IS THE DAWN CHORUS AN ADAPTATION TO WARM-UP THE VOICE IN ADELAIDE'S WARBLER (*SETOPHAGA ADELAIDAE*)?

JULEYSKA VAZQUEZ-CARDONA
Bachelor of Science, University of Puerto Rico, Mayagüez Campus, 2017

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

Department of Psychology
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

IS THE DAWN CHORUS AN ADAPTATION TO WARM-UP THE VOICE IN ADELAIDE’S WARBLER (*SETOPHAGA ADELAIDAE*)?

JULEYSKA VAZQUEZ – CARDONA

Date of Defence: November 18th, 2021

Dr. D. M. Logue Supervisor, Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. A. Iwaniuk Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. H. Jiskoot Thesis Examination Committee Member	Professor	Ph.D.
Dr. J.B. Leca Chair	Associate Professor	Ph.D.

ABSTRACT

Birds sing intensely during the early morning. I test predictions of the hypothesis that this phenomenon, known as the dawn chorus (*DC*), is an adaptation to warm-up the voice in male Adelaide's warblers (*Setophaga adelaidae*). I also analyze the effects of weather and date on the dawn chorus start time (*DCST*). I used songs from 29 males recorded in three different years and weather data from a MesoWest weather station. Song performance improves rapidly during the *DC*, but changes little after dawn. High song-rates lead to rapid improvements in vocal performance. Two out of three measures of performance peak at intermediate air temperatures (*Ta*). *Ta*, however, does not exert a strong effect on *DCST*. Males start to sing later with higher relative humidity, and earlier with increasing precipitation the previous day. This study lends support to the warm-up hypothesis and improves our understanding on the effects of weather on birdsong.

ACKNOWLEDGEMENTS

First, I would like to acknowledge that the University of Lethbridge, where I spent most of the time of my master's degree, sits upon the land of the Blackfoot Confederacy. I honor with gratitude the land itself and the Blackfoot / Niitsítapi people.

Not every student has the pleasure to meet good mentors. I do. Dr. David M. Logue is not only one of the greatest lecturers, presenters, and the most organized person I have ever met. He also cares about science, his research, and his students. And you may think that I am just talking about a supervisor that teaches his students the skills to succeed in the field, that meets with them frequently to discuss and review their projects, and that keeps them on track of every new finding in Animal Behaviour, which he does. But when I say that David is a great mentor, I also mean that he encourages us to have work-life balance, to work-out to be ready for the field, and to **stay hydrated!** Dr. David M. Logue gave me the opportunity to conduct a master's degree at a foreign country with a foreign language and helped me in every step of the way. I will always be grateful.

I would also like to acknowledge all the help I have received from my committee members, Hester Jiskoot and Andrew Iwaniuk. I am thankful that they accepted the invitation to be part of my thesis committee and did not hesitate to meet with me whenever I had questions. I appreciate all the knowledge and resources they have shared with me.

I acknowledge Peter C. Mower, Samantha W. Krouse, Liam Mitchell, Dawson Lehmann, Heath Petkau, Heatlie Lambert, Samantha Huang, and other members of the Birdsong lab for all the work they spent collecting and scoring the data I used in my analysis, and for their patience when training me to use the programs I needed. I am also thankful for all the good times and ideas we have exchanged. Being part of the Birdsong Lab has helped me to acquire skills that go beyond conducting research in Bioacoustics.

The Birdsong lab is not the only lab I am happy to have met. The multi-lab Animal Behaviour Research Group has been a great place to exchange ideas and to better understand what it means to be a scientist. Many members of this group have also helped me to complete my thesis project. I appreciate all the ideas and time I have received from the Barret-Henzi lab to conduct my analysis and to practice my presentation. I also acknowledge Jennifer Mather for the teaching skills, and for all the conversations and advice. Finally, I do not have enough words to explain how grateful I am for Tyler R. Bonnell, who taught me to conduct Bayesian statistical analyses and collaborated in my project in many ways. It has been a pleasure to meet such a great group of scientists.

