.56$  for  $N=1$ ;  $p=0.14$  for  $N=2$ ), indicating that the variability in Fig. 3.6a and 3.6b are within the range expected under a time-homogeneous model. At higher values of  $N$ , however, the observed  $L_1$ -distances were significantly larger than those expected under a time-homogeneous model (Fig. 3.7;  $p<0.05$  at  $N=3$ ;  $p<0.01$  at  $N\geq 4$ ), indicating that differences between the training and testing sets in these metrics were larger than expected under strict time homogeneity.

### **3.4 Discussion**

#### **3.4.1 Syntactic Complexity**

These results illustrate that the songs of Cassin's Vireos are too complex to be modeled with either a zero-order or first-order Markov model. This conclusion was borne out by the sequence simulations, the Fisher's exact tests, and the model likelihood analysis, all of which showed evidence for non-adjacent dependencies in the songs. Though sequences generated by a second-order Markov model appeared to more closely approximate the singing style of the species (Fig. 3.2), the superior fit of the less complex interpolated Markov model in the likelihood analysis suggests that the second-order Markov models may contain unnecessary levels of complexity. The interpolated model did not, however, account for the third-order

dependencies identified by the Fisher's exact tests, so it is probable that some aspects of the syntax were not completely captured by that model.

An important consideration in the interpretation of these results is the influence of sample size on the detection of additional levels of complexity. That there was a significant positive relationship between individual sample sizes and the proportion of third-order dependencies suggests that sample sizes may constrain any attempt at inferring true levels of syntactic complexity in animal vocalizations; although sample sizes were in the thousands in this study, it is possible that with additional recordings, all individuals may have shown evidence of third-order or even higher-order dependencies. It may therefore be improper to draw strong conclusions regarding the upper bounds of syntactic complexity based on a finite recording sample. Attempts to rule out lower levels of complexity, however, are not faced with the same concerns. A conservative conclusion in light of this is that the rules underlying the syntax of Cassin's Vireo songs are, at a minimum, intermediate in complexity between those that can be modeled by a first-order and a second-order Markov model, but that higher levels of complexity cannot be ruled out.

Regardless of the exact level of complexity of the songs, it appears that the syntax is governed by relatively simple rules and that transition probabilities are conditioned on the most recent phrase or phrases. Patterns that manifest themselves over longer time-scales, such as the tendency towards low recurrence intervals (Hedley 2016), appear to be emergent properties of these rules.

The patterns of complexity observed here are not unlike those observed in other species that have been the subject of thorough analyses of syntax. In Bengalese Finch, for example, Katahira et al. (2011) identified second-order dependencies, and Jin and Kozhevnikov (2011)

proposed that a Partially-observable Markov Model (POMMA) best describes the species' song sequences. Canaries have been shown to employ up to sixth-order dependencies, describable by a Prediction Suffix Tree (PST; Markowitz et al. 2013). Common to both the POMMA and PST models is the ability to model higher-order dependencies, such that each phrase type can be associated with more than one model state. This many-to-one mapping between model states and observable output is also a characteristic of second-order Markov models, and is a key feature of the interpolated Markov model that provided the best fit to the Cassin's Vireo sequences.

The similarity in syntactical complexity between these species is particularly noteworthy given the distant evolutionary relationship between Cassin's Vireo, members of the clade Corvoidea, and Bengalese Finch and Canary, of the clade Passerida – clades that have been separated by more than 25 million years of evolution (Barker et al. 2004; Prum et al. 2015). This apparent convergence towards similar syntactic properties suggests this pattern may be a widespread phenomenon in songbirds. Previous authors have posited that many-to-one mapping between internal states and vocal output may be a natural consequence of song development – either through multiple memorization of the same phrase type under different syntactic contexts (Slater 1983; Katahira et al. 2013), or through many-to-one neuronal projections between disparate song nuclei in the brain (Katahira et al. 2007; Katahira et al. 2013). Another possibility is that these similarities reflect an upper limit to the complexity of syntactic rules that can be stored in, and produced by, the brains of songbirds. Such a limit has been proposed by previous authors that have noted the relative simplicity of bird song syntax when compared with that of human language, with particular emphasis on the apparent lack of recursion in bird song (Hauser et al. 2002; Berwick et al. 2011; ten Cate and Okanoya 2012). This hypothesis is difficult to evaluate, however, without a more complete understanding of the evolutionary costs and benefits

associated with complex syntax. Gentner and Hulse (1998) showed that European Starlings produce sequences that are well approximated by a second-order Markov model, and also showed that individuals of that species could differentiate sequences generated by first- and second-order Markov models, but not those generated by second- and third-order models. They proposed that one factor constraining syntax may be the ability of receivers to process additional complexity above some limit, which correspondingly restricts the benefits of evolving more complex song output. Alternatively, song syntax may be limited by neural constraints on the part of the sender, though the neural demands associated with complex song syntax are still not well understood (Jin 2013). Future research should attempt to classify the complexity of syntax in a wider array of species, and to identify the role of complex syntax in communication, with the goal of understanding the evolution of this trait and explaining similarities and differences in singing strategies in songbirds.

### **3.4.2 Predictability of Sequences**

These results showed that, given sufficient knowledge of a bird's syntax, the identity of upcoming phrase types could be estimated with more than 54% accuracy, and more than 65% accuracy when song output was high (Fig. 3.5). Aligning with the likelihood analyses and simulations, the model with the highest accuracy was the interpolated model, though the differences in performance between this model and the first- and second-order models were less than three percentage points. It is doubtful that the differences between these three models are of biological significance, as the results show that a bird attempting to predict the identity of upcoming phrases would do nearly as well having heard a single phrase as if they had heard two or more. More likely to be important from a biological perspective are the high levels of overall determinism in the sequences. During territorial disputes, many species of birds are known to

engage in song matching, in which the songs of one bird will be immediately repeated by the other. This behavior has been proposed to have a role in the evolution of song repertoires (Byers and Kroodsma 2009) and geographic dialects (Podos and Warren 2007). In most species that have been the focus of song matching research, however, birds sing with eventual variety (e.g. Song Sparrows, Stoddard et al. 1992; Beecher et al. 2000), repeating each song type many times consecutively. In such cases, the predictability of songs is high: the next song type in a sequence is likely to be the same as the previous. To match a rival's songs, a male need only repeat the most recently heard song type.

In contrast, species that sing with immediate variety, constantly switching between song types, are faced with a greater challenge when attempting to match songs, especially when the intervals between songs are short, as in the songs of Cassin's Vireo. The ability to predict upcoming events could conceivably facilitate song matching by allowing individuals to anticipate upcoming phrases rather than react to them. Counter-singing interactions between rival males are common across territorial boundaries in this species, and I have observed many interactions that appear to contain significant amounts of phrase-type matching. The increased predictability of songs when singing at high rates may play a role in mediating territorial disputes, which often involve rapid exchange of song. Further investigations into the specific role of syntactic rules during song matching interactions are ongoing and promise to yield further insight into the functional importance of song syntax in this species.

### **3.4.3 Non-Markovian Properties**

**Time Homogeneity.** The songs of Cassin's Vireo appear to be nearly time homogeneous, but not completely so. The observed differences between the training and testing sequences in the 1-gram and 2-gram distributions were within the range of variation expected under a time-

homogeneous second-order Markov model, whereas the higher order N-grams exhibited larger differences than would be expected if the syntax were time homogeneous (Fig. 3.7). In other words, the variability in Fig. 3.6a and 3.6b is within the range expected under a time-homogeneous model, but a similar plot of higher N-grams would show somewhat more variability than expected.

These differences could have resulted from a few factors. First, it is possible that the delivery of songs depends upon external variables, such as the presence of females (Sakata and Brainard 2009), rivals (Vehrencamp et al. 2007), or predators (Langmore and Mulder 1992). As an example of this, I have observed that Cassin's Vireos sing particular phrase types near the nest as the female incubates, an observation that has also been noted in Yellow-throated Vireos (Smith et al. 1978) and is the focus of ongoing research. It follows, then, that recordings made during nesting may be enriched for transitions and N-grams containing these particular phrase types, and impoverished for others. Second, as discussed above, Cassin's Vireos regularly engage in phrase-type matching interactions with neighbors, implying yet another influence of social context on their singing behavior.

Both of these potential social influences on song occur over short time scales, on the order of seconds or minutes. There does not appear to be evidence that the syntax of the species drifts over long periods, as this would likely have led to much larger differences between the training set and testing set, especially in the individuals that were recorded in two breeding seasons (Fig. 3.6 and Fig. 3.7). The observations presented here align with the findings of Warren et al. (2012), who showed that Bengalese Finch syntax is stable over time, but can change in response to external factors. They showed that transition probabilities could be altered using aversive stimuli, but that probabilities returned to baseline values following the

experiment, suggesting that although birds can alter their syntax over short time-scales, their preferred syntax remains consistent throughout their adult life.

The relatively stable nature of the repertoire and syntax over time aligns with current understandings of bird song development in which many species of birds undergo a sensitive period of song acquisition, before their songs crystallize and remain relatively unchanged throughout their adult life (Margoliash 2002), with modifications in adulthood being minor adjustments rather than gross reorganizations of song syntax or repertoire composition (James and Sakata 2014).

**Repetition of Phrase Types.** Previous attempts to model the syntax of birds have often modeled separately the transitions between vocal units and the patterns of repetition of units (Sasahara et al. 2006; Jin and Kozhevnikov 2011; Katahira et al. 2013; Markowitz et al. 2013). For many species, this is critical because the repeat distributions are clearly non-Markovian (Kershenbaum et al. 2014), and appear to be governed by processes distinct from those governing transitions between phrase types. I did not model these two processes separately, justified by the observation that repetitions of phrase types are uncommon in this species, accounting for less than 7% of consecutive phrases. Triplets of the same phrase type were even rarer, accounting for less than 2% of consecutive phrases, while quadruplets accounted for 1% of the observed song corpus. A second-order Markov model captures the probabilities of repetitions up to length three, so the models evaluated here likely approximated the distribution of repeated phrases reasonably well. I did, however, occasionally observe highly anomalous singing behavior characterized by long bouts of repetitions. In one case, an individual repeated the same phrase type 38 times, a sequence that would be infinitesimally improbable under any of the models examined here. The current models may be improved somewhat by treating repetitions

differently from transitions between phrase types (Jin and Kozhevnikov 2011), but given the rarity of these events in this species, it is likely that any improvements attained in doing so would be minor.

These results illustrate a moderate level of complexity in the syntax of Cassin's Vireo songs. The sequences examined show abundant evidence of second-order dependencies, and hint towards a level of complexity somewhere between that of a first-order and second-order Markov model, though there was also some evidence for higher levels of complexity. It is clear that there are ample opportunities for future research, especially regarding the functional significance of a highly structured syntax during song matching and other social contexts.

### **3.5 Materials and Methods**

#### **3.5.1 Recording Techniques and Individual Identification**

All field research was approved by the Institutional Animal Care and Use Committee at the University of California, Los Angeles (protocol number ARC # 2013-041-01). Banding was carried out under federal bird banding permit #23809. I collected recordings between April 25 and June 28, 2013, and between May 5 and June 25, 2014 at a field site on private land in the foothills of the Sierra Nevada Mountains in California, USA (10 S 706584 4262742, datum WGS 84). All recordings were made with a Marantz PMD 661 solid state digital recording unit and a Sennheiser MKH20-P48 microphone with a Telinga parabolic reflector, and recordings were stored as 16 bit WAV files with a sampling rate of 44.1 kHz. I recorded birds opportunistically by approaching a known breeding territory and recording the singing male until he either stopped singing or flew too far away to be recorded. In 2013, males were identified based on their association with known breeding territories, and identifications were then confirmed from recordings based on the observation that birds possess individually distinctive



repertoires that are organized into diagnostic sequences (Arriaga et al. 2014; Arriaga et al. 2016). In May 2014, males were captured and marked with unique colored leg bands, which helped identify the birds during subsequent recordings. Individuals are here referred to by a code representing their color bands, except in three birds that were not captured, referred to as ‘Gate’, ‘Meadow’ and ‘Gully’.

### **3.5.2 Recording Annotation**

Recording annotation methods were the same as those described in Hedley (2016), and many of the recordings analyzed here were also included in that study. All recording annotation was completed using the program Praat (Boersma and Weenink 2014). I visually inspected the spectrogram of each recording, and marked the boundaries of each phrase with approximately +/- 0.01s accuracy. Each phrase was classified as one of 126 phrase types that have been observed in the study population, based on distinctive acoustic characteristics that are readily visible on the spectrogram; phrase types were denoted with unique two-letter codes. A spectrogram showing four phrases with their corresponding annotations is shown in Fig. 3.1.

Although the identification of phrase types is subjective, Tan et al. (2013; 2015) have shown that the categories used here were objectively identifiable by computer algorithms, and a comparison between visual inspection and algorithmic classification showed that the two methods agreed on more than 99% of occasions when tested on recordings with high signal-to-noise ratios (Hedley 2016). Similarly, a comparison of annotations made by two observers on 100 phrases from each of the fourteen individuals in this study showed that the two observers assigned the same phrase type label for 99% of phrases (1386/1400 phrases). The visual inspection method from a single observer was used for all annotations due to potential concerns

about noise-robustness of the classification algorithm and the prohibitive time and personnel requirements of using multiple observers.

In accordance with the methods of Hedley (2016), I discarded all recordings containing fewer than 50 phrases, which amounted to four percent of the total phrases in the original recordings. The resulting annotated dataset comprised 221 recordings containing 57377 phrases from fourteen individuals (mean=4098, range=1335-10336 phrases per individual). All sound recordings are available on Figshare (10.6084/m9.figshare.3081814), along with corresponding textgrid files containing phrase annotations, for use with Praat (Boersma and Weenink 2014).

### 3.5.3 Simulations

**Model Parametrization.** Each bird's recording corpus was divided into a training set and a testing set by concatenating the sequence from all recordings of a given individual in chronological order and identifying the midpoint of the sequence. The recording containing this phrase was placed in the training set along with all preceding recordings, while the remaining recordings were held out as the testing set. As such, each recording was assigned to either the training set or the testing set, never split in two. This procedure for dividing a bird's recording corpus in is referred to as the 'train-test paradigm' throughout this study.

The zero-order Markov model was the simplest model investigated. In this model, the probability of observing a given phrase type in a sequence is independent of the identity of the preceding phrases, and is proportional to its frequency of occurrence in the training data. The probability of each phrase type was calculated as

$$P_i = n_i/N$$

Where  $n_i$  is the number of times phrase type  $i$  was observed in the training set, and  $N$  was the total number of phrases in the training set.

The model with intermediate complexity was a first-order Markov model, in which the probability of observing a given phrase type depends on the identity of the immediately preceding phrase. This model can be visualized as a  $C \times C$  matrix, where each row and column represents a phrase type in the bird's repertoire. The rows represent the identity of the preceding phrase, while the columns represent the identity of the subsequent phrase. The probability of transitioning from phrase type  $i$  to phrase type  $j$  is estimated as

$$P_{ij} = P(j|i) = n_{ij}/n_i.$$

Where  $n_{ij}$  is the number of times phrase type  $j$  followed phrase type  $i$  in the training sequence, and  $n_i$  is the total number of bigrams observed that began with phrase type  $i$ . The value of  $n_i$  was sometimes slightly less than  $n_i$  when phrase type  $i$  was the terminal phrase of a recording. Typically, however, these two values were equal.

The most complex model examined here was a second-order Markov model, in which the probability of observing a given phrase type depends on the identity of the two preceding phrases. This model can be represented as a matrix containing  $C^2$  rows and  $C$  columns, where each contains a unique pair of the  $C$  phrase types in the training set. Transition probabilities are estimated as

$$P_{ijk} = P(k|ij) = n_{ijk}/n_{ij}.$$

Where  $n_{ijk}$  is the number of times the trigram  $ijk$  was observed in the training set, and  $n_{ij}$  is the total number of trigrams beginning with the bigram  $ij$ . Considerations of sample sizes precluded looking at higher order Markov processes.

**Sequence Simulations.** Simulations were designed such that the only difference between a simulated dataset and the training data was the mode of syntax generation: since each training set was comprised of multiple recordings containing varying numbers of phrases, each

simulation contained the same number of simulated recordings with the same numbers of phrases. Furthermore, the first phrase type (or two phrase types, in the case of the second-order model) in each simulated recording was made to be identical to those observed in the training set. Unlike subsequent analyses, transition matrices were not smoothed prior to simulations (see below). One ‘simulation’ therefore was equivalent in file structure, length and initial state to the training set from which it was parameterized. Each individual’s training set was simulated 1000 times using the stochastic random sampling function *sample* in R [53].

**Model Assessment.** The resulting simulations were compared with the testing set in two ways. First, the N-gram distribution of the simulated data was compared to that of the testing data for N-grams between one and seven. The N-gram distribution is the relative frequency of all unigrams (for N=1), bigrams (N=2), trigrams (N=3), etc. in the sequence. The 1-gram distribution was the relative frequency of each phrase type in the sequence, calculated by dividing the number of observed instances of each phrase type by the total length of the sequence. The 2-gram distribution was calculated similarly, by dividing the number of observed instances of each bigram by the total number of bigrams in the sequence. Higher-order N-gram distributions were calculated in a similar fashion. Because each N-gram distribution contained relative frequencies, the entries of each N-gram distribution summed to one. The lack of smoothing prior to the sequence simulations constrained the N-grams that could be observed generated at low values of N: the zero-order model was constrained to the 1-grams in the training set; the first-order model was constrained to the 2-grams in the training set; and the second-order model was constrained to the 3-grams in the training set. Each of these models could generate a greater variety of N-grams at higher values of N, provided they were comprised of lower N-

grams present in the training set. Differences between simulated sequences and observed sequences were therefore expected to be most prominent at higher values of N.

Comparison of N-gram distributions was conducted by taking the  $L_1$ -distance between the two distributions, where  $L_1 - distance = \sum_i |x_i - y_i|$ ,  $x_i$  is the frequency of occurrence of a particular N-gram in one dataset, and  $y_i$  is the frequency of occurrence of that same N-gram in the second dataset. Because this approach measures the difference between two distributions that each sum to one, this measure of similarity can take values from 0 to 2. Two distributions with an  $L_1$ -distance of 0 are identical, those with a distance of 1 show exactly 50% concordance in their probability distributions, and those with a value of 2 do not overlap at all; the  $L_1$ -distance therefore corresponds with the amount of probability mass that must be redistributed to render the two distributions identical. This same analysis was conducted using the sum of squared differences, which more heavily weights outliers, and the results were qualitatively similar;  $L_1$ -distances were used for their ease of interpretation.

The second way I compared simulated sequences with the observed data was through the distribution of recurrence intervals. The recurrence interval was defined as the number of intervening phrases between two occurrences of the same phrase type (Kroodsma 1975). This characteristic was chosen because it evaluates sequential structure across the full length of each recording, and so complements the characteristics assessed by the N-gram distributions, which analyzed the structure across short chunks of song, up to seven phrases in length. Recurrence intervals have been shown to be diagnostic characteristics of bird song for various species. Some species sing with high recurrence intervals, reflecting a tendency to cycle through the repertoire (e.g. Marsh Wren, Verner 1975; Western Meadowlark, Falls and Krebs 1975; Chaffinch, Slater 1983), while other species favor low recurrence intervals, presenting small subsets of their

repertoires at a time (e.g. Rock Wren, Kroodsma 1975). Suitable syntactic models should generate songs with species-typical recurrence intervals.

The songs of Cassin's Vireos favor low recurrence intervals, meaning that after a phrase type is delivered once, it is likely to recur shortly thereafter in a sequence (Hedley 2016). I calculated the observed distribution of recurrence intervals for each simulated sequence and for the training set and testing set, then compared the training set and simulated data to the testing set using the  $L_1$ -distance in the same manner as above.

### **3.5.4 Fisher's Tests for Higher-order Dependencies**

To identify second-order dependencies, I calculated a second-order Markov transition matrix using the complete corpus of each bird's songs. For each phrase type in a bird's repertoire, I compared the probability distributions associated with two states, where the states differed in the identity of the penultimate phrase type but not in the ultimate phrase type. I used a Fisher's exact test to determine whether the observed transition probabilities differed as a function of the identity of the penultimate phrase type while keeping the ultimate phrase type constant (Fig. 3.3). The two states selected for comparison were the two containing the largest number of observations because, despite large overall sample sizes in this study, the number of observations contributing to the estimation of each transition probability was often small. Considerations of sample size are examined further in the discussion. In about 3% of comparisons, sample sizes were too large to calculate exact p-values using Fisher's exact tests, so Chi-square tests for independence were used instead. I used Bonferroni corrections to control for multiple comparisons within each bird's repertoire. The total number of comparisons was often less than the repertoire size because some phrase types were only observed to be associated with a single second-order state, making comparisons unfeasible for that phrase type.

As a point of comparison, I simulated a complete recording corpus for each bird under a first-order Markov model and conducted the same analysis on the simulated data, with the expectation that a sequence generated by a first-order Markov model would show little or no evidence of higher-order dependencies.

I conducted the same analysis to investigate third-order dependencies by calculating the third-order Markov matrix, and comparing states that had the same ultimate and penultimate phrase type, but differed in their antepenultimate (third most recent) phrase type. Similar to above, only the two states with the largest sample sizes for a given pair of ultimate and penultimate phrase types were compared. The total number of comparisons was somewhat less than the total number of bigrams observed in each bird's recording corpus, because some bigrams were only ever preceded by a single antepenultimate phrase type and therefore could not be included in this analysis. Again, I simulated each bird's corpus, this time using a second-order Markov model, and conducted the same analysis on the simulated sequence to rule out the possibility of spurious relationships.

### 3.5.5 Model Likelihood

**Model Evaluation.** Another measure of the fit of a model to unseen data is the negative of the log-likelihood of a sequence given a particular model. Negative log-likelihood is measured by multiplying the probability of observing each phrase type at each point in the testing set, given the hypothesized model of sequence generation parametrized from the training set, then taking the negative of the natural logarithm.

The negative-log-likelihood of the sequence as a whole is given by

$$-\ln L(\text{sequence}|\text{model}) = -\ln \prod_m P(x_m|\text{model})$$

Where  $P(x_m|model)$  is the probability of observing phrase type  $x$  at position  $m$  given the model in question. Probabilities of each observation given each candidate model are given by  $P_i$ ,  $P_{ij}$ , and  $P_{ijk}$  above for the zero-order, first-order and second-order models respectively.

**Forward Selection.** The forward selection method used here was derived from the method outlined in James et al. (2013), and shares similarities with the methods used by Markowitz et al. to construct Prediction Suffix Trees that were used to identify long range dependencies in Canary song (Markowitz et al. 2013). Starting with the zero-order model, I evaluated all possible first-order states that could be added to the model, selecting the state that provided the greatest reduction in negative-log-likelihood when applied to the training set. The model was then evaluated on the testing set. Subsequently, first- or second-order states could be added to the model, with the constraint that a second-order state was only considered if its corresponding first-order state had already been included. The process concluded once all of the second-order states had been added to the model. The ‘interpolated model’ was that which had the lowest negative-log-likelihood when applied to the testing set, and could contain states with any combination of zero-order, first-order and second-order properties.

**Evaluation Paradigms.** The fit of each model to the data was assessed in two ways: first, the model was parametrized on the training set and evaluated on the testing set according to the train-test paradigm described earlier. This method was used to select the interpolated Markov model. The second method used was leave-one-out cross-validation, which I refer to as the ‘LOOCV paradigm’. In this case, the model was iteratively parametrized on all but one recording from a bird’s corpus, and was evaluated on the held out file. This method acted as a check against the possibility of overfitting when using a single training set.



For each testing set, the first two phrases in each recording were not included in the calculation, since the second-order model requires knowledge of the two preceding phrase types. I divided the resulting likelihood value by the total length of the sequence to control for the different lengths of sequences in the recordings of each bird, such that the values presented represent the negative-log-likelihood averaged across each phrase in the sequence.

**Smoothing.** Some of the transitions observed in the testing set were not observed in the training set, and therefore had estimated probabilities of zero, making the likelihood calculation above meaningless. To overcome this issue, I used backoff smoothing to provide non-zero probabilities to each cell in the transition matrix. Backoff smoothing uses information from simpler models to estimate probabilities for unseen observations in a more complex model (Jurafsky and Martin 2000). For example, if the transition *aa-ab-ac* was not observed in the training set, the algorithm uses the probability obtained from the first-order model for the transition *ab-ac* to estimate the probability. If this transition was also not observed, the algorithm ‘backs off’ further to the zero-order model, using the probability of the phrase type *ac* to provide a non-zero estimate. Observed probabilities must be lowered to account for the added weight to the unobserved transitions, and for this I used Witten-Bell discounting, which reduces each probability by the factor  $N/(N + T)$ , where  $N$  is the number of observations upon which the probability was conditioned, and  $T$  is the number of different phrase types that were observed under the given condition (Witten and Bell 1991). The remaining weight, equal to  $T/(N + T)$ , is redistributed amongst the unobserved phrase types according to the backoff algorithm described above.

### 3.5.6 Predictability of Sequences

To assess the ability of each model to predict upcoming phrase types, I used the transition matrices to predict the most likely upcoming phrase type based on recently observed sequences, and assessed the accuracy of each model based on the extent to which the predicted phrase types agreed with those observed in the data. The zero-order model always output the most common phrase type in the training set, independent of the preceding sequence, while the first- and second-order models predicted the most commonly observed phrase type given the most recent and two most recent phrases, respectively. The interpolated model used information from one or two preceding phrase types to make its prediction, depending on whether the preceding sequence pertained to a zero-order, first-order or second-order state in the model.

### **3.5.7 Time Homogeneity**

Previous work assessing the time homogeneity of Markov processes has relied upon dividing observed sequences into multiple subsets and comparing transition matrices derived from each subset using Chi-square tests (De Stavola 1988; Tan and Yilmaz 2002). Such methods would not be suitable here, however, since the transition matrix for Cassin's Vireo songs contained several thousand elements, many of which were rarely or never observed. Instead, I used simulations to evaluate the extent to which the differences between the observed training and testing sets deviated from the differences that would be expected under a time-homogeneous process.

I simulated each bird's recording corpus 100 times using a second-order Markov model. In contrast to previous simulations, the entire recording corpus was used for model parametrization. Recording lengths, initial phrase types, and sample sizes of the original data were maintained as above, and the simulated corpus was subsequently divided into 'simulated training' and 'simulated testing' sets corresponding to the files in the training and testing sets in

the train-test paradigm, such that the simulated and observed sets differed only in their means of generation – the simulated sets by a time-homogeneous second-order Markov process, the observed sets by the birds themselves. As in the earlier simulations, smoothing was not employed here.

For each individual at each N-gram value, I calculated the  $L_1$ -distance between each of the 100 pairs of simulated training and simulated testing sets using the  $L_1$ -distance. This resulted in a distribution of  $L_1$ -distances that would be expected under a time-homogeneous second-order Markov model, with the expectation that a time-inhomogeneous process would show significantly larger  $L_1$ -distances than the simulated data. To test for significance, I calculated the mean and standard deviation of the simulated  $L_1$ -distances for each individual at each N-gram, and calculated corresponding z-scores for their observed  $L_1$ -distances at that N-gram. I then conducted a one sample t-test to examine whether the observed z-scores differed significantly from zero, which would suggest a time-inhomogeneous process.

## List of Figures

**Fig. 3.1.** Spectrogram of four phrases from the songs of the ‘Gully’ individual. The sequence depicted shows the phrase types *ai*, *en*, *ds*, and *ai*. The clear resemblance of the first and last phrases illustrates the high levels of within-type stereotypy in the songs of the species.

**Fig. 3.2.**  $L_1$ -distances for the three candidate models compared with those of the training set. Lines connect results from a given individual for the first-order, second-order, and training results, but were not drawn from the zero-order model for clarity and because this model so clearly diverged from the rest. Values on the left represent the seven N-gram distributions, while the values on the right represent the recurrence interval (RI) distributions.

**Fig. 3.3.** A representation of second-order (a) and third-order (b) dependencies in the songs of the bird ‘AGBk’. Bars represent the probability of observing a given phrase type given the identity of the preceding phrase types indicated below each graph. **a)** ultimate phrase type was *db*, and the probability distribution of upcoming phrase types changed depending on whether the penultimate phrase type was *da* or *cj* (Chi-square test for independence,  $\chi^2(9, N=450)=116.94$ ,  $p<0.0001$ ). **b)** penultimate and ultimate phrase types were *cr* and *fq* in both cases, and the probability distribution differed with the identity of the antepenultimate phrase type, whether *bq* or *cg* (Fisher’s exact test,  $p<0.0001$ ). Sample sizes indicate the number of observations upon which the probabilities were conditioned.

**Fig. 3.4.** Illustration of the methodology and results of forward selection. Black lines show results for the individual ‘AGBk’; other individuals are shown in gray. **(a)** states were added to

the model by selecting the state providing the largest improvement when applied to the training set (dotted line). The supplemented model was then tested on the testing set (solid line), and the method repeated until all second-order states had been added to the model. The best (interpolated) model was that which gave the minimum negative log-likelihood when applied to the testing set (solid circles). **(b)** structure of resulting interpolated models, showing the proportion of a bird's repertoire exhibiting zero-order, first-order and second-order properties.

**Fig. 3.5.** The relationship between the predictability of upcoming phrases and singing rate of Cassin's Vireos (gray bars). Predictions were derived from the 'interpolated' Markov model using the train-test paradigm. Sample sizes for each interval are shown (black circles).

**Fig. 3.6.** Phrase-type use, bigram occurrence, and transition probabilities observed in training sets and corresponding testing sets. Each point represents a phrase type **(a)**, bigram **(b)**, or transition probability **(c)** in a bird's repertoire, and different individuals are represented by different colors. Individuals recorded in both 2013 and 2014 ( $n=6$  individuals) are denoted by circles, while those recorded in only a single year are denoted by triangles. Transition probabilities **(c)** were only calculated for probabilities conditioned upon at least 50 observations. Dotted lines have the equation  $y=x$ .

**Fig. 3.7.** Comparison of  $L_1$ -distances observed (closed circles) to the  $L_1$ -distances expected under a second-order Markov model, which has the property of being time homogeneous (open circles). 'Observed' points were calculated by comparing the training set with the testing set for each individual, while the 'Time Homogeneous' points are average  $L_1$ -distances from 100

simulated training and testing sets for each individual. Significant differences are denoted with asterisks (\*= $p < 0.05$ , \*\*= $p < 0.01$ ).

## List of Tables

**Table 3.1.** Summary of each individual's recording corpus, along with the results from the Fisher's tests for second- and third-order dependencies.

**Table 3.2.** Results from likelihood and predictability analyses. L-0, L-1, L-2, and L-Int represent the average per-phrase negative log-likelihood under the zero-order, first-order, second-order, and interpolated Markov models, respectively. P-0, P-1, P-2 and P-Int indicated the prediction accuracy for the same models. Bolded values highlight the most likely model and the model with the best ability to predict upcoming phrases for each individual under the two evaluation paradigms.