There is a group of students that made me feel at home since the day one, and whom I consider friends. Francisco Jiménez-Gómez (mi compi), Peter C. Mower, Samantha Huang, Sydney Chertoff, Camilla Cenni, and many others: every coffee break, ride, hike, trivia night, and bonfire, have warmed-up my heart.

Finally, I would like to acknowledge the support I have received from my family and friends. My parents, grandparents, and Puerto Rican friends have kept me emotionally stable no matter how physically distant we are. To David Pere, I appreciate the emotional support and the help with R code, reviewing my writing, and making cute figures for my project. Les amo!

TABLE OF CONTENT

LIST OF TABLES

CHAPTER 2

LIST OF FIGURES

CHAPTER 2

RH is relative humidity, wind speed is the speed of wind, and Precipitation is precipitation accumulation. DCST is the time that the dawn chorus starts relative to the time of sunrise.

Figure 10. *Mixed-quantile regression (tau = 0.10) for the trade-offs between (a) note duration and gap duration, (b) note bandwidth and note duration, and (c) gap bandwidth and gap duration, that underlie the performance variables (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. Semi-transparent turquoise dots show the density of notes within songs from 2012. Purple dots indicate the density of notes within songs from 2017. 10th quantile regression lines (black) represent the estimates of performance limits with the intercepts and slopes that Logue et al. (2020) reported.*

Figure 11. *Distribution of the deviation scores for (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation from 2012 and 2017. Deviation scores were measured at the note level, and then averaged for each song.*

Figure 12. *Estimates of the effects of time relative to sunrise on the acoustic performance metrics (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation in male Adelaide's warbler song. Lower values of performance metrics indicate better performance. The black lines represent the estimate of the effect on the performance variable, the turquoise shaded areas represent the 95% confidence intervals of the estimate, and the semi-transparent purple dots represent songs. Estimates and confidence intervals were derived from Bayesian mixed models (see text).*

Figure 13. *Effects of air temperature on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

Figure 14. *Song order effects on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

Figure 15. *Effects of relative humidity on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

Figure 16. *Effects of Julian date on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

Figure 17. *Effects of time relative to sunrise on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. The black line represents the estimate of the effect on each acoustic trait, the turquoise shade represents the 95% confidence interval of the estimate, and the semi-transparent purple dots are songs. Estimates and confidence intervals were derived from Bayesian mixed models (see text).*

Figure 18. *Air temperature effects on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

Figure 19. *Song order effects on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

LIST OF ABBREVIATIONS

Bird ID	Bird identification
Bird ID / ST	Song-type within Bird identification
DAGs	Directed acyclic graphs
DCST	Dawn chorus start time
FEX	Frequency excursion
FM	Frequency modulation
Gap BW	Silent gap bandwidth
Gap duration	Duration of the silent gap
JD	Julian date (ordinal date)
MesoWest	MesoWest weather station
Note BW	Note bandwidth
Order	Song order
Precipitation	Precipitation accumulation
RH	Relative humidity
ST	Song-type
Ta	Air temperature
Tb	Body temperature
Time	Time relative to sunrise
WAIC	Watanabe-Akaike information criterion
Y	Generic dependent variable

CHAPTER 1: Introduction

Many songbird species sing intensely for a short period during the early morning. There are several hypotheses to explain the origin of this phenomenon, known as the "dawn chorus" (Staicer et al., 1996; Gil & Llusia, 2020). Some of these hypotheses propose that birds sing intensely at dawn because (1) sound propagates further at dawn (2) they have a surplus of energy that they did not spend during the night, (3) conditions are not optimal for foraging, (4) they are re-establishing their territorial claims after a night of unstable social dynamics, and (5) females prospect for mates at dawn (McNamara, et al., 1987; Henwood & Fabrick, 1979; Kacelnik, 1979; Staicer et al., 1996; Barnett & Briskie, 2006; Bradbury & Vehrencamp, 2011; Gil & Llusia, 2020). None of these hypotheses has emerged as a consensus explanation for the evolution of the dawn chorus. In this study, I will focus on a recent discovery that male Adelaide's warblers (*Setophaga adelaidae*) improve their vocal performance as they add more songs to the morning singing (Schraft et al., 2017). This finding suggests that the dawn chorus functions to improve vocal performance (Schraft et al., 2017). Testing this hypothesis is the main goal of this thesis.