## Figures

Figure 3.1

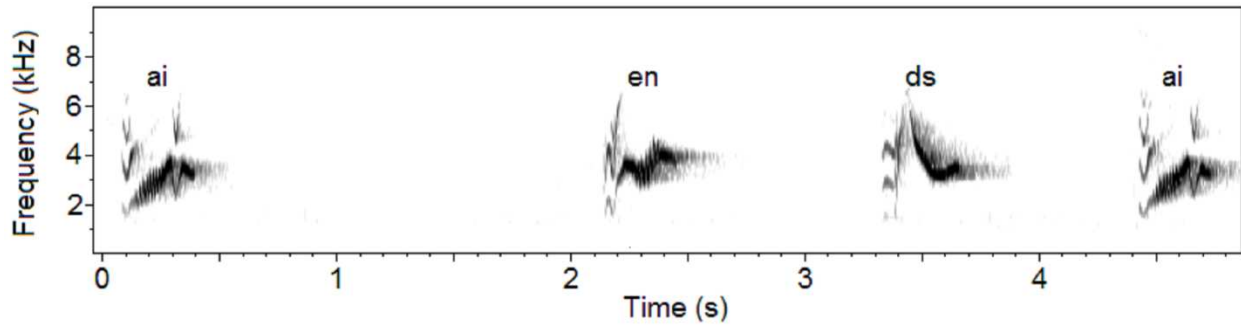


Figure 3.2

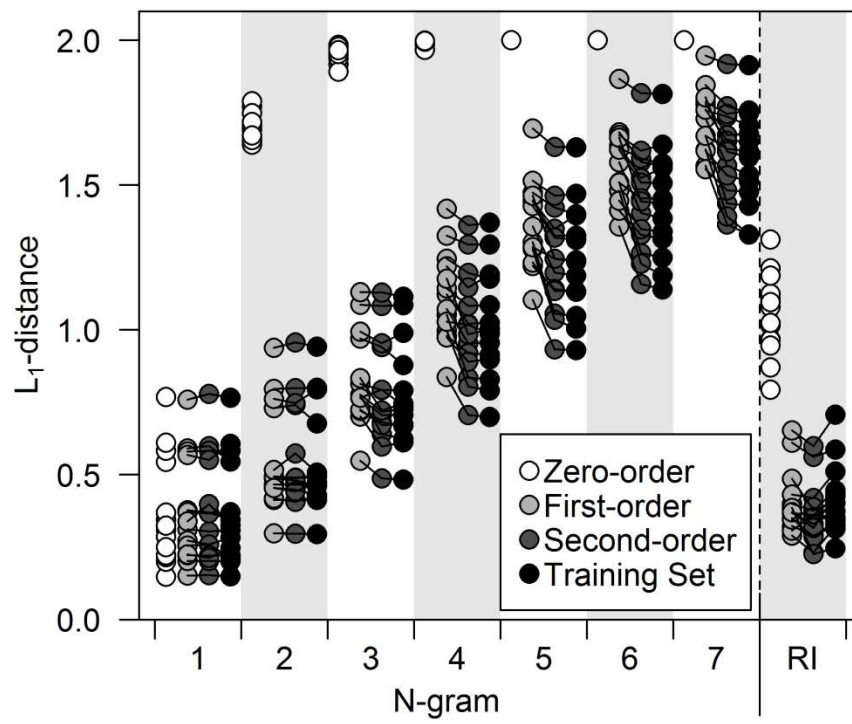




Figure 3.3

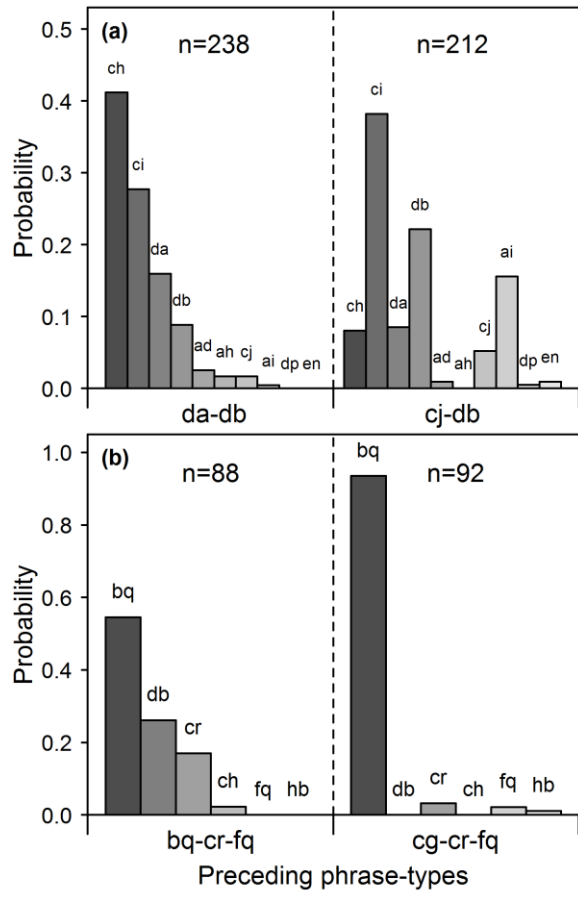


Figure 3.4

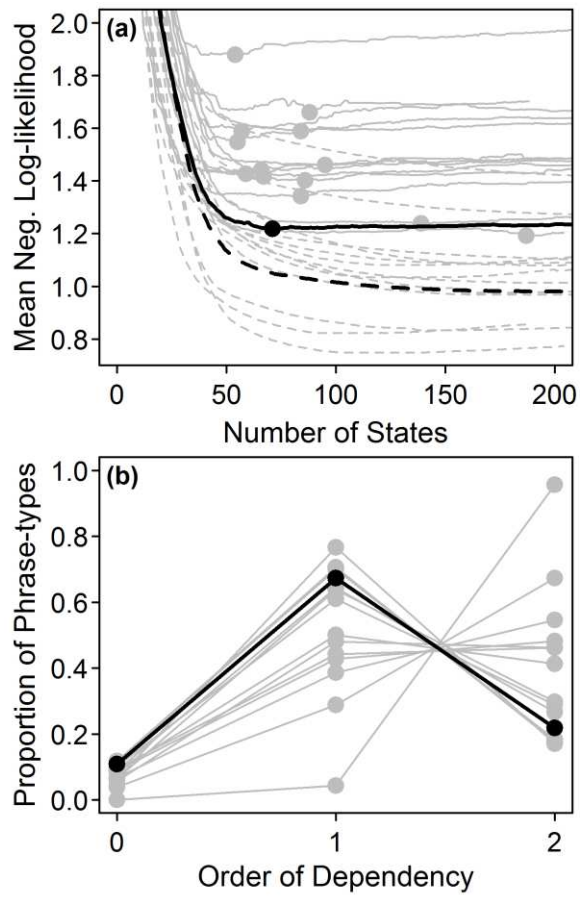


Figure 3.5

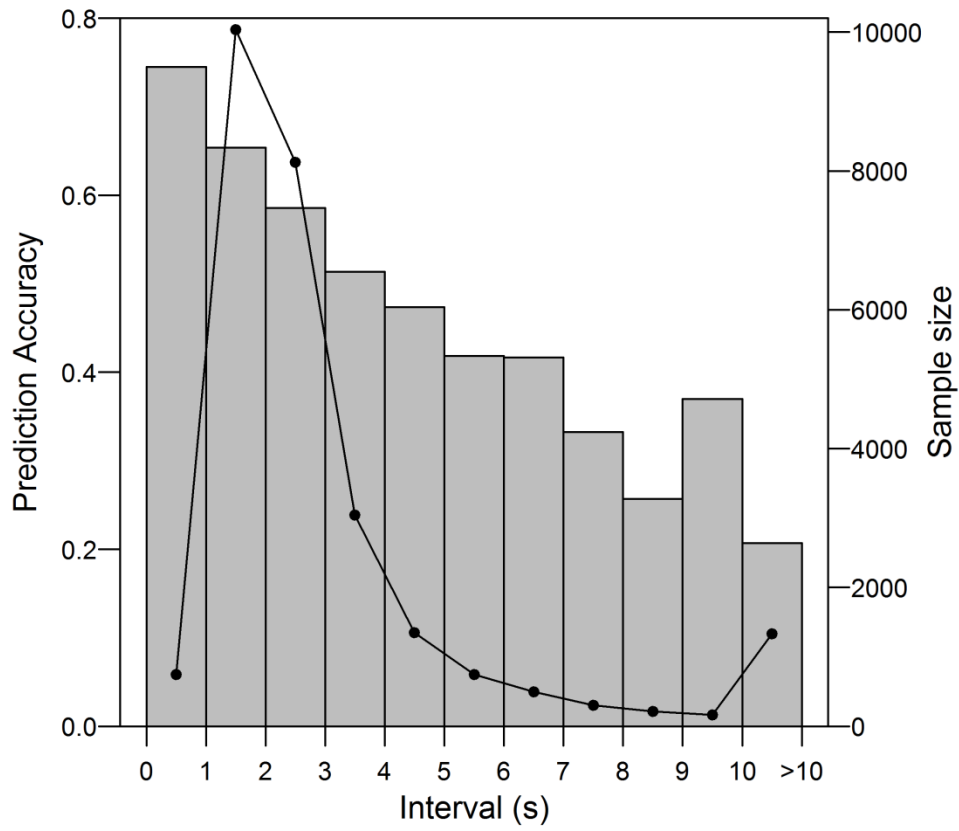


Figure 3.6

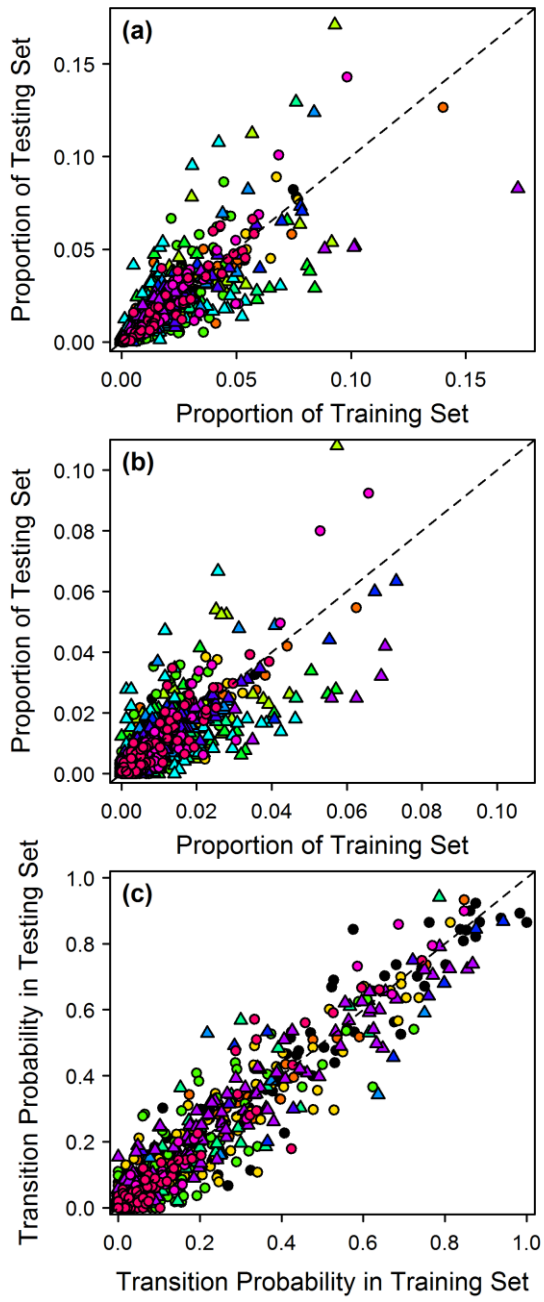


Figure 3.7

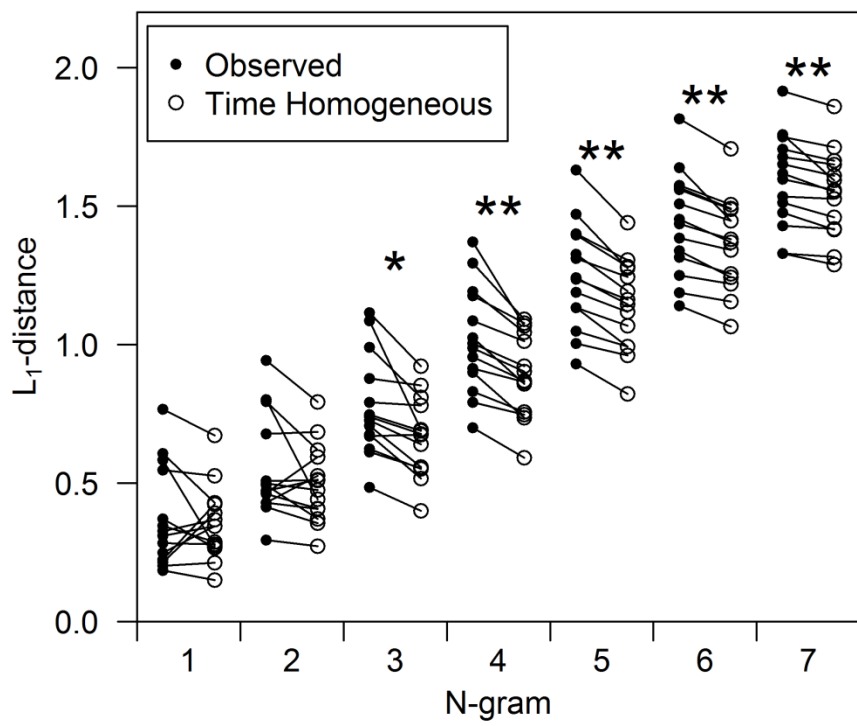


Table 3.1

Individual	No. recordings	Phrases recorded	Repertoire size	Fisher's Exact Tests			
				Order of dependency	No. comparisons (n)	Significant at p<0.05	Significant at p<0.05/n
AGBk	34	10336	55	2	51	25	17
				3	222	22	5
AGO	22	4492	52	2	46	19	11
				3	151	12	1
AOBu	20	6983	59	2	52	16	10
				3	243	14	3
AYO	9	1335	46	2	41	8	2
				3	81	0	0
BuRA	19	4919	51	2	49	18	8
				3	251	8	0
Gate	9	1570	45	2	37	9	1
				3	97	2	0
GRA	11	3918	60	2	55	24	9
				3	166	15	1
Meadow	7	1505	44	2	43	8	3
				3	95	1	0
ORA	11	2262	46	2	34	16	9
				3	89	7	2
RYA	14	3037	47	2	44	15	6
				3	144	4	0
Gully	4	2322	47	2	46	23	6
				3	156	3	0
WABk	30	8588	56	2	54	31	14
				3	235	11	3
YAW	15	3166	52	2	46	11	4
				3	145	7	1
YBuA	16	2944	52	2	48	18	6
				3	165	6	1

Table 3.2

Individual	Method	Sample size*	Likelihood				Prediction Accuracy (%)			
			L-0	L-1	L-2	L-Int	P-0	P-1	P-2	P-Int
AGBk	Train-test	5170	3.67	1.27	1.25	<b>1.22</b>	7.9	58.6	61.1	<b>61.6</b>
	LOOCV	10032	3.69	1.23	1.18	<b>1.16</b>	5.9	58.7	<b>63.2</b>	62.4
AGO	Train-test	2414	3.68	1.34	1.28	<b>1.24</b>	12.7	57.2	<b>63.4</b>	62.7
	LOOCV	4288	3.63	1.30	1.23	<b>1.19</b>	13.4	56.8	63.8	<b>64.1</b>
AOBu	Train-test	3883	3.68	1.62	1.65	<b>1.59</b>	7.7	46.9	45.4	<b>47.5</b>
	LOOCV	6634	3.66	1.56	1.58	<b>1.52</b>	6.1	47.9	49.5	<b>50.1</b>
AYO	Train-test	721	3.34	1.44	1.45	<b>1.42</b>	17.1	56.9	55.9	<b>58.1</b>
	LOOCV	1187	3.50	1.53	1.61	<b>1.51</b>	12.9	56.1	55.3	<b>57.7</b>
BuRA	Train-test	2698	3.53	1.89	2.06	<b>1.88</b>	6.8	40.4	41.0	<b>42.6</b>
	LOOCV	4460	3.59	1.81	1.88	<b>1.78</b>	4.6	41.8	42.7	<b>43.8</b>
Gate	Train-test	915	3.85	1.60	1.76	<b>1.59</b>	5.1	58.6	60.3	<b>61.4</b>
	LOOCV	1396	3.54	1.49	1.58	<b>1.46</b>	4.7	58.1	59.7	<b>60.4</b>
GRA	Train-test	1986	3.82	1.47	1.52	<b>1.44</b>	13.0	53.2	56.0	<b>56.8</b>
	LOOCV	3562	3.80	1.38	1.37	<b>1.35</b>	10.3	56.5	<b>60.1</b>	59.3
Meadow	Train-test	781	3.90	1.69	1.72	<b>1.66</b>	3.1	<b>52.7</b>	51.8	50.3
	LOOCV	1290	3.66	1.37	1.42	<b>1.35</b>	7.4	62.4	60.0	<b>62.6</b>
ORA	Train-test	1252	3.56	1.29	1.21	<b>1.19</b>	12.3	59.7	<b>64.2</b>	64.1
	LOOCV	2056	3.65	1.19	1.11	<b>1.09</b>	10.2	63.4	68.9	<b>69.3</b>
RYA	Train-test	1578	3.64	<b>1.43</b>	1.55	<b>1.43</b>	7.0	59.5	61.9	<b>62.2</b>
	LOOCV	2820	3.64	1.29	1.35	<b>1.27</b>	7.4	64.5	63.7	<b>65.6</b>
Gully	Train-test	1195	3.71	1.58	1.68	<b>1.55</b>	6.3	54.4	55.1	<b>57.0</b>
	LOOCV	1742	3.72	1.54	1.59	<b>1.50</b>	6.1	54.4	56.0	<b>57.8</b>
WABk	Train-test	4714	3.85	1.48	1.53	<b>1.46</b>	8.2	52.4	54.6	<b>54.7</b>
	LOOCV	8302	3.61	1.35	1.40	<b>1.33</b>	13.1	54.1	<b>56.2</b>	55.4
YAW	Train-test	1710	3.39	<b>1.34</b>	1.43	<b>1.34</b>	14.2	<b>62.6</b>	61.9	<b>62.6</b>
	LOOCV	2955	3.48	1.37	1.44	<b>1.36</b>	11.8	60.1	60.4	<b>61.1</b>
YBuA	Train-test	1558	3.54	1.49	1.50	<b>1.40</b>	6.0	57.2	61.9	<b>62.8</b>
	LOOCV	2760	3.59	1.53	1.51	<b>1.43</b>	5.1	57.2	61.9	<b>62.2</b>

\*Sample size for the train-test paradigm was the number of phrases in an individual's training set, while for the

LOOCV paradigm, the sample size was the number of phrases in an individual's total recording corpus minus the average number of phrases in each recording.

## References

- Arriaga J, Sanchez H, Hedley R, et al. (2014) Using song to identify Cassin's Vireo individuals. A comparative study of pattern recognition algorithms. *MCPR*. pp 291–300
- Arriaga JG, Sanchez H, Vallejo EE, et al. (2016) Identification of Cassin's Vireo (*Vireo cassinii*) individuals from their acoustic sequences using an ensemble of learners. *Neurocomputing* 175:966–979.
- Barker FK, Cibois A, Schikler P, et al. (2004) Phylogeny and diversification of the largest avian radiation. *Proc Natl Acad Sci U S A* 101:11040–11045. doi: 10.1073/pnas.0401892101
- Beecher MD, Campbell SE, Burt JM, et al. (2000) Song-type matching between neighbouring song sparrows. *Anim Behav* 59:21–27. doi: 10.1006/anbe.1999.1276
- Bell TC, Cleary JG, Witten IH (1990) Text compression. Prentice-Hall, Upper Saddle River, NJ
- Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ (2011) Songs to syntax: the linguistics of birdsong. *Trends Cogn Sci* 15:113–121. doi: 10.1016/j.tics.2011.01.002
- Boersma P, Weenink D (2014) Praat: doing phonetics by computer. Version 5.3.82. <http://www.praat.org/>
- Bolhuis JJ, Everaert M (2013) *Birdsong, speech, and language: exploring the evolution of mind and brain*. MIT Press, Cambridge, MA
- Borror DJ (1981) The songs and singing behavior of the Red-eyed Vireo. *Condor* 83:217–228.
- Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. *Anim Behav* 77:13–22. doi: 10.1016/j.anbehav.2008.10.003
- Chatfield C, Lemon RE (1970) Analysing sequences of behavioural events. *J Theor Biol* 29:427–445. doi: 10.1016/0022-5193(70)90107-4
- De Stavola BL (1988) Testing departures from time homogeneity in multistate Markov processes. *J R Stat Soc* 37:242–250.
- Dobson C, Lemon R (1979) Markov sequences in songs of American thrushes. *Behaviour* 68:86–105.
- Dobson CW, Lemon RE (1977) Markovian versus rhomboidal patterning in the song of Swainson's thrush. *Behaviour* 62:277–297.
- Falls JB, Krebs JR (1975) Sequence of songs in repertoires of Western Meadowlarks (*Sturnella neglecta*). *Can J Zool* 53:1165–1178. doi: 10.1139/z75-135



- Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Anim Behav* 56:579–594. doi: 10.1006/anbe.1998.0810
- Gil D, Slater PJB (2000) Song organisation and singing patterns of the Willow Warbler, *Phylloscopus trochilus*. *Behaviour* 137:759–782.
- Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298:1569–79. doi: 10.1126/science.298.5598.1569
- Hedley RW (2016) Composition and sequential organization of song repertoires in Cassin’s Vireo (*Vireo cassinii*). *J Ornithol* 157:13–22. doi: 10.1007/s10336-015-1238-x
- Howes-Jones D (1985) The complex song of the Warbling Vireo. *Can J Zool* 63:2756–2766.
- James G, Witten D, Hastie T, Tibshirani R (2013) *An Introduction to Statistical Learning*. doi: 10.1007/978-1-4614-7138-7
- James LS, Sakata JT (2014) Vocal motor changes beyond the sensitive period for song plasticity. *J Neurophysiol* 112:2040–2052. doi: 10.1152/jn.00217.2014
- Jin DZ (2013) The neural basis of birdsong syntax. *Prog. Cogn. Sci. from Cell. Mech. to Comput. Theor.* 1951:
- Jin DZ, Kozhevnikov AA (2011) A compact statistical model of the song syntax in Bengalese finch. *PLoS Comput Biol* 7:e1001108. doi: 10.1371/journal.pcbi.1001108
- Jurafsky D, Martin JH (2000) *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. doi: 10.1162/089120100750105975
- Katahira K, Okanoya K, Okada M (2007) A neural network model for generating complex birdsong syntax. *Biol Cybern* 97:441–448. doi: 10.1007/s00422-007-0184-y
- Katahira K, Suzuki K, Kagawa H, Okanoya K (2013) A simple explanation for the evolution of complex song syntax in Bengalese finches. *Biol Lett* 9:20130842. doi: 10.1098/rsbl.2013.0842
- Katahira K, Suzuki K, Okanoya K, Okada M (2011) Complex sequencing rules of birdsong can be explained by simple hidden Markov processes. *PLoS One* 6:e24516. doi: 10.1371/journal.pone.0024516
- Kershenbaum A, Bowles AE, Freeberg TM, et al. (2014) Animal vocal sequences : not the Markov chains we thought they were. *Proc R Soc B* 281:20141370.
- Kershenbaum A, Garland EC (2015) Quantifying similarity in animal vocal sequences: which metric performs best? *Methods Ecol Evol* 6:1452–1461. doi: 10.1111/2041-210X.12433

- Kroodsma DE (1975) Song patterning in the Rock Wren. *Condor* 77:294–303.
- Langmore NE, Mulder RA (1992) A novel context for bird song: Predator calls prompt male singing in the kleptogamous superb fairy-wren, *Malurus cyaneus*. *Ethology* 90:143–153. doi: 10.1111/j.1439-0310.1992.tb00828.x
- Lemon RE, Chatfield C (1973) Organization of song of Rose-breasted Grosbeaks. *Anim Behav* 21:28–44.
- Lemon RE, Chatfield C (1971) Organization of song in cardinals. *Anim Behav* 19:1–17. doi: 10.1016/S0003-3472(71)80129-X
- Lemon RE, Dobson CW, Clifton PG (1993) Songs of American Redstarts (*Setophaga ruticilla*): sequencing rules and their relationships to repertoire size. *Ethology* 93:198–210. doi: 10.1111/j.1439-0310.1993.tb00989.x
- Li M, Vitanyi P (1993) An introduction to Kolmogorov complexity and its applications. Springer, New York, NY
- Margoliash D (2002) Evaluating theories of bird song learning: Implications for future directions. *J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol* 188:851–866. doi: 10.1007/s00359-002-0351-5
- Markowitz JE, Ivie E, Kligler L, Gardner TJ (2013) Long-range order in canary song. *PLoS Comput Biol* 9:e1003052. doi: 10.1371/journal.pcbi.1003052
- Martin D (1990) Songs of the Fox Sparrow. III. Ordering of song. *Wilson Bull* 102:655–671.
- Podos J, Warren PS (2007) The evolution of geographic variation in birdsong. *Adv Study Behav* 37:403–458. doi: 10.1016/S0065-3454(07)37009-5
- Prum RO, Berv JS, Dornburg A, et al. (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573. doi: 10.1038/nature15697
- R Core Team (2014) R: a language and environment for statistical computing.
- Sakata JT, Brainard MS (2009) Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *J Neurophysiol* 102:2485–2497. doi: 10.1152/jn.00340.2009
- Sasahara K, Kakishita Y, Nishino T, et al. (2006) A reversible automata approach to modeling birdsongs. 2006 15th Int Conf Comput 80–85. doi: 10.1109/CIC.2006.14
- Slater PJB (1983) Sequences of song in Chaffinches. *Anim Behav* 31:272–281.

- Smith WJ, Pawlukiewicz J, Smith ST (1978) Kinds of activities correlated with singing patterns of the Yellow-throated Vireo. *Anim Behav* 26:862–884. doi: 10.1016/0003-3472(78)90153-7
- Soma M, Garamszegi LZ (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. *Behav Ecol* 22:363–371. doi: 10.1093/beheco/arq219
- Stoddard PK, Beecher MD, Campbell SE, Horning CL (1992) Song-type matching in the Song Sparrow. *Can J Zool* 70:1440–1444.
- Tan B, Yilmaz K (2002) Markov chain test for time dependence and homogeneity: An analytical and empirical evaluation. *Eur J Oper Res* 137:524–543. doi: 10.1016/S0377-2217(01)00081-9
- Tan LN, Alwan A, Kossan G, et al. (2015) Dynamic time warping and sparse representation classification for birdsong phrase classification using limited training data. *J Acoust Soc Am* 137:1069–1080. doi: 10.1121/1.4906168
- Tan LN, Kossan G, Cody ML, et al. (2013) A sparse representation-based classifier for in-set bird phrase verification and classification with limited training data. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* pp 763–767
- ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philos Trans R Soc B Biol Sci* 367:1984–1994. doi: 10.1098/rstb.2012.0055
- Vehrencamp SL, Hall ML, Bohman ER, et al. (2007) Song matching, overlapping, and switching in the Banded Wren: the sender's perspective. *Behav Ecol* 18:849–859. doi: 10.1093/beheco/arm054
- Verner J (1975) Complex song repertoire of male long-billed Marsh Wrens in eastern Washington. *Living Bird* 14:263–300.
- Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32:15414–15425. doi: 10.1523/JNEUROSCI.1254-12.2012
- Witten IH, Bell TC (1991) The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Trans Inf Theory* 37:1085–1094. doi: 10.1109/18.87000
- Wolfram S (2002) *A new kind of science*. Wolfram Media, Champaign, IL

## CHAPTER 4

### Song Mirroring in Birds: Evidence from a Songbird with a Complex Song

#### 4.1 Abstract

Many species of songbirds possess repertoires of discrete song types, and establishing the roles of these repertoires in counter-singing interactions is central to our understanding of the evolution of bird song. Species can be classified into two categories: eventual variety singers, which repeat each song type many times before introducing another, and immediate variety singers, which rarely repeat song types consecutively. Song matching, a behavior involving exchange of the same song type between two singing individuals, appears to be more common in eventual than immediate variety singers. Another behavior, that we call *song advancing*, involves a bird responding to a rival's song with the next song type in the bird's own preferred sequence, and has only been identified in immediate variety singers. We propose that song matching and song advancing are underpinned by a common mechanism, called *song mirroring*, where upon perceiving a rival's song type, birds progress sequentially as though they had delivered the perceived song type themselves. Song matching may therefore naturally result from an eventual variety mode of singing, while song advancing is an equivalent behavior in immediate variety singers. We conducted playback experiments on Cassin's Vireo, an immediate variety singer, to test this hypothesis and evaluate the role of song syntax in male-male interactions. Subjects responded with song advancing responses, consistent with song mirroring, and this behavior became more common as they approached the speaker. We discuss the implications of this hypothesis and these results for the evolution of bird song and aggressive signaling in birds.

#### 4.2 Introduction

Counter-singing interactions between neighboring songbirds facilitate the establishment and maintenance of territory boundaries (Yasukawa 1981). In species that possess repertoires of discrete song types, such interactions can involve rapid exchange of information, where individuals use their repertoires to signal their willingness to escalate an encounter in seemingly sophisticated ways (Burt et al. 2001; Burt et al. 2002; Molles 2006; Searcy and Beecher 2009). One way of communicating aggressive intentions is through song matching, where, upon hearing a rival's song type, a bird responds with that same song type (Krebs et al. 1981; Beecher et al. 2000). Evidence suggests that matching provides a highly specific signal that can be directed unambiguously at a rival (Vehrencamp 2001), and that this behavior is central to the vocal communication systems of many songbird species.

Though song matching has been widely documented, Whitney (1991) observed that it appears to be more prevalent in species that sing with eventual variety, where song types are arranged into bouts of repeated utterances, than in species that sing with immediate variety, where song types are rarely repeated consecutively. In free-living Wood Thrushes, which sing with immediate variety, he demonstrated that subjects avoided matching playback song types (Whitney 1991). Like other immediate variety singers, Wood Thrush deliver their songs in non-random sequences, and experiments in captivity revealed that instead of matching, they responded to a playback song type by singing the next song type in their own preferred sequence (Whitney 1985). This behavior, which we refer to as *song advancing* (Fig. 4.1b), has been demonstrated in captive Marsh Wrens (Kroodsma 1979), Common Nightingales (Todt 1971), and Common Blackbirds (Todt 1981), all species that sing with immediate variety, but has not been demonstrated in eventual variety singers.

A unifying explanation for these results can be reached by invoking an alternative interpretation of both song matching and song advancing. If, upon hearing a rival's song type, the receiver proceeds as though they had delivered that song type themselves, the discrepancy between the two singing modes is explained: in species that sing with eventual variety, the receiver would repeat the perceived song type (song matching), and in species that sing with immediate variety, the receiver would advance onward through their preferred progression of song types (song advancing).

In line with this perspective, we propose that song matching and song advancing are subsets of a broader phenomenon that we call *song mirroring*, underpinned by a common proximate mechanism characterized by a blurring of one's own songs and those originating from a rival. This hypothesis, hereafter referred to as the Song Mirroring Hypothesis (SMH), predicts that song matching and song advancing should share similarities from the perspective of the sequential arrangement of songs. Specifically, the tendency to engage in matching or advancing should correspond to the transition probabilities governing song sequences: species with low rates of song type switching (i.e. eventual variety singers) should engage in matching, and species with high rates of switching (i.e. immediate variety singers) should engage in song advancing (Fig. 4.1). Furthermore, in species with high entropy in their transition probabilities, song advancing will not comprise a single outcome, but several alternate outcomes whose rates of occurrence will correspond with their transition probabilities during solo singing. Given the relatively few examples of song advancing in the literature, the function of this behavior remains unclear. It is plausible, however, that song advancing serves a similar purpose to song matching, either conveying aggression (Vehrencamp 2001) or facilitating comparisons between the songs of the two birds by an eavesdropping third party (Logue and Forstmeier 2008).

This paper aims to test the SMH in Cassin's Vireo (*Vireo cassinii*), an immediate variety singer. When examining song sequences in this species, two phenomena become apparent. First, it is clear that phrase types (equivalent to song types in other species) are not delivered in random order, but in clusters, in such a way that certain phrase types consistently appear together in sequences (Fig. 4.2a). These clusters have been shown to be stable over time with respect to the phrase types contained therein (Hedley 2016a). Second, cluster composition appears to be shared between individuals (Fig. 4.2b). That is, not only do neighbors overlap in their song repertoires (Hedley 2016b), their sequencing rules appear to be shared as well.

When two individuals interact across territorial boundaries, they often sing roughly parallel sequences (Fig. 4.2c). We hypothesized that these bouts of parallel singing are facilitated by the shared syntactic patterns described above. For this to be effective, however, a mechanism must exist to synchronize the sequences. Either song matching or song advancing may suffice for this purpose; the SMH predicts that song advancing should predominate in this species, where switching rates are high (~93%, Hedley 2016b). We conducted playbacks of sequences of phrase types with typical or altered sequencing to test the SMH and probe the mechanisms by which this species interacts during counter-singing.

## **4.3 Results**

### **4.3.1 Literature Review**

We identified 20 species where song matching has been shown to occur above chance levels. These species were members of ten different families, some of which have been separated by over 25 million years of evolution (e.g. Paridae vs. Dasyornithidae, Prum et al. 2015). All are members of the Oscine songbirds, typified by a tendency to learn their songs; to our knowledge song matching has not been demonstrated in Suboscine songbirds or other bird species. Of

species that engage in song matching, fourteen sing with eventual variety, and six with immediate variety.

Song advancing has been documented in four species in three families. All of these species sing with immediate variety. A summary of species that engage in song matching and song advancing is presented in Table 4.1. The paucity of published examples of song advancing relative to song matching is not unexpected, given the ease with which song matching can be detected and studied relative to song advancing (see Discussion). Because of this potential bias, the relative importance of these two behaviors across taxa cannot be evaluated without concerted efforts to document song advancing in a greater variety of species.

#### **4.3.2 Playback Experiments**

To examine the role of structured and shared syntax in counter-singing interactions, and to test the predictions of the SMH, we presented eleven male Cassin's Vireos with paired playback trials on different days, where the two treatments differed in being assembled with either *typical* or *atypical* sequencing of phrase types (See Materials and Methods and Fig 4.2). Each trial consisted of playback of 125 phrases of five phrase types over five minutes.

#### **4.3.3 Physical Response to Playback**

All individuals exhibited a strong physical response, approaching closer than 18 meters of the speaker in every trial (Fig 4.3), and responded vocally with an average of 111 phrases (range=12 to 145) per five-minute playback interval. There were no clear indications of differences in the physical response to the trials with typical phrase type sequencing patterns compared to those with atypical sequencing. The minimum distance from the subject to the speaker during the five-minute trials did not differ between the two treatments (Wilcoxon signed-rank test,  $p=1$ ), nor did the amount of time spent within ten meters of the speaker



(Wilcoxon signed-rank test,  $p=0.16$ ), the latency to approach within ten meters of the speaker (Wilcoxon signed-rank test,  $p=0.80$ ), or the number of changes in position, either towards or away from the speaker (Wilcoxon signed-rank test,  $p=0.80$ ).

#### 4.3.4 Matching or Advancing Above Chance Levels

For all analyses presented here, we employed a logistic regression model that took into consideration each bird's known syntax determined from a held out *training set* of recordings made under non-experimental conditions (see Materials and Methods). This model incorporates Markov transition probabilities to estimate the probability of matching (or advancing) by chance. Chance probabilities varied throughout the experiment as a function of the most recent phrase type delivered by the bird, so the model effectively controlled for the influence of syntax on these probabilities.

To investigate whether birds matched or advanced above chance levels, we ran a logistic regression model including only the chance probability plus an intercept (see Materials and Methods). In this case, the intercept,  $\alpha_0$ , reflected the tendency to match (or advance) relative to chance levels by all birds across all trials. The model was run seven times. The first run investigated whether immediate matching (IM) occurred above chance levels, where an immediate match was defined as a phrase type that matched the most recent playback phrase type. The second run examined the tendency to respond with a delayed match (DM), defined as a phrase type that matched any of the previous playback phrase types. Runs three through seven investigated song advancing: the third tested whether subjects responded more than expected by chance with the phrase type that most commonly followed the playback phrase type in that individual's training set ( $SA_1$ ), the fourth tested whether they advanced to their second ranked subsequent phrase type ( $SA_2$ ), and so on until the seventh run, which tested whether they

responded with the phrase type that was the fifth most common successor to the playback phrase type in the training set (SA<sub>5</sub>).

We found no evidence that subjects engaged in immediate matching (Fig. 4.4a,  $\alpha_0 = -0.026 \pm 0.20$  SD,  $p(\alpha_0 > 0) = 0.46$ ) or delayed matching (Fig. 4.4a,  $\alpha_0 = 0.10 \pm 0.079$  SD,  $p(\alpha_0 > 0) = 0.90$ ) at greater than chance levels across treatments. Instead, we found strong evidence of song advancing, where subjects responded to playbacks by advancing ahead from the playback phrase type to the most common subsequent phrase type in their training sets (Fig. 4.4a,  $\alpha_0 = 1.14 \pm 0.14$  SD,  $p(\alpha_0 > 0) \geq 0.9999$  for SA<sub>1</sub>). Subjects also advanced to their second and third ranked phrase types at reduced, but still significant, levels (Fig. 4.4a,  $\alpha_0 = 0.48 \pm 0.16$  SD,  $p(\alpha_0 > 0) = 0.9985$  for SA<sub>2</sub> and  $\alpha_0 = 0.50 \pm 0.15$  SD,  $p(\alpha_0 > 0) = 0.9992$  for SA<sub>3</sub>), but did not advance to their fourth or fifth preferred phrase types above chance levels (Fig. 4.4a,  $\alpha_0 = 0.017 \pm 0.18$  SD,  $p(\alpha_0 > 0) = 0.54$  for SA<sub>4</sub> and  $\alpha_0 = -0.31 \pm 0.23$  SD,  $p(\alpha_0 > 0) = 0.084$  for SA<sub>5</sub>).

These results support the SMH in this species. Under the SMH, birds should respond to a perceived phrase type as if they sang it themselves. Suppose, however, that a bird follows phrase type A with types B, C or D, with probabilities 60%, 20%, and 10% respectively, leaving a 10% chance of transitioning to any other phrase type. The SMH predicts that the bird's response should be to advance from the perceived phrase type A to any of these three phrase types in the order of preference of B>C>D, and that phrase types other than these three should occur at approximately chance levels. In our analysis, the decline in coefficient estimates from SA<sub>1</sub> to SA<sub>5</sub> roughly paralleled the transition probabilities for the first through fifth ranked subsequent phrase types, where median probabilities in our population were 58%, 20%, 8%, 4%, and 2%, respectively (Fig. 4.5).

Given the concordance between the coefficient estimates and transition probabilities, we expect that song advancing in response to a playback phrase type to the first, second, or third ranked subsequent phrase types are probably not three separate behaviors. Instead, they are likely to be underpinned by song mirroring, where upon hearing a given phrase type, the recipient progresses from that phrase type onward as they would in typical song, which in this species involves progressing to one of a few options, each with some associated probability. For this reason, we hereafter combine these three responses into a single behavior, song mirroring, which is here characterized by responding to a phrase type with any of the three most common phrase types that would follow it in typical song. Subsequent analyses will only present results for immediate matching (IM), delayed matching (DM), and song mirroring (SM).

#### 4.3.5 Treatment Effects

To examine whether singing behavior differed in response to playbacks with typical and atypical phrase sequencing, we ran models that included the chance probability plus an intercept ( $\alpha_0$ ) and a treatment coefficient ( $\alpha_{imt}$ ). Here, the intercept represented the tendency to match (or advance) above chance levels in the atypical trials, and the treatment coefficient reflected the magnitude of the difference between atypical and typical trials.

Immediate matching and song mirroring showed no significant treatment effects (Fig. 4.4b,  $\alpha_{imt} = -0.31 \pm 0.40$  SD,  $p(\alpha_{imt} > 0) = 0.21$  for IM and  $\alpha_{imt} = -0.12 \pm 0.21$  SD,  $p(\alpha_{imt} > 0) = 0.28$  for SM), but delayed matching showed a significant positive treatment coefficient (Fig. 4.4b,  $\alpha_{imt} = 0.34 \pm 0.16$  SD,  $p(\alpha_{imt} > 0) = 0.98$ ), indicating a stronger tendency to engage in the behavior in typical than atypical trials.

The increased occurrence of delayed matching in the typical trials can be attributed to the fact that the phrase types in these trials normally occur together in sequence – song mirroring in

response to any of these phrase types will likely lead the bird to deliver another of these phrase types (i.e. a delayed match), a result that would not be true in the atypical trials on account of the aberrant sequencing of phrase types in those trials. In support of this, 71 of 129 instances of song mirroring coincided with delayed matches in typical trials, while just six of 114 instances of song mirroring coincided with delayed matches in the atypical trials. The two treatments differed by a total of 36 instances of delayed matching, so the entirety of the treatment effect can be explained as a peripheral consequence of song mirroring.

Given that song mirroring involves responding with one of the phrase types most likely to follow the stimulus phrase in a preferred sequence, a reasonable hypothesis regarding this behavior is that it might serve to pre-empt an upcoming phrase type that has yet to be delivered. Pre-emption of a rival, however, would only be expected to be successful if the two participants deliver their songs in similar orders. To test this, we investigated the rate of these *pre-emptive matches* of the upcoming playback phrase type when the subjects engaged in song mirroring versus when they did not (Fig. 4.6a). Each phrase type was classified as either an instance of song mirroring or not, as above. Not mutually exclusively, each phrase type was also classified as either a pre-emptive match if it was the same phrase type as the playback phrase that followed it, or not.

Across all trials, when a subject engaged in song mirroring, pre-emptive matches occurred at a rate of 10.1% (24/237) compared with just 1.5% (13/872) when the subject did not engage in song mirroring, a highly statistically significant difference (Fisher's Exact Test, 24/237 vs 13/872,  $p < 0.0001$ ). Twenty-three of the 24 pre-emptive matches that occurred in conjunction with song mirroring occurred in the trials with typical sequencing (23 anticipatory matches out of 125 instances of song mirroring in typical trials). This was a significantly greater rate of pre-

emptive matching than in atypical trials, where just one pre-emptive match resulted from 112 instances of song mirroring (Fig. 4.6b; Fisher's exact test, 23/125 vs 1/112,  $p < 0.0001$ ).

To summarize, engaging in song mirroring appears to often lead birds to sing the upcoming playback phrase type before it has emanated from the speaker, but song mirroring only leads to pre-emption when the playback phrase types are organized according to locally typical sequencing. When playback phrase types are arranged according to locally atypical sequencing patterns, song mirroring still occurs, but does not lead to pre-emption of the playback phrase type (Fisher's exact test, 1/112 pre-emption rate when song mirroring occurred in atypical trials vs 13/463 pre-emption rate when song mirroring did not occur in atypical trials,  $p = 0.31$ ).

#### **4.3.6 Effects of Distance on Song Mirroring**

To investigate whether rates of matching or mirroring changed as a function of the subject's distance from the playback speaker, we ran models comprised of the chance probabilities plus an intercept ( $\alpha_0$ ) and a distance coefficient ( $\alpha_{dist}$ ). The intercept reflected the tendency to match (or mirror) relative to chance levels when distance=0, and the distance coefficient reflected the per-meter change in this tendency as distance increased.

Neither immediate nor delayed matching were significantly associated with the distance from the speaker ( $\alpha_{dist} = -0.0045 \pm 0.019$  SD,  $p(\alpha_{dist} > 0) = 0.41$  for IM and  $\alpha_{dist} = -0.003 \pm 0.009$  SD,  $p(\alpha_{dist} > 0) = 0.36$  for DM). Song mirroring was negatively associated with the subject's distance from the playback ( $\alpha_{dist} = -0.026 \pm 0.011$  SD,  $p(\alpha_{dist} > 0) = 0.01$ ). Our model predicted that subjects engaged in song mirroring at chance levels when far (>40m) from the speaker, increasing steadily as the bird approached (Fig. 4.6c).

## **4.4 Discussion**

### **4.4.1 Evidence for the Song Mirroring Hypothesis**

Cassin's Vireos, which sing with immediate variety, did not song match above chance levels, but instead engaged in song advancing, providing support for the SMH in this species. In other species where song advancing has been described, it has been presented as though the birds advance to a single subsequent element in their preferred sequence (Kroodsma 1979; Whitney 1985). We showed that Cassin's Vireos may respond to a perceived phrase type by advancing to any of the three phrase types that would typically follow the stimulus phrase in their normal song sequences, and that the tendency to do so reflects the Markov properties of their syntax (Fig. 4.5). The most natural interpretation of this is that song advancing is underpinned by song mirroring, where their vocal response corresponds to the response they would give had they delivered the stimulus phrase type themselves. In Cassin's Vireo, this may constitute advancing to any of a few phrase types in their sizable repertoire, as is the tendency during solo singing.

We believe that song mirroring, coupled with shared syntactic patterns, explains the tendency for this species to sing in parallel under natural conditions. Song mirroring accomplishes this in two ways: by increasing the rate of delayed matches (Fig. 4.4b), leading to like phrase types occurring in close temporal proximity, but not adjacently; and by increasing the rate of pre-emptive matches (Fig. 4.6), where a subject pre-empts the upcoming phrase type of the playback speaker, or presumably, under natural conditions, its rival. In natural interactions (Fig. 4.2c), the tendency to align their sequences may be still greater than during playback of pre-determined sequences, since both individuals may engage in song mirroring in a reciprocal fashion. However, it is also possible that song mirroring may occur asymmetrically if it is related to motivational state or dominance (Kroodsma 1979).

#### **4.4.2 Song Mirroring as a Signal**

The precise information conveyed through song mirroring in Cassin's Vireo remains unclear. The relationship between distance and song mirroring was strong (Fig. 4.6), and similar findings from song matching studies have been presented as evidence of aggressive content in matched singing (Krebs et al. 1981; Vehrencamp 2001; Vehrencamp et al. 2007; Akçay et al. 2013). In Cassin's Vireo, however, song mirroring manifests itself as song advancing, which differs from song matching in important ways. The detection of song advancing by a receiver requires familiarity with a rival's syntax, since advancing involves responding with unlike song types related only by intrinsic sequencing rules. The same cannot be said for song matching, which is readily detectable by human observers (Kroodsma 1971) and by birds (Nielsen and Vehrencamp 1995; Burt et al. 2001). Song advancing may play a similar communicative role in Cassin's Vireo by aligning shared song sequences, but the potentially reduced detectability of the behavior means receiver-perspective studies will be necessary to assess the salience of the signal.

Regarding song matching, the SMH presents an alternative perspective on the proximate mechanisms that may underlie the behavior in species that sing with eventual variety, and we propose that it may be song mirroring, not song matching in the traditional sense, that is widespread in songbirds. This perspective does not conflict with previous research regarding the signal content of matched singing. Recent studies support the notion that song matching is a conventional signal of aggression: matching often precedes attack (Akçay et al. 2013), and signal honesty is maintained by the threat of retaliation (Molles and Vehrencamp 2001; Vehrencamp 2001). Instead of contradicting these results, the SMH provides a plausible explanation for why song matching may come to be viewed as aggressive in the first place. The SMH implies that a bird needn't 'select' a song from their repertoire as a response; instead, the matching response may emerge as a byproduct of a bird's tendency to repeat song types. The ability of a perceived

song to trigger this response may vary as a function of a singer's intrinsic characteristics – including dominance status, attentiveness to their acoustic environment, or motivational state – or extrinsic characteristics, such as the amplitude and degradation of the acoustic signal, that are related to the distance between sender and receiver (Morton 2012). A type-matching bird, therefore, may represent a greater threat to a territory-holder than a non-matching bird, on account of their closer proximity and demonstrated attentiveness.