The primary functions of male song in birds are mate attraction and territory defence (Catchpole & Slater, 2008). Selection may favor songs with high performance if they provide reliable information about the sender's condition or motivation to potential mates or rivals (Cardoso et al., 2007; DuBois et al., 2008; Irschick et al., 2008; Janicke et al., 2008; Cardoso et al., 2009; Byers et al., 2010; Bradbury & Vehrencamp, 2011; Botero & de Kort, 2013; Dinh et al., 2020). There is evidence that chestnut-sided warblers (*Setophaga pensylvanica*), dusky warblers (*Phylloscopus fuscatus*), swamp sparrows (*Melospiza georgiana*), and canaries (*Serinus canaria*) all prefer males that sing with high performance (Vallet & Kreutzer, 1995; Drăgănoiu et al., 2002; Forstmeier et al., 2002; Ballentine et al., 2004; Byers, 2007). Performance is also salient in male-male interactions. It often increases during real or simulated contests. Male nightingales (*Luscinia megarhynchos*), dark-eyed juncos (*Junco hyemalis*), skylarks (*Alauda arvensis*), and swamp sparrows respond with higher performance songs to playbacks that simulate male intrusions (Kunc et al., 2006; DuBois et al., 2008; Cardoso et al., 2009; Geberzahn & Aubin, 2014). During male-male interactions, males may be showing off their song performance to female eavesdroppers (Trillo & Vehrencamp, 2005; Kunc et al., 2006; Logue & Forstmeier, 2008) or signalling their motivation or resource holding potential to their opponent. Sexual or social selection favoring higher performance singing may have driven the evolution of male Adelaide's warblers' intense singing before dawn (Schraft et al., 2017).

1.1 Performance and motor constraints

When we listen to a saxophonist play a solo or watch a hockey player score a goal we may be impressed by their performance. In the human domain, performance is about "how well a person…does…an activity" (Cambridge University Press, n.d.). All athletes and musicians are trying to reach the finish line first, score the most goals, or impress the audience with their prowess, however, performance differs across individuals. One way to explain this variation is that individuals are subject to physiological constraints that limit their performance (Podos, 1996, Podos, 1997; Byers et al., 2010). The same kinds of constraints limit performance in non-human animals, including singing birds. If selection favors high song performance, it may evolve to approach one or more limits of performance. Some individuals get closer to the limit than others (Podos, 2001; Ballentine et al., 2004; Logue et al., 2020). It thus appears that the approach to a motor constraint represents a motor challenge, allowing us to define bird song performance as "the degree of challenge to the motor system, the respiratory system, or other physiological processes involved in singing" (Cardoso, 2017, p. e29).

Several systems are involved in sound production (e.g., HVC and RA regions of the brain, the respiratory system, the syrinx, and the vocal tract; Catchpole & Slater, 2008). There may be motor constraints in any of those systems or the coordination between them that will limit song production (Gil & Gahr, 2002; Podos, 2017; Logue et al., 2020, Podos & Sung, 2020). These constraints will generate trade-offs between acoustic traits and, hence, shape the structure of the songs (Podos, 1997; Hoese et al., 2000; Podos, 2001; Podos & Nowicki, 2004; Podos, 2017; Logue et al., 2020).

If a motor ability is constrained, acoustic properties that rely on that ability will trade-off against one another (Cardoso, 2017). For example, Podos (1997) established a metric known as "vocal deviation" based on an observed trade-off between trill rate and frequency bandwidth (Figure 1). At the limit of performance, sparrows can produce songs with high trill rates and narrow frequency bandwidths or low trill rates and wide frequency bandwidths, but they do not seem capable of singing high trill rates and wide frequency bandwidths at the same time (Figure 1; Podos, 1997). The speed with which a bird can modulate its vocal tract is limited, causing those parameters to trade-off at the limit of performance. Most studies of birdsong performance focus on these kinds of acoustic trade-offs rather than direct physiological observations (Ballentine, et al., 2004; Cardoso et al., 2007; Podos, 2017; Logue et al., 2020; Dinh et al., 2020; Podos & Sung, 2020). This approach has been used to study the relationship between song performance and the singer's quality, contextual variation in performance, and changes in performance over time, which is the topic of my thesis.