A further consideration is that the SMH, if broadly applicable, provides insight into the evolution and maintenance of eventual variety singing, a problem that remains unresolved in the birdsong literature (Price 2013). Models that invoke sexual selection through female choice as a driver of the evolution of song repertoire size predict that birds should sing with immediate variety to reveal their repertoire rapidly (Byers and Kroodsma 2009). Why, then, should so many species deliver songs with eventual variety, effectively obscuring the sizes of their repertoires? If song matching is underpinned by song mirroring, eventual variety may evolve precisely to facilitate song matching. As pointed out above, song matching is more easily detected and less ambiguous than song advancing, since it requires no prior knowledge of a rival's sequencing tendencies. Convergence among a population of songbirds towards eventual variety singing would increase the tendency to song match, thereby clarifying the focus of each bird's attention and increasing the effectiveness of threat signals, with the potential cost that repertoires are revealed more gradually.

#### **4.4.3 A Potential Neural Mechanism**

Recent work has revealed the existence of neurons in the songbird HVC that are active during both perception and production of particular song types (Prather et al. 2008). These so-called 'mirror neurons' have been proposed to play a role in song matching (Morton 2012), as it



is plausible that activation of these neurons during perception may lead birds to select the same song type as a response.

The SMH fits naturally within this framework. Models of songbird syntax are typically depicted as branching networks, where vertices represent song or syllable types, and directional edges represent preferred transitions (Jin and Kozhevnikov 2011); repetitions are likely governed by feedback mechanisms, but are incorporated within the same network architecture (Wittenbach et al. 2015). These branching networks are presumed to represent analogous networks of neurons in the brain (Jin 2013). If a perceived song type can activate the neurons encoding that song type in the receiver, as seems to be the case in Swamp Sparrows (Prather et al. 2008), it is plausible that the receiver's response may be to advance sequentially from that song type onward, whether that involves repeating the perceived song type or advancing to another.

#### **4.4.4 Future Research**

The SMH provides testable predictions about the ways in which songbirds should interact as a function of their singing mode. In immediate variety singers, the SMH predicts that song advancing, rather than song matching, should predominate. Though counter-singing remains understudied in immediate variety singers, published research occasionally provides counter-examples to the SMH, where song matching appears to occur (Kroodsma 1975). Importantly, such research has typically not considered song advancing at all, so additional studies including this lesser-known alternative should aim to evaluate the relative importance of matching and advancing in these species. Cassin's Vireo, for example, show behavior superficially resembling song matching (Fig. 4.2c), but our results demonstrate that this is a consequence of song advancing, not intentional matching.

Furthermore, song advancing and song matching are not necessarily mutually exclusive: even birds singing with immediate variety sometimes repeat song types, and in such cases, song matching may be a natural consequence of song mirroring. Banded Wrens, for example, generally sing with immediate variety, but exhibit switching rates of only  $\sim 0.5$  during counter-singing interactions (Molles and Vehrencamp 1999), which may account for observations of song matching in that species (Burt and Vehrencamp 2005), even if it is underpinned by song mirroring. Analytical approaches that explicitly incorporate sequential structure, such as the approach taken here, can be employed to better understand the nuances of vocal interactions in immediate variety singers.

In species that sing with eventual variety, the SMH again provides a testable hypothesis. Like immediate variety singers, eventual variety singers have preferred sequences of song types within their repertoires (Lemon and Chatfield 1971; Falls and Krebs 1975; Krebs 1976; Slater 1983). If the SMH is correct, we expect that, even when birds do not match song types, they may secondarily advance to their next preferred song type at greater than chance levels (Kroodsma 1979).

In sum, the SMH is supported by the evidence presented here for Cassin's Vireo, and we propose that song mirroring may represent a heretofore overlooked phenomenon common to many songbird species, encompassing those that sing with both eventual and immediate variety, and may help explain the use of song matching as a threat signal as well as the widespread, convergent evolution towards eventual variety as a singing mode. At a minimum, the results presented here illustrate the insights that can be gained by considering the preferred sequential ordering of songs in analyses of counter-singing interactions. Further efforts to test and refine the

SMH promise to improve our understanding of counter-singing interactions and the evolution of complex communication systems in general.

## **4.5 Materials and Methods**

### **4.5.1 Study Site and Species**

Research was conducted in a one-square-kilometer valley of mixed conifer-deciduous forest near Volcano, CA, USA (UTM: 10 S 706584 4262742, datum WGS 84). Experiments were approved by the Animal Research Committee at UCLA (ARC # 2013-041-03A). Cassin's Vireo males possess repertoires comprised of an average of 51 phrase types (range 31 to 60), that they deliver in structured sequences with immediate variety (Hedley 2016b). Phrase types are short (<0.7s long), highly stereotyped, shared widely throughout the population, and can be readily identified by a trained observer with >99% accuracy (Hedley 2016a). Details regarding the singing style of this species are examined in more detail in Hedley (2016a; 2016b).

### **4.5.2 Playback Design**

Eleven color-banded males were subjected to two playback trials, one with *typical* sequencing, and one with *atypical* sequencing (Fig. 4.2a). For typical trials, we selected eleven sets of five phrase types that were commonly clustered together in the song sequences of individuals in our study population. For atypical trials, we selected eleven sets of five phrase types that never, or exceedingly rarely, occurred adjacently in sequences. In the held out training set of 62395 phrases (See below and Appendix A), transitions between phrase types assembled into typical trials occurred 19697 times, whereas transitions between phrase types in atypical trials occurred just 220 times.

Each trial was designed specifically for each subject, taking into consideration their known repertoire, but not their syntax. Each individual was assigned typical and atypical

playback files containing either three or four shared phrase types, the remaining one or two being unshared, a design that simulated a real territorial intruder (Hedley 2016b). The number of shared and unshared phrase types was consistent between each subject's typical and atypical trials, allowing equal opportunity for matching in both trials. Subjects were presented with different phrase types in the typical and atypical trials, and each of the 55 phrase types used in the experiment occurred in exactly one typical and one atypical trial.

The five phrase types chosen for a given trial were arranged into a sequence of 25 phrases, where each type occurred five times. Within this sequence, phrase order was randomized in such a way that consecutive phrases were never of the same type. Phrases were delivered every two seconds, as is typical for this species, so the sequence took 50 seconds to complete, and was followed by ten seconds of silence. This one-minute segment repeated five times, comprising a five minute playback period.

We sourced phrase type exemplars from recordings of twelve individuals in the same study area during 2013 and 2014, but the ten phrase types presented to each bird needn't have originated from the same individual. Instead, phrase types were combined into artificial sequences to simulate an intruder with an unfamiliar repertoire and syntax. Three individuals were unintentionally presented with one to three of their own phrase types. Since subjects may respond differently to their own songs and those of a rival (Falls 1985; Stoddard et al. 1992), we ran the main analyses again without these individuals, confirming that the main results were robust (Appendix A).

### **4.5.3 Playback Field Methods**

We conducted playbacks from April 28 to May 31, 2015 between 7:00 and 11:00am. An attempt was made to conduct both trials during the same phase of nesting, either during territory

establishment prior to nesting (n=2 individuals) or during incubation (n=9 individuals) when the male was off the nest. Ten subjects received the two trials separated by one to three days, but one individual's nest failed between trials, so the second trial was postponed until incubation began on a subsequent nest 23 days later. We placed a Jawbone Jambox speaker about 1.8m up in a tree, well within the bird's known territory, and placed flags ten meters on each side to aid in distance estimation. The speaker was placed in the same location for both trials, and presentation order of the two trials was randomized. Phrase amplitude was normalized and broadcast at ~80 dB measured 1m from the speaker.

A single observer conducted all trials. The observer stood ~20m from the speaker, recorded the subject's songs with a Sennheiser MKH20-P48 microphone and Telinga parabolic reflector onto a Marantz PMD661 recording device, and dictated the bird's estimated distance from the speaker into the microphone. The experimental period lasted twelve minutes: the first two minutes were passive observation, the subsequent five minutes coincided with the broadcast of the playback file, and the last five minutes comprised a second period of passive observation. Two observers examined the spectrogram of the resulting .wav file in the program Praat (Boersma and Weenink 2014). The observers annotated to phrase type the phrases from the speaker and subject by assigning each phrase type a two-letter code (Hedley 2016b), and noted the subject's estimated distance from the speaker when each phrase was delivered. When the observers differed in their annotation of a phrase type (<1% of phrases), they discussed the discrepancy to arrive at a consensus.

#### **4.5.4 Motivation for Statistical Analyses**

Assessing whether a behavior occurs more often than expected by chance requires an estimation of this chance probability. Kroodsma (1975) stated that the probability of matching a

song type depends on four factors: 1) the frequency of occurrence of that phrase type overall; 2) the transition probability from the bird's most recent phrase type to the playback phrase type; 3) the amount of time since the bird most recently sang the phrase type in question; and 4) the vocal behavior of other males within earshot. We employed a model that incorporated properties of syntax from the songs of the subjects, and thereby effectively controlled for 1), 2) and 3). The influence of other males 4) is precisely what we hope to understand with playback experiments, and while it is possible that songs from non-subject males may affect the subject, such effects are likely to be minimal relative to the effect of the playback speaker, which was positioned within the territory to simulate a strong territorial intrusion.

#### 4.5.5 Statistical Analyses

We considered each of the subjects' phrases during the five minute playback trials to be an independent Bernoulli trial. Considering first the immediate matching (IM) model,  $y_i$  designated whether phrase  $i$  from the bird matched ( $y_i=1$ ) or did not match ( $y_i=0$ ) the most recent phrase type from the speaker. These Bernoulli trials were modeled using a logistic regression model

$$y_i | \mu_i \sim \text{Bernoulli}(\mu_i) \quad (1)$$

$$\text{logit}(\mu_i) = \text{logit}(p(y_i = 1 | x_{i-1})) + \alpha_0 \quad (2)$$

where  $\mu_i$  was the probability of matching at phrase  $i$ ,  $p(y_i = 1 | x_{i-1})$  was the probability of matching by chance given the subject's previous phrase type  $x_{i-1}$ , and  $\alpha_0$  was an intercept. The chance probability  $p(y_i = 1 | x_{i-1})$  was calculated using a Markov model, smoothed with Backoff smoothing (Jurafsky and Martin 2000) and Witten-Bell discounting (Witten and Bell 1991) and parametrized from a held out Training Set of recordings of the same subject under non-

experimental conditions (See Appendix A). An intercept ( $\alpha_0$ ) with posterior density significantly greater than zero would therefore indicate a tendency to match above chance levels.

Additional terms were added to the model by adding  $\alpha_{mt}T_i$ , to investigate treatment effects, or  $\alpha_{dist}D_i$ , to investigate distance effects, to the right side of equation (2).  $T_i$  was a covariate that was 0 for all phrases in atypical trials, and 1 in typical trials.  $D_i$  was the subject's estimated distance, in meters, from the speaker when phrase  $i$  was delivered. Significantly non-zero  $\alpha_{mt}$  and  $\alpha_{dist}$  coefficients reflected influences of treatment and distance, respectively, on the tendency to match.

We adapted this model to assess delayed matching and song advancing. In the delayed matching (DM) case,  $y_i=1$  if the bird sang any of the (up to) five phrase types that had come from the speaker earlier in the trial, and 0 otherwise. The chance probability  $p(y_i = 1|x_{i-1})$  was the sum of the probabilities of delivering each of these phrase types given the subject's previous phrase type.

In the first song advancing case ( $SA_1$ ), for each shared phrase type that emanated from the speaker, we used the Markov model to determine the phrase type that most commonly followed the stimulus phrase type in the subject's Training Set sequences.  $y_i$  was 1 if the bird sang that phrase type at phrase  $i$ , and 0 otherwise, and the null probability  $p(y_i = 1|x_{i-1})$  was the probability of singing that phrase type given the subject's previous phrase type. Because the syntax of this species is not deterministic, the SMH predicts that a bird may advance not only to its single most common subsequent phrase type, but also to its second, third, fourth or fifth ranked phrase types, though these are expected to occur at declining rates. We considered these four possibilities ( $SA_2$ ,  $SA_3$ ,  $SA_4$  and  $SA_5$  in Figs 4.4 & 4.5) independently in the same way as the initial song advancing ( $SA_1$ ) case. In the case of song mirroring (SM),  $y_i$  was 1 if the bird

responded to a playback phrase type with any of the three phrase types that most commonly followed it in that subject's training set, and  $p(y_i = 1|x_{i-1})$  was the sum of the chance probabilities in the SA<sub>1</sub>, SA<sub>2</sub> and SA<sub>3</sub> models.

As described above, the playback sequence contained shared and unshared phrase types. Since subjects never responded with an unshared phrase type, our analysis only considered the subject able to engage in immediate matching or song advancing if a) the most recent playback phrase type was shared, and b) the bird had not delivered any intervening phrases since the most recent playback phrase. Delayed matching was less constrained, since the bird could engage in delayed matching at any time as long as at least one shared phrase type had been delivered by the playback. Immediate matching and song advancing models were therefore based on lower sample sizes (N=1125 phrases) than the delayed matching model (N=2438 phrases).

Analyses were conducted using a Bayesian framework with uninformative priors (Normal with mean=0, SD=10) and 25000 Markov Chain Monte Carlo simulations in Jags (Plummer 2003), version 3.3.0, implemented in R (R Core Team 2014) with the package R2Jags (Su and Yajima 2015).



## List of Figures

**Fig. 4.1.** The song mirroring hypothesis depicted for eventual and immediate variety singers. Two birds are shown in both a) and b), their song types denoted with letters (A-D) and their individual identities with subscripts (1 or 2). a) The SMH predicts that species that sing with eventual variety should respond to perceived song types ( $D_2$ ) with type-matched replies ( $D_1$ ). b) In contrast, species that sing with immediate variety are predicted to respond by advancing from the perceived song type ( $C_2$ ) onward through their preferred progression of song types ( $D_1$ ). The hypothetical bird depicted delivers its repertoire of four song types in the preferred order ABCDABCD...

**Fig. 4.2.** Syntactic patterns are shared between individuals. a) and b) illustrate two separate sequences of solo singing by two individuals. Phrase type ID was assigned to phrase types on the order of appearance in a). Phrase types shared between a) and b) are highlighted with a gray background. Phrase types unique to b) were assigned numbers >46 as a continuation of the numbering in a). The tendency to deliver phrase types in clusters is evident in a), and shared phrase types are often clustered similarly in the songs in b), indicating shared sequencing patterns. The box highlights a cluster of five phrase types that may be suitable for a *typical* playback trial, since the phrase types are clustered in the sequences of multiple individuals. Arrows indicate phrase types suitable for an *atypical* trial, since they do not occur together in sequences of these or other birds. c) a counter-singing interaction between two neighbors (red crosses and black circles), with shared phrase types on a gray background. When one bird delivered a shared phrase type, the other often followed with that same phrase type soon thereafter. These bouts of parallel song appear to be common in Cassin's Vireo, and our

playbacks sought to identify whether this phenomenon is facilitated by song matching or song advancing.

**Fig 4.3.** Physical response of subjects to playback. Shown are the estimated distances for all eleven subjects across all 22 trials (gray lines) during the twelve minute experimental period, including two minutes of observation prior to playback, five minutes of playback, and five minutes of subsequent observation. The solid black line traces the mean distance from the speaker among all individuals at each second: subjects clearly approached the speaker immediately after the onset of playback, then typically remained in the vicinity of the speaker for the duration of playback, and gradually left the area following termination of playback.

**Fig. 4.4.** Summary of song matching and song advancing in response to playback. a) Expected rates ( $\pm 2$  SD) of immediate matching (IM), delayed matching (DM), and song advancing ( $SA_{1-5}$ ) are shown in light gray and observed rates in dark gray. Expected values and SDs were calculated using the formula for the Poisson binomial distribution:  $E(x) = \sum_{i=1}^n p_i$  and  $SD = \sqrt{\sum_{i=1}^n (1 - p_i)p_i}$  (Appendix A and Johnson et al. 2005). b) Rates of matching and song mirroring (SM) in response to the *atypical* (light gray) and *typical* (dark gray) playback trials. Expected values for each treatment are shown with horizontal lines for each trial, calculated as in a). SM was calculated as the sum of  $SA_1$ ,  $SA_2$  and  $SA_3$  from a). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

**Fig. 4.5.** Song advancing coefficient estimates reflect transition probabilities. a) we identified phrase types that occurred  $\geq 50$  times in a given bird's recordings in the held out Training Set,

since these are expected to have well-resolved transition probabilities (N=323 phrase types across eleven individuals). For each, we ranked the five most common subsequent phrase types and calculated the probability of transitioning to each. Horizontal lines, boxes, and error bars depict medians, quartiles, and extremes of the data, respectively. b) coefficient estimates, shown with 95% credible intervals, reflect the tendency to respond to a playback phrase type with their first through fifth ranked subsequent phrase type relative to chance levels. The declining tendency to respond with the first through fifth ranked phrase types roughly parallels the decline in transition probabilities in a), supporting the song mirroring hypothesis in this species.

**Fig. 4.6.** Song mirroring led to pre-emption of the subsequent playback phrase type (pre-emptive matching, see text), and occurred at greater rates closer to the speaker. a) Across all trials, song mirroring led to a significantly greater rate of pre-emptive matches when compared with responses that were not classified as song mirroring. b) When song mirroring occurred, it coincided with significantly higher rates of pre-emptive matching in typical trials than in atypical trials, suggesting that song mirroring only facilitates the juxtaposition of phrase types when two participants share similar sequencing rules, as is the case in Cassin's Vireos in the population under study. c) Song advancing declined with increasing distance from the speaker. Average rates of song advancing expected by chance (5.4% overall) are shown with a dotted line. The solid line shows the prediction of our model of the rate of song advancing as a function of the subject's distance from the speaker. Gray shading shows the 95% credible interval. \*\*\*,  $p < 0.0001$ .

## **List of Tables**

**Table 4.1.** Singing modes of species that demonstrate song matching and song advancing. For inclusion in this table, species must appear to engage in either song matching or song advancing at greater than chance levels. Some studies demonstrated matching or advancing with rigorous statistical analyses. Others mentioned these behaviors with relatively minimal evidence, but were nonetheless included here as long as these behaviors appeared to occur above chance levels.

## Figures and Tables

Figure 4.1

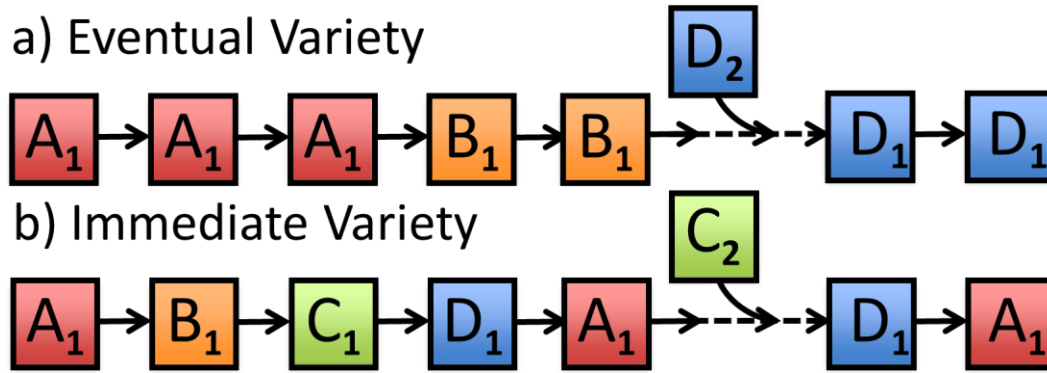


Figure 4.2

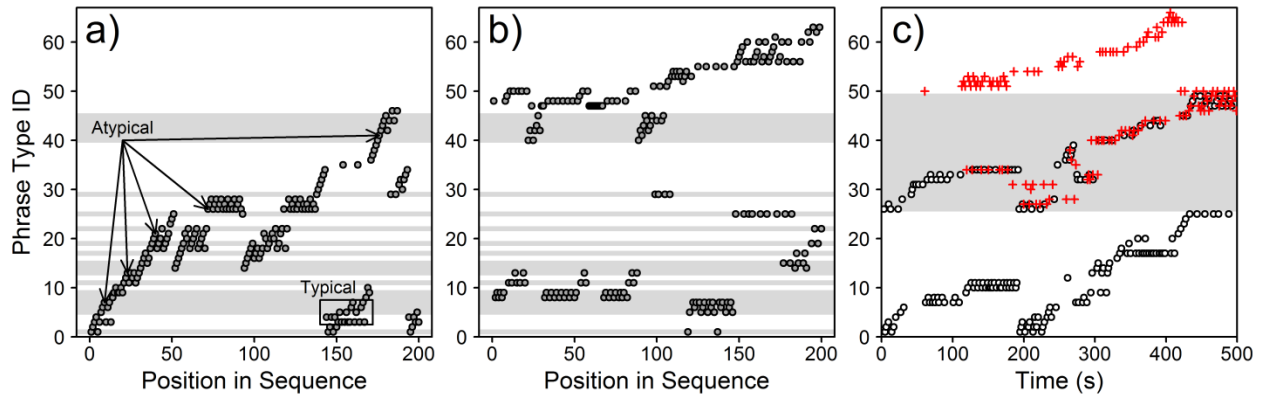


Figure 4.3

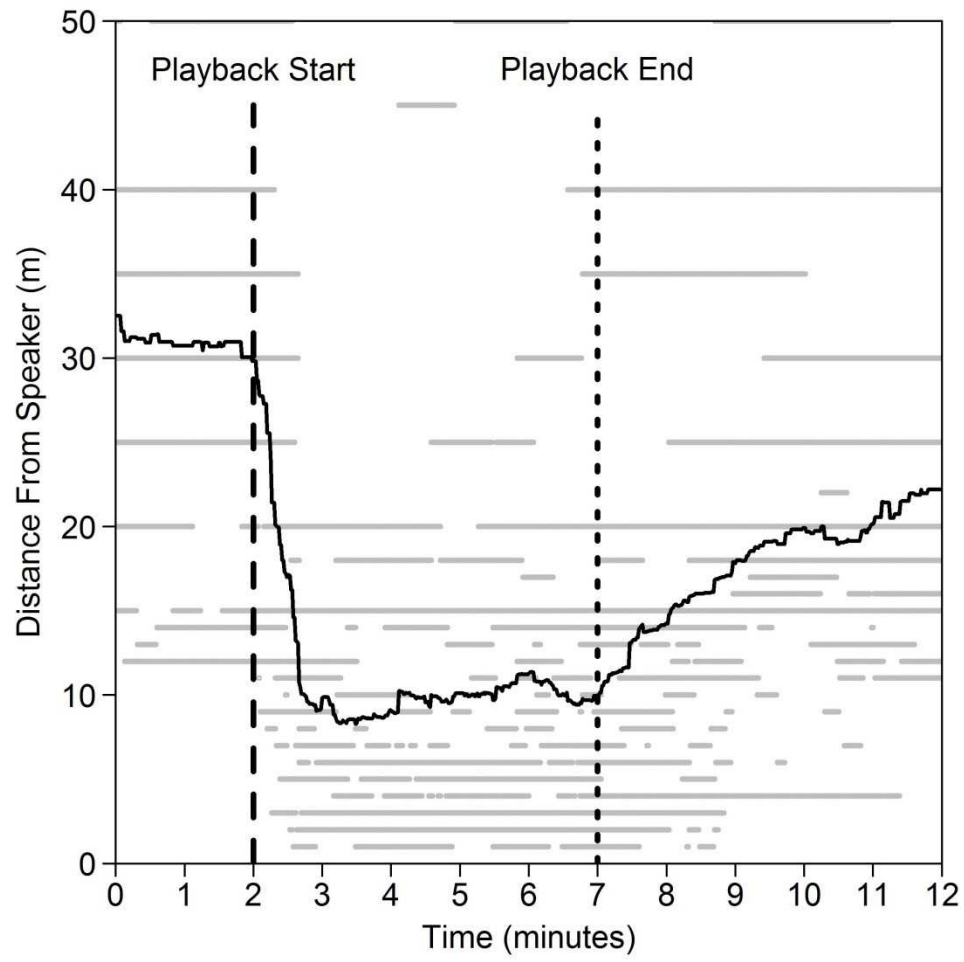


Figure 4.4

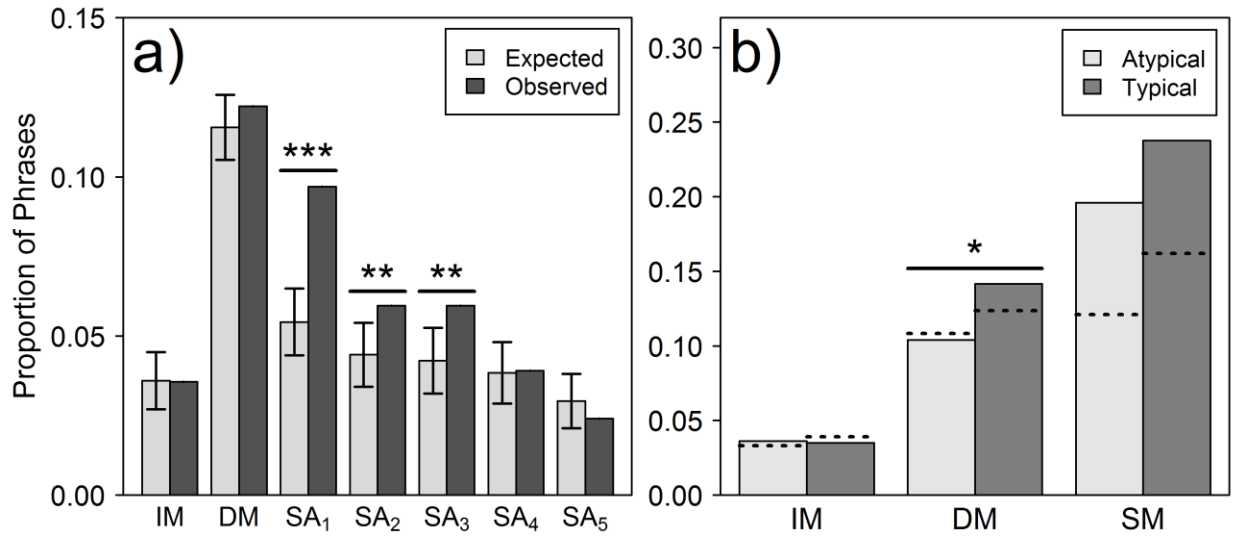


Figure 4.5

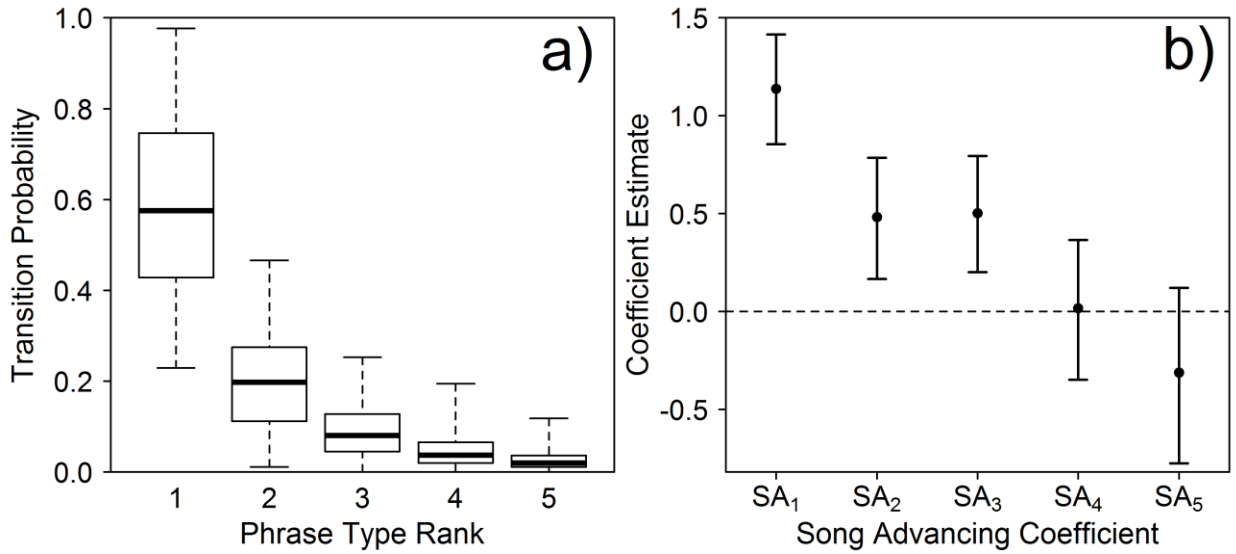


Figure 4.6

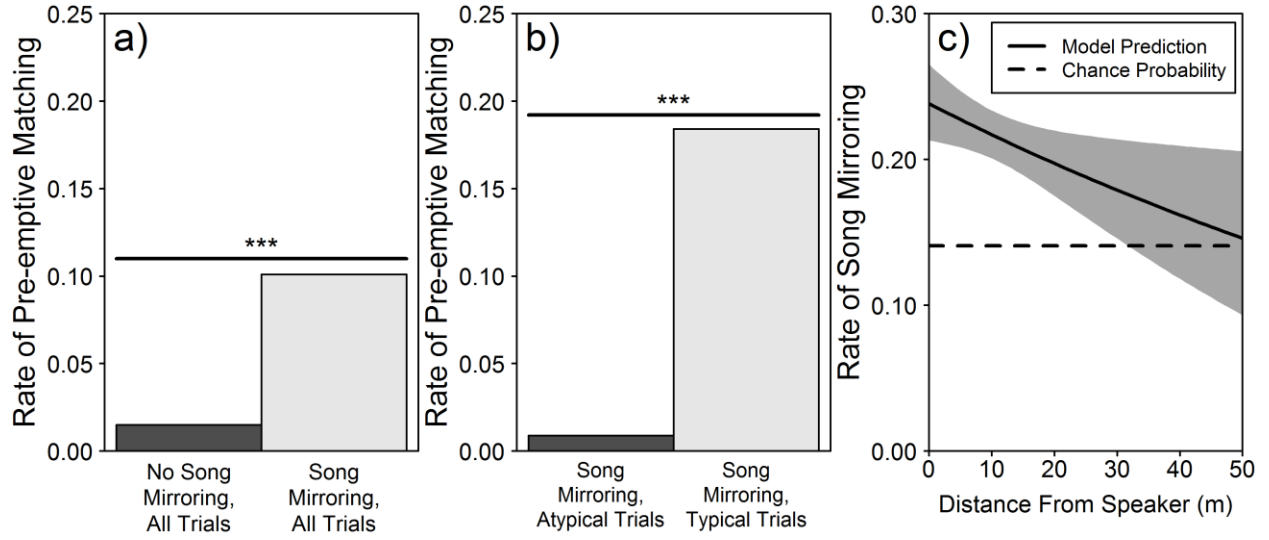




Table 4.1

<b>Singing mode</b>	<b>Family</b>	<b>Species</b>	<b>Engage in song matching</b>	<b>Engage in song advancing</b>	<b>References</b>
Eventual variety	Paridae	Great tit ( <i>Parus major</i> )	x		(Krebs et al. 1981)
		Black-capped Chickadee ( <i>Poecile atricapillus</i> )	x		(Gammon et al. 2008)
		Tufted Titmouse ( <i>Baeolophus bicolor</i> )	x		(Schroeder and Wiley 1983)
		Black-crested Titmouse ( <i>Baeolophus atricristatus</i> )	x		(Lemon 1968)
		Juniper Titmouse ( <i>Baeolophus ridgwayi</i> )	x		(Dixon 1969)
	Emberizidae	Song sparrow ( <i>Melospiza melodia</i> )	x		(Stoddard et al. 1992)
		Spotted towhee ( <i>Pipilo maculatus</i> )	x		(Kroodsma 1971)
	Icteridae	Western meadowlark ( <i>Sturnella neglecta</i> )	x		(Falls 1985)
	Cardinalidae	Northern cardinal ( <i>Cardinalis cardinalis</i> )	x		(Lemon 1974)
		Pyrrhuloxia ( <i>Cardinalis sinuatus</i> )	x		(Lemon and Herzog 1969)
	Dasyornithidae	Rufous bristlebird ( <i>Dasyornis broadbenti</i> )	x		(Rogers 2004)
	Sturnidae	Indian Hill Mynah ( <i>Gracula religiosa</i> )	x		(Bertram 1970)
	Fringillidae	Chaffinch ( <i>Fringilla coelebs</i> )	x		(Hinde 1958)
Troglodytidae	Carolina wren ( <i>Thryothorus ludovicianus</i> )	x		(Simpson 1985)	
Immediate variety	Troglodytidae	Rock wren ( <i>Salpinctes obsoletus</i> )	x		(Kroodsma 1975)
		Marsh wren ( <i>Cistothorus palustris</i> )	x	x	(Kroodsma 1979)
		Banded wren ( <i>Thryothorus pleurostictus</i> )	x**		(Molles 2006)
	Turdidae	Wood thrush ( <i>Hylocichla mustelina</i> )		x	(Whitney 1991)

	Common blackbird ( <i>Turdus merula</i> )	x	x	(Hultsch and Todt 1981)
Mimidae	Tropical mockingbird ( <i>Mimus gilvus</i> )	x		(Price and Yuan 2011)
Muscicapidae	Nightingale ( <i>Luscinia megarhynchos</i> )	x	x	(Todt 1971; Geberzahn et al. 2002)

\*\*Only delayed matching occurred at greater than chance levels, not immediate matching.

## Appendix A

### Estimation of Null Probabilities of Matching and Advancing

**Background.** Suppose a bird possesses a repertoire of four song types, which it delivers in roughly stereotyped sequences in the order ABCDABCDAB...

If we present this hypothetical bird with song type B, traditional analyses have assumed that the probability of responding with B by chance is 0.25, justified by the observation that the four song types are delivered at approximately equal rates. However, it is clear that the probability of responding with B varies as a function of the subject's previously delivered song type. Had the bird delivered song type A immediately prior to the playback of song type B, the probability of subsequently responding with B would be close to 1 by chance alone. This is because the sequence A-B appears to be preferred by the bird, so a response with B is expected, even in the absence of a playback stimulus.

In contrast, if the bird delivered song type C immediately prior to the presentation of B, the probability of responding with B by chance would be near zero, since the bird appears to rarely, if ever, deliver the sequence C-B under natural conditions. In this case, a response with B is much less likely to occur by chance, and therefore more likely to have been elicited by the playback stimulus.

Previous research has employed binomial tests (Stoddard et al. 1992; Rogers 2004), G-tests (Gammon et al. 2008; Price and Yuan 2011), or chi-square tests (Lemon 1974; Kroodsma 1975; Falls 1985) to evaluate the significance of matching. These tests differ in the calculation of chance probabilities, but all assume samples to be independent (Sokal and Rohlf 1995). The above example, however, demonstrates that this independence is violated in most, if not all, bird song analysis: the probability of singing any given song type depends upon the identity of the

previous song type. Researchers have circumvented this issue, in some cases, by considering only the first song of each bird's response in their analysis; playbacks conducted on different days or on different subjects can reasonably be assumed to be independent. While this is a suitable solution in some cases, it restricts our ability to study species in which the nuance of their response derives from the sequential delivery of song types, rather than any one song in isolation.

Instead, a more appropriate statistical distribution appears to be the *Poisson binomial* distribution which, like the binomial distribution, models the expected number of successes of a series of Bernoulli trials. Unlike the binomial distribution, the probabilities of success in the Poisson binomial distribution needn't be identical between trials. Our model derives from the Poisson binomial distribution, where the probability of success (i.e. matching or advancing) at phrase  $i$  varies depending on the phrase type that the bird delivered immediately prior at phrase  $i-1$ . Our analysis aimed to assess whether the introduction of playback phrases alters their response in any predictable fashion, which would be reflected in observed numbers of matches or advances above or below expected levels.

The expected number of successes deriving from a Poisson Binomial distribution can be approximated as  $E(x) = \sum_{i=1}^n p_i$ , and the variance as  $\sigma^2 = \sum_{i=1}^n (1 - p_i)p_i$ . Additional details regarding this distribution can be found in Johnson et al. (2005), pages 144-147.

**Markov Models and the Training Set.** The Poisson binomial distribution requires an estimation of the probability of success at each time step, or in this case, for each phrase with which the bird responds. For this purpose, we used a first-order Markov model trained on a set of recordings of the subject under non-experimental conditions. A first-order Markov model is comprised of a single state for each phrase type in a bird's repertoire, and a set of transition

probabilities describing the probability of transitioning from the current phrase type to any other (Chatfield and Lemon 1970). For individuals with a repertoire size of  $C$ , a Markov model is represented as a  $C \times C$  matrix, where rows represent the preceding state, columns represent the following state, and each of the  $C^2$  cells contain a transition probability of moving from each preceding state to each following state. The probability that an arbitrary phrase type  $s$  follows phrase type  $r$  is calculated as

$$p(x_i = s | x_{i-1} = r) = \frac{n_{rs}}{n_{r\bullet}}$$

Where  $i$  indexes the position in a sequence,  $n_{rs}$  is the number of times the bigram  $rs$  was observed in the training set, and  $n_{r\bullet}$  is the number of times phrase type  $r$  was followed by any phrase type in the training set (i.e. the number of times that  $r$  occurred, except as the terminal phrase in a recording).

Reliance upon this model assumes a) that transition probabilities can be approximated effectively with a first-order Markov model, b) that transition probabilities are roughly stable over time and c) that the sample sizes in the training set recordings are sufficient to parametrize the Markov model.

Condition a) has been discussed in detail in Hedley (2016a). The songs of fourteen Cassin's Vireo, including seven of the subjects of the current study, were analyzed for the complexity and time homogeneity of their syntax. The best fit model was intermediate in complexity between a first- and second-order model, so it is clear that the first-order Markov model does not perfectly model the syntax of this species. However, a first-order Markov model can be derived as a weighted average of the more complex second-order Markov model, so it provides a reasonable approximation of transition probabilities and is not expected to systematically biased towards either high or low probabilities. We chose to use first-order

Markov models for their ease of use and because of their lower sample sizes required for parametrization.

Condition b) was also discussed in Hedley (2016a): transition probabilities were shown to be stable, showing little within-individual variation when calculated from two non-overlapping sets of recordings from the same bird, even when those recordings spanned multiple years. The conclusion was that transition probabilities are largely fixed in adult life, though the possibility was left open that these probabilities may vary on shorter time scales in response to social or contextual cues.

Condition c), the sample sizes required for model parametrization, has not been addressed elsewhere. For the purposes of our model, each of the ~2500 cells in the Markov transition matrix required a non-zero probability, even though most of these transitions are exceedingly rare and are not likely to be observed in a training set of finite size. In linguistics, this problem is referred to as the “zero-frequency problem” (Witten and Bell 1991) and is commonly overcome with the use of smoothing, which adds probability mass to transitions that were not observed in the training set, based on each phrase type’s overall frequency of occurrence. Here we used Backoff smoothing (Jurafsky and Martin 2000) for this purpose, coupled with Witten-Bell discounting (Witten and Bell 1991).