1.2 Warming-up the voice to improve performance

An individual's display performance can increase over both long and short timescales (Dinh et al., 2020). Over time frames of months and years, practice underlies both song learning and improvement in performance with age (de Kort et al., 2009; Vehrencamp et al., 2013, Zipple et al., 2019, Botero et al., 2009; Wada, 2010). Displays can also improve over shorter timescales. In many taxa, individuals display with a higher performance after they warm-up. In humans, singers warm-up to improve vocal performance, and athletes warm-up to improve athletic performance (Amir et al., 2005; Fradkin et al., 2010). There is now evidence that warming-up improves birds' display performance as well. Specifically, ruffed grouse (*Bonasa unbellus*) drumming rates increase with the cumulative number of displays during the day (Déaux et al., 2020), and vocal performance increases in swamp sparrows and Adelaide's warblers over the course of morning singing (Schraft et al., 2017; Dinh et al., 2020).

The mechanisms that underlie the improvement in performance with recent practice are rarely understood. Bishop (2003a, 2003b) hypothesizes that the warm-up period causes an improvement in performance due to mechanisms such as an increase in the speed of nerve conduction, faster metabolic reactions, an increase in oxygen consumption, etc. Some of these mechanisms are associated with muscle or core temperature (Bishop, 2003a; Bishop, 2003b). A more specific hypothesis is that the vocal performance benefits of warming-up can be attributed to an increase in temperature, and consequent reduction in viscous resistance of the muscles responsible for sound production (e.g., the vocalis muscle). Tests of this hypothesis in human beings are inconclusive due to variability across subjects (Elliot et al., 1995; Motel et al., 2003). Further, it is not clear whether the mechanisms that underlie warm-up in humans are comparable to those in birds (Dinh et al., 2020). Zebra finches (*Taenopygia guttata*) increase the fundamental frequency and amplitude of their songs during the morning singing (Wood et al., 2013). It is hypothesized that this increase in fundamental frequency and amplitude may be due to an increase in body temperature (*Tb)* that modulates the peripheral nervous system and the muscles involved in respiration, ultimately affecting song (Wood et al., 2013; Tan & Knight, 2018; Wu et al., 2021). Further investigation is required to describe the physiological causes of the improvement in performance.

There may be external factors that affect physiological components that influence performance. In humans (*Homo sapiens*), extreme air temperatures affect their exercise performance by impairing physiological responses such as oxygen uptake and heart rate (No & Kwak, 2016; Racinais et al., 2017). It is also necessary to further investigate the effects of extrinsic factors on the improvement in birdsong performance.

1.3 Effects on singing behaviour and song performance

Age and anatomical variation affect song performance. For example, banded wrens (*Thryophilus pleurostictus*) increase trill performance and consistency with age (de Kort et al., 2009; Vehrencamp et al., 2013). Dusky warblers (*Phylloscopus fuscatus*) are better at keeping a high amplitude when they are older (Forstmeier et al., 2002). The tropical mockingbird (*Mimus gilvus*) also produces more consistent syllable-types with age (Botero et al., 2009). Conversely, longitudinal studies in male swamp sparrows have shown that the quality of their songs – measured as song consistency and song-rate – increases until an intermediate age and declines afterwards (Zipple et al., 2019). Turning to morphological characters, bill morphology constrains the vocal performance of Darwin's finches (Family: *Thraupidae*) – species with large bills cannot change pitch as rapidly as species with small bills (Podos, 2001). Another recent study with tanagers (Family: *Thraupidae*), showed that body size, bill size, and bill shape impact song structure measures that may be associated with performance such as the peak frequency, frequency shift, note length, pause length, and trill rate (Demery et al., 2021). However, we still know very little about the effects of abiotic environmental variables on vocal performance.