Adding probability to unobserved transitions requires a corresponding reduction in the probability attributed to observed transitions, since the probability distribution associated with each model state must sum to one. Witten-Bell Discounting (Witten and Bell 1991) accomplishes this by reducing observed transition probabilities by  $N/(N+T)$ , where  $N$  is the total number of observations from which the transition probability was calculated (i.e. the row sum in the transition matrix), and  $T$  is the total number of observed transitions for the given state (i.e. the

number of non-zero cells in the row). Transition probabilities for states that were parametrized with large numbers of observations were therefore discounted less than those with fewer observations. The remaining weight,  $T/(N+T)$ , is distributed among the unobserved transitions for that state according to the weighting determined by Backoff smoothing. Backoff smoothing works by ‘backing off’ to lower levels of model complexity to estimate the probabilities of unobserved transitions (Jurafsky and Martin 2000). That is, if transition A-B was observed in the training set, the estimated transition probability,  $p_{AB}$ , was used, albeit discounted via Witten-Bell discounting. If transition A-B was not observed in the training set, it was allocated near-zero probabilities proportional to the overall rates of occurrence of the latter phrase type in the training set (i.e.  $p(x_i = B | x_{i-1} = A) \propto p(B)$ ), such that the sum of their probabilities summed to the remaining weight,  $T/(N+T)$ , leftover from the discounting step.

The number of phrases in each individual’s training set varied from 1498 to 14101 (mean=5672 phrases), and the number of recording files from nine to 84 files (mean=27 files). The training sets from five individuals spanned three years, those of two individuals spanned two years, and those of the remaining four individuals included only recordings from 2015.

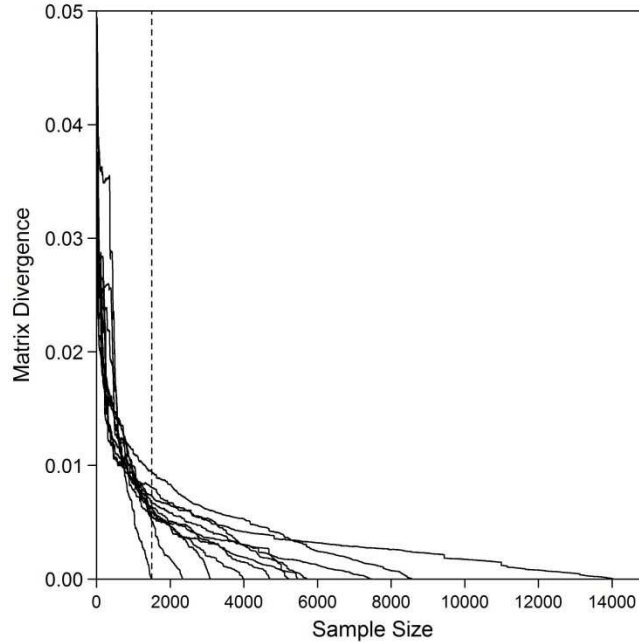
The last question to be considered was whether the sample sizes in the training set were sufficient to accurately estimate the Markov transition probabilities. We examined the changes in estimated transition probabilities as a function of sample size, with the expectation that, after sufficient sampling, the transition probabilities should reach an equilibrium. We began with a transition matrix full of zeroes, adding single observations at a time to the matrix. At each step, smoothing and discounting was applied as above, and the resulting matrix was compared to the matrix for that individual containing the full dataset. This comparison was done by calculating the mean per-transition divergence between the two matrices, with

$$\text{Matrix Divergence} = \frac{1}{C^2} \sum_{j=1}^C \sum_{i=1}^C |x_{ij} - y_{ij}|$$

where  $C$  is the number of rows and columns in the transition matrix, and  $x_{ij}$  and  $y_{ij}$  are the estimated transition probabilities in the cell at row  $i$  and column  $j$ , in the matrix with reduced sample sizes and that parametrized from the bird's full training set, respectively. By plotting this metric as a function of the sample size included in the reduced transition matrix, the influence of sample size on estimated transition probabilities can be seen (Fig 4.7). It is evident that transition probabilities continue to change even after several thousand phrases have been sampled.

However, the rate of change in all individuals was very rapid for the first ~1000 phrases, and incremental thereafter. Importantly, all of the individuals in our experiments had training sets with sample sizes well beyond 1000 phrases, so we presume the transition probabilities used in our analysis reasonably estimate the transition probabilities that would be expected with an arbitrarily large sample size. Furthermore, even if transition probabilities were only approximately correct, there is no reason to expect them to be systematically biased either high or low, so deviation from true probabilities should have minimal effects on our conclusions.





**Fig. 4.7.** Influence of sample size on transition probability estimation. The patterns for the eleven individuals are shown with solid lines, and the vertical dotted line shows the sample size for the individual with the smallest training set. As observations were added to the transition matrix, the transition probabilities converged towards the probabilities derived from the full training set for that individual, reflected in lower matrix divergence values (see text). Though transition probabilities continued to change as observations were added, the changes from samples 1,000 to 14,000 are of much smaller magnitude than the changes from samples 1 to 1,000.

### **Analysis excluding individuals subjected to ‘self-song’**

**Background.** Several papers have illustrated that birds respond differently to playbacks of their own songs in comparison to playbacks of stranger or neighbor songs (Falls 1985; Stoddard et al. 1992). Our playback tracks involved artificially concatenated phrases from different individuals to simulate an unfamiliar intruder into the territory. The phrases were derived from recordings in 2013 and 2014, and most of the phrases originated from individuals no longer present at our study site. Three individuals, however, were unintentionally subjected to one or more of their own phrase types. Because this may influence the response, we analyzed the results presented in the main text again, excluding these three individuals.

**Results: matching or advancing above chance levels.** All of our results from the main analysis were robust to the removal of the three individuals: immediate matching was still not significant ( $\alpha_0=0.060\pm 0.23$  SD,  $p(\alpha_0>0)=0.61$ ); delayed matching was not significant ( $\alpha_0=0.076\pm 0.094$  SD,  $p(\alpha_0>0)=0.79$ ); advancing to any of the three most preferred subsequent phrase types occurred above chance levels (SA<sub>1</sub>:  $\alpha_0=1.12\pm 0.17$  SD,  $p(\alpha_0>0)\geq 0.9999$ ; SA<sub>2</sub>:  $\alpha_0=0.39\pm 0.19$  SD,  $p(\alpha_0>0)=0.98$ ; SA<sub>3</sub>:  $\alpha_0=0.47\pm 0.19$  SD,  $p(\alpha_0>0)=0.99$ ); and advancing to the fourth and fifth preferred subsequent phrase types did not occur above chance levels (SA<sub>4</sub>:  $\alpha_0=0.071\pm 0.22$  SD,  $p(\alpha_0>0)=0.64$ ; SA<sub>5</sub>:  $\alpha_0=-0.27\pm 0.27$  SD,  $p(\alpha_0>0)=0.16$ ).

Additionally, the individuals that were presented with one or more ‘self’ phrase types engaged in song mirroring (SA<sub>1</sub>, SA<sub>2</sub> and SA<sub>3</sub> above, combined) at nearly identical levels to the individuals presented with only ‘other’ phrase types: subjects presented with their own phrase types engaged in mirroring at a rate of 21.5%, while other individuals responded with song mirroring at a rate of 21.7%.

**Results: treatment effects.** As in the main analysis, immediate matching showed no evidence of treatment effects ( $\alpha_{imt}=-0.002\pm 0.47$  SD,  $p(\alpha_{imt}>0)=0.50$ ), and delayed matching showed a significantly positive treatment effect ( $\alpha_{imt}=0.53\pm 0.19$  SD,  $p(\alpha_{imt}>0)=0.998$ ). Song mirroring did not show a significant treatment effect ( $\alpha_{imt}=-0.21\pm 0.26$  SD,  $p(\alpha_{imt}>0)=0.21$ ).

**Results: distance effects.** The effects of distance were robust to the removal of individuals. The distance effect was not significant for immediate or delayed matching (IM:  $\alpha_{dist}=-0.0081\pm 0.020$  SD,  $p(\alpha_{dist}>0)=0.67$ ; DM:  $\alpha_{dist}=-0.0040\pm 0.010$  SD,  $p(\alpha_{dist}>0)=0.34$ ), but was significant in the case of song advancing (SM:  $\alpha_{dist}=-0.035\pm 0.013$  SD,  $p(\alpha_{dist}>0)=0.0033$ ).

**Conclusion.** Our intent here was not to assess whether Cassin’s Vireos respond differently to songs according to the identity of the singer – additional experiments, designed for

this purpose, would be necessary to determine whether or not this is true. Rather, we aimed to address the potential concern that our overall results may have been driven by distinct responses of the three individuals that were presented with ‘self’ phrase types. We have demonstrated here that our primary results are robust to the removal of these three individuals. It is worth noting that only a minority of the ten playback phrase types presented to these three individuals were ‘self’ phrases (one of ten phrase types for two birds, and three of ten for the other bird); the majority of the playbacks were comprised of phrase types from other individuals, which may additionally explain why these birds did not appear to respond differently. We included these three individuals in our main analysis, to avoid reducing sample sizes unnecessarily.

## References

- Akçay Ç, Tom ME, Campbell SE, Beecher MD (2013) Song type matching is an honest early threat signal in a hierarchical animal communication system. *Proc R Soc B* 280:20122517. doi: 10.1098/rspb.2012.2517
- Beecher MD, Campbell SE, Burt JM, et al. (2000) Song-type matching between neighbouring Song Sparrows. *Anim Behav* 59:21–27. doi: 10.1006/anbe.1999.1276
- Bertram B (1970) The vocal behaviour of the Indian hill myna, *Gracula religiosa*. *Anim Behav Monogr* 3:79–192. doi: 10.1016/S0066-1856(70)80003-6
- Boersma P, Weenink D (2014) Praat: doing phonetics by computer.
- Burt JM, Bard SC, Campbell SE, Beecher MD (2002) Alternative forms of song matching in Song Sparrows. *Anim Behav* 63:1143–1151. doi: 10.1006/anbe.2002.3011
- Burt JM, Campbell SE, Beecher MD (2001) Song type matching as threat: a test using interactive playback. *Anim Behav* 62:1163–1170. doi: 10.1006/anbe.2001.1847
- Burt JM, Vehrencamp SL (2005) Dawn chorus as an interactive communication network. In: McGregor PK (ed) *Anim. Commun. Networks*. Cambridge, UK, pp 320–343
- Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. *Anim Behav* 77:13–22. doi: 10.1016/j.anbehav.2008.10.003
- Chatfield C, Lemon RE (1970) Analysing sequences of behavioural events. *J Theor Biol* 29:427–445. doi: 10.1016/0022-5193(70)90107-4
- Dixon KL (1969) Patterns of singing in a population of the Plain Titmouse. *Condor* 71:94–101.
- Falls JB (1985) Song matching in Western Meadowlarks. *Can J Zool* 63:2520–2524. doi: 10.1139/z85-373
- Falls JB, Krebs JR (1975) Sequence of songs in repertoires of Western Meadowlarks (*Sturnella neglecta*). *Can J Zool* 53:1165–1178. doi: 10.1139/z75-135
- Gammon D, Hendrick M, Baker M (2008) Vocal communication in a songbird with a novel song repertoire. *Behaviour* 145:1003–1026.
- Geberzahn N, Hultsch H, Todt D (2002) Latent song type memories are accessible through auditory stimulation in a hand-reared songbird. *Anim Behav* 64:783–790. doi: 10.1006/anbe.2002.3099
- Hedley RW (2016a) Complexity, predictability and time homogeneity of syntax in the songs of Cassin's Vireo (*Vireo cassinii*). *PLoS One* 11:e0150822. doi: 10.1371/journal.pone.0150822

- Hedley RW (2016b) Composition and sequential organization of song repertoires in Cassin's Vireo (*Vireo cassinii*). *J Ornithol* 157:13–22. doi: 10.1007/s10336-015-1238-x
- Hinde R (1958) Alternative motor patterns in Chaffinch song. *Anim Behav* 6:211–218.
- Hultsch H, Todt D (1981) Repertoire sharing and song-post distance in nightingales (*Luscinia megarhynchos* B.). *Behav Ecol Sociobiol* 8:183–188. doi: 10.1007/BF00299828
- Jin DZ (2013) The neural basis of birdsong syntax. *Prog. Cogn. Sci. from Cell. Mech. to Comput. Theor.* 1951:
- Jin DZ, Kozhevnikov AA (2011) A compact statistical model of the song syntax in Bengalese finch. *PLoS Comput Biol* 7:e1001108. doi: 10.1371/journal.pcbi.1001108
- Johnson NL, Kemp AW, Kotz S (2005) *Univariate Discrete Distributions*. John Wiley and Sons, Inc., Hoboken, New Jersey, USA
- Jurafsky D, Martin JH (2000) *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. doi: 10.1162/089120100750105975
- Krebs J, Ashcroft R, Orsdol K (1981) Song matching in the Great Tit *Parus major* L. *Anim Behav* 29:918–923.
- Krebs JR (1976) Habituation and song repertoires in the Great Tit. *Behav Ecol Sociobiol* 1:215–227. doi: 10.1007/BF00299198
- Kroodsma D (1979) Vocal dueling among male Marsh Wrens: Evidence for ritualized expressions of dominance/subordination. *Auk* 96:506–515.
- Kroodsma D (1971) Song variations and singing behavior in the Rufous-sided Towhee, *Pipilo erythrophthalmus oregonus*. *Condor* 73:303–308.
- Kroodsma DE (1975) Song patterning in the Rock Wren. *Condor* 77:294–303.
- Lemon E (1968) Coordinated singing by Black-crested Titmice. *Can J Zool* 46:1163–1167.
- Lemon RE (1974) Song dialects, song matching and species recognition by cardinals *Richmondia cardinalis*. *Ibis (Lond)* 116:545–549.
- Lemon RE, Chatfield C (1971) Organization of song in cardinals. *Anim Behav* 19:1–17. doi: 10.1016/S0003-3472(71)80129-X
- Lemon RE, Herzog A (1969) The vocal behavior of Cardinals and Pyrrhuloxias in Texas. *Condor* 71:1–15.
- Logue DM, Forstmeier W (2008) Constrained performance in a communication network: Implications for the function of song-type matching and for the evolution of multiple ornaments. *Am Nat* 172:34–41.

- Molles L (2006) Singing complexity of the Banded Wren (*Thryothorus pleurostictus*): Do switching rate and song-type diversity send different messages? *Auk* 123:991–1003.
- Molles LE, Vehrencamp SL (2001) Songbird cheaters pay a retaliation cost: evidence for auditory conventional signals. *Proc Biol Sci* 268:2013–2019. doi: 10.1098/rspb.2001.1757
- Molles LE, Vehrencamp SL (1999) Repertoire Size, Repertoire Overlap, and Singing Modes in the Banded Wren (*Thryothorus pleurostictus*). *Auk* 116:677–689. doi: 10.2307/4089329
- Morton ES (2012) Putting distance back into bird song with mirror neurons. *Auk* 129:560–564. doi: 10.1525/auk.2012.12072
- Nielsen B, Vehrencamp S (1995) Responses of Song Sparrows to song-type matching via interactive playback. *Behav Ecol Sociobiol* 37:109–117.
- Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc 3rd Int Work Distrib Stat Comput* 1–10. doi: ISSN 1609-395X
- Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 451:305–310. doi: 10.1038/nature06492
- Price J (2013) Why is birdsong so repetitive? Signal detection and the evolution of avian singing modes. *Behaviour* 150:1–19. doi: 10.1163/1568539X-00003051
- Price JJ, Yuan D (2011) Song-type sharing and matching in a bird with very large song repertoires, the Tropical Mockingbird. *Behaviour* 148:673–689. doi: 10.1163/000579511X573908
- Prum RO, Berv JS, Dornburg A, et al. (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573. doi: 10.1038/nature15697
- R Core Team (2014) R: a language and environment for statistical computing.
- Rogers D (2004) Repertoire size, song sharing and type matching in the Rufous Bristlebird (*Dasyornis broadbenti*). *Emu* 104:7–13.
- Schroeder DJ, Wiley RH (1983) Communication with shared song themes in Tufted Titmice. *Auk* 100:414–424.
- Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. *Anim Behav* 78:1281–1292. doi: 10.1016/j.anbehav.2009.08.011
- Simpson B (1985) Effects of location in territory and distance from neighbours on the use of song repertoires by Carolina Wrens. *Anim Behav* 33:793–804.
- Slater PJB (1983) Sequences of song in Chaffinches. *Anim Behav* 31:272–281.

- Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W. H. Freeman & Co., New York, NY
- Stoddard PK, Beecher MD, Campbell SE, Horning CL (1992) Song-type matching in the Song Sparrow. *Can J Zool* 70:1440–1444.
- Su Y-S, Yajima M (2015) R2jags: using R to run “JAGS.”
- Todt D (1981) On functions of vocal matching: Effect of counter-replies on song post choice and singing. *Z Tierpsychol* 93:73–93.
- Todt D (1971) Aquivalente und konvalente gesangliche Reaktionen einer extrem regelmabig singenden Nachtigall (*Luscinia megarhynchos* L.). *Zeitschrift f?r Vergleichende Physiol* 71:262–285. doi: 10.1007/BF00298139
- Vehrencamp SL (2001) Is song-type matching a conventional signal of aggressive intentions? *Proc Biol Sci* 268:1637–42. doi: 10.1098/rspb.2001.1714
- Vehrencamp SL, Hall ML, Bohman ER, et al. (2007) Song matching, overlapping, and switching in the Banded Wren: the sender’s perspective. *Behav Ecol* 18:849–859. doi: 10.1093/beheco/arm054
- Whitney C (1991) Avoidance of song matching in the Wood Thrush: a field experiment. *Wilson Bull* 103:96–101.
- Whitney CL (1985) Serial order in Wood Thrush song. *Anim Behav* 33:1250–1265. doi: 10.1016/S0003-3472(85)80185-8
- Witten IH, Bell TC (1991) The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Trans Inf Theory* 37:1085–1094. doi: 10.1109/18.87000
- Wittenbach JD, Bouchard KE, Brainard MS, Jin DZ (2015) An adapting auditory-motor feedback loop can contribute to generating vocal repetition. *PLoS Comput Biol* 11:1–29. doi: 10.1371/journal.pcbi.1004471
- Yasukawa K (1981) Song and territory defense in the Red-winged Blackbird. *Auk* 98:185–187.

## CHAPTER 5

### General Discussion

#### 5.1 Repertoires and Syntax in Cassin's Vireo Songs

The findings included here support the assertion of Byers and Kroodsma (2009) that song repertoires are not simple indices of male quality but form the basis of a sophisticated communication system governed by complex rules. As described in Chapter 2, the sequential delivery of songs – with phrase types clustered into themes and with strong biases towards some phrase types over others – does not fit the expectations of repertoire evolution driven by female choice. If female receivers were driving repertoire evolution, the tendency of males to obscure large portions of their repertoires would be unexpected. Instead, it appears that songs are more often directed at other males, and serve a primary function in the establishment and maintenance of territory boundaries. The results from Chapter 4 support this idea, by demonstrating that males respond very strongly to playback of songs, physically approaching the speaker and engaging in counter-singing exchanges with complex dynamics, including the hitherto undocumented phenomenon of song mirroring. Interestingly, females never responded to song playbacks, including the 22 trials conducted during data collection for Chapter 4, and the dozens of song broadcasts used to capture males for banding prior to data collection. This does not imply that females are completely insensitive to conspecific song, but does demonstrate that songs do not warrant the same urgency of response for females as they do for males.

Considered with respect to the three axes of song complexity outlined in the introduction, the songs of Cassin's Vireos appear to contain a considerable degree of complexity. The first component of complexity, in this case the complexity and performance of individual phrase types (section 1.1.1), remains largely unexplored. However, I demonstrated in Chapter 2 that



sharing of phrase types is common in this species, and individuals at my study site appear to draw their repertoires from a pool of locally typical phrase types, performing renditions of these phrase types with high degrees of stereotypy. Though it is not clear the costs and benefits associated with these shared phrase types, it seems likely that birds with anomalous songs would be diminished in their ability to interact with neighbors during counter-singing, which might present a strong constraint on song innovation at the level of individual phrase types. In other species, reduced sharing between territory neighbors, resulting from learning errors or innovation, has been associated with reduced territory tenure (Hiebert et al. 1989; Beecher et al. 2000). Further playback experiments that perturb the normal acoustic structure of phrase types would help to illuminate the consequences of song sharing and phrase type stereotypy for territorial interactions.

The second component of song complexity, repertoire size, was also described in Chapter 2. Repertoire sizes in this species were found to be much larger than previously documented by James (1981), and also showed a great degree of variability between individuals, ranging from 31 to 60 phrase types per individual across nineteen males recorded over three years. Because the number of individuals included in this study was fairly small, the consequences of this variation remain largely unknown. For example, it is not known if larger repertoire sizes correspond with increased longevity (Reid et al. 2005), territory quality (Catchpole 1986), reproductive success (Potvin et al. 2015), or other reproductive variables, as has been documented in other species.

Studies of bird song face an unavoidable tradeoff between the number of individuals sampled and the recording effort directed at each individual. While the focus in this thesis on a relatively small sample of individuals limited the statistical power to detect relationships between song characteristics and ecological covariates, the large per-individual samples facilitated a

detailed analysis of the third axis of song complexity outlined in Chapter 1 – the often overlooked aspect of song syntax. Analyses presented in Chapter 3 revealed that syntax in this species is too complex to be modeled with a first-order Markov model, that song sequences are predictable with over 55% accuracy, and that sequencing rules appear to be mostly stable within and between years. This is not to suggest, however, that syntax in this species is fully understood. As an example, the syntactic models employed in Chapter 3 were discrete-time models, meaning transitions between phrase types were analyzed without respect to the length of the silent interval between successive phrases. The existence of a relationship between predictability and inter-phrase interval (Fig 3.5), however, implies that the syntactic rules of these birds are not impervious to the effects of time. Improvements to the syntactic models employed here will likely result from the use of more sophisticated models that explicitly incorporate the lengths of silent intervals between phrases. While I suspect that improvements thus obtained would be minor, they would nonetheless help clarify some of the aspects of syntax that remain unexplored.

## **5.2 The Importance of Syntax for Understanding Bird Song**

Chapter 4 represented the culmination of a bottom-up approach to the study of the songs of this species. That is, the methods employed in that analysis relied upon the assessments of repertoires and the understanding of syntax developed in Chapters 2 and 3. Prior familiarity with the syntax of interacting individuals proved valuable in understanding the dynamics of counter-singing behavior in this species; indeed, the phenomenon of song mirroring could not have been detected without a detailed understanding of the syntax of the subjects under study.

Analyses explicitly incorporating syntax have been lacking from studies of bird song playback experiments and counter-singing, a shortcoming that may obscure the detection of such

phenomena. For example, Song Sparrows represent a model species for the study of singing behavior in the wild, but to my knowledge the sequential organization of song has never been studied in this species. As a result, it is not known if transitions between song types occur randomly or according to an internal syntax. The lack of attention to this aspect of Song Sparrow song is particularly vexing given the finding that switching between song types occurs at increased rates during territorial interactions (Kramer et al. 1985) and is temporally coordinated between the two participants (Kramer and Lemon 1983). While this has led to the claim that switching itself may constitute an agonistic signal (Bradbury and Vehrencamp 2011), research has thus far not endeavored to examine the nature of these switches between song types in any detail.

Similar to Song Sparrows, the songs of the Banded Wren have been studied extensively, with studies investigating neighbor recognition (Molles and Vehrencamp 2001b), song performance (Illes et al. 2006; Vehrencamp et al. 2013), and the use of song under different behavioral contexts (Trillo and Vehrencamp 2005). Several studies have attempted to decipher the strategies employed during counter-singing interactions. Banded Wrens responded with aggression when their song types were interactively matched during these exchanges (Molles and Vehrencamp 2001a), and the apparent occurrence of song matching during natural encounters has been used as evidence that song matching is a foundational component of the signaling behavior of this species (Burt and Vehrencamp 2005; Vehrencamp et al. 2007; Bradbury and Vehrencamp 2011). However, when males were presented with playback songs simulating a territorial intrusion, they failed to match the speaker more often than expected by chance (Molles 2006). If song matching is a signal in this species, as claimed, it is surprising that birds do not appear to intentionally engage in this behavior when prompted; clearly there is more to the story.

My results from Chapter 4 show striking similarities between the singing behavior of Banded Wrens and Cassin's Vireos: both species sing with immediate variety, observations of natural encounters in both species reveal apparently coordinated exchanges of the same songs in quick temporal succession (Fig 4.2), and neither species intentionally matched playback songs (Fig 4.4). In Cassin's Vireo, the nearly synchronized exchanges of the same phrase type appears to arise from a combination of song mirroring and shared sequencing rules, such that a bird interactively "jumps ahead" of its rival in a shared sequence of phrase types. The rival may "match" in response, but not due to any intent to do so, only because they continued to sing their preferred sequence which had been pre-empted by the instigating bird.

It is plausible that this same mechanism may apply in Banded Wrens, which might account for the apparently distinct results obtained from natural counter-singing interactions (Burt and Vehrencamp 2005) and playback experiments (Molles 2006). However, no efforts have been made to assess the sequencing rules employed by Banded Wrens. If song mirroring occurs in Banded Wrens, the absence of information regarding the sequencing rules employed by this species make it currently impossible to detect, with the consequence that our current understanding of counter-singing interactions is likely to be incomplete.

The lack of attention to song sequencing rules in Song Sparrows and Banded Wrens is not unique. Indeed, the general lack of information regarding the syntax of birds outside of the laboratory indicates that this aspect of song has generally been considered a low priority target for research. If behavioral ecologists truly seek to understand bird song as a communication system, the practice of ignoring certain aspects of song organization in favor of others will likely hinder progress. Further attention should be given to the development of analyses, like the

analyses presented in Chapter 4, that consider responses to conspecific song in light of all aspects of their normal song, encompassing their repertoires of songs and their syntax.

### **5.3 Functions of Song Syntax**

With respect to the information encoded in the syntax of songs, my playback experiments did not reveal obvious differences in the physical responses to playback of song with normal and abnormal syntax; both treatments were met with strong physical responses. This result contrasts with the results of others that have shown differential responses to songs with altered syntax (e.g. Holland et al. 2000; Briefer et al. 2013). An interesting and potentially meaningful difference exists between the syntax of Cassin's Vireo songs and that employed by other species that have been subjected to playbacks of altered and unaltered syntax, including the Eurasian Wrens (Holland et al. 2000), Skylarks (Briefer et al. 2013), Eurasian Starlings (Gentner and Hulse 1998; Gentner 2008) and California Thrashers (Taylor et al., unpublished). In the latter four species, syntax links vocalizations delivered in rapid succession, leading to variability within a single continuous burst of song. Intervals between syllables in these species are typically a fraction of a second (Briefer et al. 2013; Cody et al. 2016). In Cassin's Vireo song, in contrast, syntax governs transitions between phrase types that are typically separated by two seconds or more of silence – far longer than the length of the phrases themselves.

I suspect that these differences are likely to have important consequences for the role of syntax in communication. If the ordering of vocalizations, in and of itself, encodes information, shorter silent intervals and more continuous song should facilitate efficient exchange of this information. Long silent intervals, as common in Cassin's Vireo song, would have the opposite effect, obstructing the exchange of information. Not only might a receiver have to wait a few seconds to hear the subsequent vocalization of the singer, longer silent intervals would increase

the probability that background noise, such as the song of another species, would intervene and obscure one or more of the phrases in a sequence. A similar challenge would be faced by a human attempting to make sense of human speech if one word were delivered every two seconds – meaningful statements could take minutes to unfold, and would surely be more challenging to interpret.

When compared with the songs of more continuous singers, the widely separated phrases of Cassin's Vireos appear to be poorly adapted for the encoding of information. On the other hand, the timing of the songs appears well-suited for counter-singing exchanges, since intervals between phrases are sufficiently long to allow a rival to interject, leading to alternation between the two participants. I propose that the song syntax of Cassin's Vireos is primarily involved in counter-singing exchanges by facilitating more efficient exchanges through the use of shared sequences. The results of Chapter 4 seem to support this view, as song mirroring and shared sequences both seem to be important in counter-singing behavior. Shared sequencing tendencies between neighboring birds have been documented in only a few species (Slabbekoorn and Smith 2002; Kershenbaum et al. 2014), but are likely to be far more common than currently appreciated. Laboratory experiments have shown that song ordering, like the acoustic structure of songs, is memorized early in life in some songbirds (Todt and Hultsch 1998; Nowicki et al. 2001). A simple explanation for the patterns observed in Cassin's Vireo songs may be that shared sequences are acquired early in life, leading adolescent birds to conform to local song sequencing norms as they mature, and that these shared delivery patterns are then employed during counter-singing exchanges as adults. The ontogeny of this behavior, however, remains speculative, and would benefit from additional research.

#### **5.4 Directions for Future Research**

This thesis has answered some basic questions regarding the singing behavior of Cassin's Vireos, but has also made clear several deficits in our current knowledge of this species. First and foremost, the functional aspects of song mirroring remain unknown. It is not clear, for example, the extent to which they perceive song mirroring when a rival engages in this behavior, and if they can perceive it, the type of information that it conveys. This question can best be answered with interactive playback. Interactive playbacks involve the broadcast of phrase types in an interactive way, contingent upon the phrase types delivered by the subject. They have been successfully employed to study the songs of several species (Otter et al. 1999; Molles and Vehrencamp 2001a; Burt et al. 2001), often to establish the response of birds to being matched by a rival; a similar approach might yield useful information if applied to song mirroring in Cassin's Vireo. Relatedly, the occurrence of song mirroring in Cassin's Vireo begs the question of whether this type of singing behavior is widespread among songbirds, or peculiar to this species.

Second, the temporal organization of song warrants further attention. This is true with respect to the influence of timing on syntax, as mentioned above, and with regards to the temporal arrangement of song during counter-singing interactions. In a review of counter-singing behavior in songbirds, Todt and Naguib (2000) proposed that two primary categories of behavior govern the vocal interactions of birds: pattern-specific responses and time-specific responses. Pattern-specific responses are those in which the perception of a rival's song affects the acoustic structure (i.e. pattern) of a bird's response, a response category that includes both song matching and song advancing. In contrast, time-specific responses are those in which the perception of a song influences the timing of the response. This includes the tendency for birds to avoid temporal overlap in songs (Ficken et al. 1974), or in some cases the tendency to intentionally

overlap others (Mennill and Ratcliffe 2004), though the extent to which overlap is intentional is still debated (Searcy and Beecher 2011). Some studies have demonstrated an interaction between time- and pattern-specific responses. Falls (1985), for example, found that matching responses in Western Meadowlarks had different time-delays with respect to the playback than non-matching responses. Song matching also appeared to occur with particular time signatures in a study of Eurasian Blackbirds (Wolffgramm and Todt 1982). Time-specific responses have not been examined in Cassin's Vireo, but are likely to form another component of the song exchanges in this species.

A third aspect of song that remains to be explored is the influence of social context on phrase type delivery. Preliminary data on this topic suggests that Cassin's Vireos employ different phrase types near the nest than when far from the nest, and that the song of a neighbor may influence this pattern in complex ways. This tendency is particularly noteworthy immediately prior to exchange of incubation duties between males and females. During these exchanges, the male will often sing his 'nest phrases' while the female responds with nearly inaudible calls, which may imply a role for song in the coordination of nesting behavior, as has been hypothesized in other species (Halkin 1997; Boucaud et al. 2016).

Lastly, geographic patterns of song in this species are worthy of further study. My observations suggest the presence of vocal dialects in this species, such that birds in the Sierra Nevada sound quite different from birds in other parts of the breeding range. Examination of a few recordings dating back to the 1960's also suggests a remarkable stability to the structure of particular phrase types over time. The results from Chapter 4 clearly demonstrate that this phrase-type stability is important in communication in this species, but experimentation related to neighbor-stranger discrimination and perception of local and foreign dialects would clarify the



importance of these dialects. A more complete understanding of the migratory behavior of this species, which migrates to Mexico each winter, might further illuminate these dialect patterns with respect to the year-round movements of the birds.

To summarize, study of the songs of Cassin's Vireos has revealed complexity along various axes of song organization, and has pointed towards the ways that this complexity comes into play during vocal interactions between neighboring males. These results bring us closer to understanding the cause and effect between the perception of auditory stimuli and subsequent vocal responses. Still, there remain various aspects of song that are completely unexplored in this species, so topics for future research abound. With sufficient effort, a complete picture of the rich song system of Cassin's Vireo should emerge, providing additional insights into the functional importance of song complexity in communication in this species, and in animals more broadly.

### References

- Beecher M, Campbell S, Nordby J (2000) Territory tenure in Song Sparrows is related to song sharing with neighbours, but not to repertoire size. *Anim Behav* 59:29–37. doi: 10.1006/anbe.1999.1304
- Boucaud ICA, Valere PA, Aguirre Smith MLN, et al. (2016) Interactive vocal communication at the nest by parent Great Tits *Parus major*. *Ibis (Lond 1859)*. doi: 10.1111/ibi.12374
- Bradbury JW, Vehrencamp SL (2011) *Principles of Animal Communication*, 2nd ed. Sunderland, MA
- Briefer EF, Rybak F, Aubin T (2013) Does true syntax or simple auditory object support the role of skylark song dialect? *Anim Behav* 86:1131–1137. doi: 10.1016/j.anbehav.2013.09.019
- Burt JM, Campbell SE, Beecher MD (2001) Song type matching as threat: a test using interactive playback. *Anim Behav* 62:1163–1170. doi: 10.1006/anbe.2001.1847
- Burt JM, Vehrencamp SL (2005) Dawn chorus as an interactive communication network. In: McGregor PK (ed) *Anim. Commun. Networks*. Cambridge, UK, pp 320–343
- Byers BE, Kroodsma DE (2009) Female mate choice and songbird song repertoires. *Anim Behav* 77:13–22. doi: 10.1016/j.anbehav.2008.10.003

- Catchpole C (1986) Song repertoires and reproductive success in the Great Reed Warbler *Acrocephalus arundinaceus*. *Behav Ecol Sociobiol* 19:439–445.
- Cody ML, Stabler E, Sanchez HM, Taylor CE (2016) Structure, syntax, and “small-world” organization in the complex songs of California Thrashers (*Toxostoma redivivum*). *Bioacoustics*
- Falls JB (1985) Song matching in Western Meadowlarks. *Can J Zool* 63:2520–2524. doi: 10.1139/z85-373
- Ficken R, Ficken M, Hailman J (1974) Temporal pattern shifts to avoid acoustic interference in singing birds. *Science* (80- ) 183:762–763.
- Gentner TQ (2008) Temporal scales of auditory objects underlying birdsong vocal recognition. *J Acoust Soc Am* 124:1350–9. doi: 10.1121/1.2945705
- Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual vocal recognition in European Starlings, *Sturnus vulgaris*. *Anim Behav* 56:579–594. doi: 10.1006/anbe.1998.0810
- Halkin SL (1997) Nest-vicinity song exchanges may coordinate biparental care of Northern Cardinals. *Anim Behav* 54:189–98. doi: 10.1006/anbe.1996.0415
- Hiebert SM, Stoddard PK, Arcese P (1989) Repertoire size, territory acquisition and reproductive success in the Song Sparrow. *Anim Behav* 37:266–273. doi: 10.1016/0003-3472(89)90115-2
- Holland J, Dabelsteen T, Paris A (2000) Coding in the song of the wren: importance of rhythmicity, syntax and element structure. *Anim Behav* 60:463–470. doi: 10.1006/anbe.2000.1529
- Illes AE, Hall ML, Vehrencamp SL (2006) Vocal performance influences male receiver response in the Banded Wren. *Proc Biol Sci* 273:1907–12. doi: 10.1098/rspb.2006.3535
- James RD (1981) Factors affecting variation in the primary song of North American Solitary Vireos (Aves: Vireonidae). *Can J Zool* 59:2001–2009.
- Kershenbaum A, Blumstein DT, Roch M a., et al. (2014) Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol Rev* 91:13-52. doi: 10.1111/brv.12160
- Kramer H, Lemon R, Morris M (1985) Song switching and agonistic stimulation in the Song Sparrow (*Melospiza melodia*): five tests. *Anim Behav* 33:135–149. doi: 10.2173/bna.704
- Kramer HG, Lemon RE (1983) Dynamics of territorial singing between neighboring Song Sparrows (*Melospiza melodia*). *Behaviour* 85:198–223. doi: 10.1017/CBO9781107415324.004

- Mennill DJ, Ratcliffe LM (2004) Overlapping and matching in the song contests of Black-capped Chickadees. *Anim Behav* 67:441–450. doi: 10.1016/j.anbehav.2003.04.010
- Molles L (2006) Singing complexity of the Banded Wren (*Thryothorus pleurostictus*): Do switching rate and song-type diversity send different messages? *Auk* 123:991–1003.
- Molles LE, Vehrencamp SL (2001a) Songbird cheaters pay a retaliation cost: evidence for auditory conventional signals. *Proc Biol Sci* 268:2013–2019. doi: 10.1098/rspb.2001.1757
- Molles LE, Vehrencamp SL (2001b) Neighbour recognition by resident males in the banded wren, *Thryothorus pleurostictus*, a tropical songbird with high song type sharing. *Anim Behav* 61:119–127. doi: 10.1006/anbe.2000.1561
- Nowicki S, Searcy WA, Hughes M, Podos J (2001) The evolution of bird song: male and female response to song innovation in Swamp Sparrows. *Anim Behav* 62:1189–1195. doi: 10.1006/anbe.2001.1854
- Otter K, McGregor PK, Terry AMR, et al. (1999) Do female Great Tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proc R Soc B Biol Sci* 266:1305–1309. doi: 10.1098/rspb.1999.0779
- Potvin DA, Crawford PW, MacDougall-Shackleton SA, MacDougall-Shackleton EA (2015) Song repertoire size, not territory location, predicts reproductive success and territory tenure in a migratory songbird. *Can J Zool* 1–34.
- Reid JM, Arcese P, Cassidy ALE V, et al. (2005) Fitness correlates of song repertoire size in free-living Song Sparrows (*Melospiza melodia*). *Am Nat* 165:299–310. doi: 10.1086/428299
- Searcy WA, Beecher MD (2011) Continued scepticism that song overlapping is a signal. *Anim Behav* 81:e1–e4. doi: 10.1016/j.anbehav.2010.10.013
- Slabbekoorn H, Smith TB (2002) Habitat-dependent song divergence in the Little Greenbul: an analysis of environmental selection pressures on acoustic signals. *Evolution* 56:1849–58.
- Todt D, Hultsch H (1998) How songbirds deal with large amounts of serial information: retrieval rules suggest a hierarchical song memory. *Biol Cybern* 79:487–500. doi: 10.1007/s004220050498
- Todt D, Naguib M (2000) Vocal interactions in birds: The use of song as a model in communication. *Adv Study Behav* 29:247–296. doi: 10.1016/S0065-3454(08)60107-2
- Trillo PA, Vehrencamp SL (2005) Song types and their structural features are associated with specific contexts in the Banded Wren. *Anim Behav* 70:921–935. doi: 10.1016/j.anbehav.2005.02.004

- Vehrencamp SL, Hall ML, Bohman ER, et al. (2007) Song matching, overlapping, and switching in the Banded Wren: the sender's perspective. *Behav Ecol* 18:849–859. doi: 10.1093/beheco/arm054
- Vehrencamp SL, Yantachka J, Hall ML, de Kort SR (2013) Trill performance components vary with age, season, and motivation in the Banded Wren. *Behav Ecol Sociobiol* 67:409–419. doi: 10.1007/s00265-012-1461-x
- Wolffgramm J, Todt D (1982) Pattern and time specificity in vocal responses of blackbirds *Turdus merula*. *Behaviour* 81:264–286. doi: 10.1017/CBO9781107415324.004