Abiotic environmental variables can impact signalling across taxa. Changes in air temperature (*Ta*) and precipitation are well known to affect signalling in ectotherms, which depend on external heat to regulate their body temperature (*Tb*; Costanzo, 2011; Levesque & Marshall, 2021). For example, *Ta* positively affects the chirp rate and peak frequency of tropical crickets (*Acanthogryllus asiaticus*) as well as the speed, duration, and frequency of jumping spiders' (*Habronattus clypeatus*) signals (Brandt et al., 2018; Singh, R. et al., 2020).

Endothermic organisms depend on metabolic heat generation to regulate their *Tb,* but weather parameters can also affect their signals (Clarke & Rothery, 2008). For example, the great-leaf nosed bat (*Hipposideros armiger*) lowers the frequency of their echolocation calls when their *Tb* decreases with decreasing *Ta* (Wu et al., 2021). In bornean gibbons (*Hylobatidae muelleri*), males produce fewer solo vocalizations after rainfall. They also start calling later if it rained the previous night (Clink et al., 2020).

Weather also affects vocal behaviour in birds. A study of six species indicates that high *Ta* promotes earlier dawn chorus start times, and rain tends to delay the start of the dawn chorus (Bruni et al., 2014). House sparrows (*Passer domesticus*) also start singing earlier with increasing *Ta* (Hasan & Badri, 2016). Conversely, song thrushes (*Turdus philomelos*), cerulean warblers (*Setophaga cerulea*), house wrens (*Troglodytes aedon*), Alström's warblers (*Phylloscopus soror*), streak-breasted scimitar babblers (*Pomatorhinus ruficolli*), and brownish-flanked bush warblers (*Cettia fortipes*) all start singing later with increasing temperatures (Da Silva et al., 2014; Stuart et al., 2019; MacDonald & Islam, 2021; Puswal et al., 2021). Warbler species from the last example vocalize less when *Ta* increases (MacDonald & Islam, 2021; Puswal et al., 2021). Blackbirds (*Turdidae*), bulbuls (Family: *Pycnonotidae*), European robins (*Erithacus rubecula*), song thrushes, great tits (*Parus major*), blue tits (*Cyanistes caeruleus*), common chaffinches (*Fringilla coelebs*), and cerulean warblers start singing later when it rains (Hasan, 2010; Da Silva et al., 2014; MacDonald & Islam, 2021). Finally, the cerulean warblers decrease their singing rate when it is raining (MacDonald & Islam, 2021). Weather also affects the structure of birdsongs. Zebra finches decrease the duration of their syllables with high *Ta* and produce more similar, accurate, and less consistent songs at high *Ta* than at low *Ta* (Coomes & Derryberry, 2021). These examples show that weather, particularly *Ta* and precipitation, affect signalling behavior and song structure in birds.

There is, however, little evidence about the effects of weather on display performance in birds. The only such study that I am aware of addresses a non-vocal acoustic display. Déaux et al. (2020) found that overnight *Ta* modulates the performance of a non-vocal display in male ruffed grouse. The rate of this male's drumming display – in which the male repeatedly beats his wings, generating a loud sound – is lower in colder conditions (Déaux et al., 2020).

Empirical studies typically reveal that animals' performance peaks at some intermediate *Ta* and decreases at more extreme temperatures (shown in figure 1 in Schulte et al., 2011; Cunningham et al., 2013). This relationship holds true for both physiological performance (e.g., heart rate, enzyme activity) and whole-organism performance (e.g., running speed, fecundity, metabolic rate, signalling behaviour; Schulte et al., 2011). An example of peaked performance in a signalling animal comes from the jumping spider, which displays most vigorously at 40°C (Brandt et al., 2018). It is clear that peaked relationships between *Ta* and performance are common in ectotherms but little is known about the thermal performance curve in non-human endotherms (Levesque & Marshall, 2021). Levesque & Marshall (2021) proposed that the thermal performance curve in endotherms varies with context. They proposed various ways to characterize temperature-performance relationships in endotherms, including direct measurement of performance at different environmental temperatures. They also presented "thermoregulatory polygons" to represent this relationship in endotherms, that look similar to the performance curve in ectotherms (shown in figure 3 in Levesque & Marshall, 2021).

Most of these studies about the effects of weather on signalling behaviour are focused on birds in extreme temperatures. I will study the changes of song performance over time and the effects of weather on song performance and singing behaviour in Adelaide's warbler, a tropical bird in the Caribbean that does not typically experience extreme temperatures.

1.4 Study species

Adelaide's warbler is a tropical, year-round territorial and socially monogamous Passerine, endemic to Puerto Rico (Figure 2; Toms, 2020). It is primarily insectivorous, consuming Hemiptera, Coleoptera, and Lepidoptera (Toms, 2020). Their habitats include dry forests in Southwest Puerto Rico and in Vieques, Puerto Rico; wet limestone forests; and secondary forests on the North side of the island (Toms, 2020). Females rarely sing, but males sing at high rates throughout the year (Staicer, 1991).

Male Adelaide's warblers sing repertoires of about 30 song-types, all of which are frequency modulated trills (Figure 3; Logue et al., 2020). An individual's song-types can be divided into two categories, A and B. Males sing category A songs in repeat mode (i.e., same song-types), during the daytime, with long latencies between songs, and throughout the year. Conversely, males sing category B songs in switch mode (i.e., switching from a song-type to another), mostly before sunrise, with short latencies between songs, and only during the breeding season, which occurs from March through June (Staicer, 1991; Staicer 1996a; Staicer, 1996b). Song structures are similar among category A and category B songs and different males can use the same song-type in different categories (Staicer, 1996a; Kaluthota et al., 2019). Detailed comparison of A versus B songs, however, reveals subtle structural differences on average

(Staicer 1996a; Staicer, 1996b; Kaluthota et al., 2019). Kaluthota et al. (2019) found evidence that Adelaide's warbler songs may not belong to clear-cut categories.

A measure of motor performance known as frequency excursion (*FEX*), which measures the speed of both voiced and unvoiced frequency modulation (Podos et al. 2016), increases over the course of their morning singing. A linear model indicates that the cumulative number of songs that a male has sung is a better predictor of *FEX* than is the time of day. This finding suggests that male Adelaide's warblers warm-up their voices during the dawn chorus (Schraft et al., 2017).

Adelaide's warbler is an ideal species to test the hypothesis that the dawn chorus functions to warm-up the voice. Males sing frequency modulated trills that are difficult to produce, resulting in at least three kinds of trade-offs between acoustic traits (Logue et al., 2020). Males exhibit a pronounced dawn chorus during the breeding season (Staicer, 1991). Finally, most studies of vocal performance and singing behaviour focus on temperate species. By studying a year-round territorial, tropical species I hope to expand the scope of song performance research.

1.5 Wrapping - up

In summary, birdsong requires high levels of motor performance, singers vary in their ability to achieve high levels of performance, and signal receivers may respond to variation in song performance. Male Adelaide's warblers increase their vocal performance with recent practice but the role of the dawn chorus in warming-up remains unclear. We hypothesize that the dawn chorus is an adaptation to warm-up the voice (Schraft et al., 2017). The physiological mechanisms to warm-up the voice and the effects of external factors on vocal performance and singing behaviour remain unknown.

In this thesis, I characterize variation of male Adelaide's warblers singing performance over the course of the morning. I also consider the effects of weather parameters, *Ta* in particular, on song performance and singing behaviour. Based on the hypothesis that the dawn chorus functions to warm-up the voice, I predict that (1) performance increases more rapidly during the dawn chorus compared to the rest of the morning, and that (2) singing at a high rate will cause performance to increase. Based on the hypothesis that physiological warming is the mechanism of behavioural warm-up, I predict that (3) low air temperatures will be associated with low performance. Finally, I examine the effects of weather parameters on the time male Adelaide's warblers start to sing in the morning.

CHAPTER 2:

Is the dawn chorus an adaptation to warm-up the voice in Adelaide's warbler (*Setophaga adelaidae*)?

2.1 Introduction

2.1.1 Hypothesis that the dawn chorus is an adaptation to warm-up the voice

Birdsong functions in mate attraction and territory defense (Catchpole & Slater, 2003). The performance of this display may indicate the condition or motivation of the signaller (Cardoso et al., 2007; DuBois et al., 2008; Irschick et al., 2008; Janicke et al., 2008; Cardoso et al., 2009; Byers et al., 2010; Bradbury & Vehrencamp, 2011; Botero & de Kort, 2013; Dinh et al., 2020). Male Adelaide's warblers improve their song performance over the course of the morning (Schraft et al., 2017). There is a short period early in the morning known as the dawn chorus in which birds sing at high rates compared to the rest of the morning (Gil & Llusia, 2020). Schraft et al. (2017) showed a positive linear relationship between a measure of recent practice and a measure of song performance, but they did not exactly characterize how performance changes over the course of the morning. Therefore, Schraft et al. (2017) were not able to show if the improvement in performance is attributable to the dawn chorus *per se* or if performance increases gradually over the course of the whole morning. Nevertheless, they hypothesized that the dawn chorus is an adaptation to warm-up the vocal apparatus (Schraft et al., 2017).

In the present study, I describe changes of performance over the course of the morning and test predictions of the hypothesis that the dawn chorus functions to help birds warm-up. First, I predict that song performance will improve more rapidly during the dawn chorus compared to the rest of the morning. Second, I predict that high song-rates, like those that characterize dawn singing, will drive more rapid improvements in performance than low song-rates. To better describe the change of song performance over the course of the morning singing and test these predictions, I used a larger dataset, models that allow non-linear relationships (Mackenzie et al., 2005), and detailed measures of vocal performance.

2.1.2 Measuring performance

Birdsong studies have identified several traits that may represent vocal performance. One class of vocal performance metrics is based on theoretical expectations of constrained vocal performance (Cardoso et al., 2007; Cardoso et al., 2009; Podos et al., 2016). Podos et al. (2016) developed one such metric of birdsong performance called frequency excursion (*FEX*). *FEX* estimates the rate of frequency modulation, both voiced (i.e., in notes within songs) and unvoiced (i.e., in silent gaps between notes), in a song. Higher *FEX* is associated with rapid adjustment of the vocal apparatus (Podos et al., 2016). This is the measure that Schraft et al. (2017) used to demonstrate that male Adelaide's warblers warm-up during morning singing. The use of any single metric of vocal performance may not be optimal because avian vocal performance is limited by multiple motor constraints (Gil & Gahr, 2002; Podos, 2017; Logue et al., 2020, Podos & Sung, 2020, Goller, 2021). Specific limitations with *FEX* are that it lumps two potentially independent performance variables, voiced frequency modulation, and unvoiced frequency modulation, and does not seem to directly measure sound density. Sound density is the proportion of sound in a song. It is likely constrained by birds' ability to replenish their air supply while singing (Cardoso, 2009, Logue et al. 2020).

The other class of vocal performance metrics are based on observed constraints on sound production. Song production requires rapid, precise changes in the syrinx, respiratory system, and upper vocal tract (Nowicki et al., 1992, Podos & Nowicki, 2004; (Catchpole & Slater, 2008; Logue et al., 2020). Constraints on the speed, accuracy, or coordination of these changes constrain song structures (Podos, 1997; Hoese et al., 2000; Podos, 2001; Podos & Nowicki, 2004; Podos, 2017; Logue et al., 2020). Constrained attributes of bird song generate trade-offs in the acoustic properties that rely on them (Cardoso, 2017). Vocal deviation is an example of an empirically derived performance metric because it is based on observed evidence of constrained sound production (Podos, 1997). At the limit of performance, there is a trade-off between trill rate and frequency bandwidths. Birds do not seem capable of singing high trill rates and wide frequency bandwidths at the same time (Figure 1; Podos, 1997). This trade-off at the limit of performance may be due to a constraint in the speed that a bird can modulate its vocal tract (Goller, 2021).

Logue et al. (2020) identified three trade-offs at the note-level of a song in male Adelaide's warblers that appear to represent motor constraints. Deviation scores derived from these constraints measure vocal performance more precisely than vocal deviation or *FEX* because (1) they include a measure of sound density and (2) they separate the frequency modulation in notes from the frequency modulation in silent gaps. I refer to the deviation scores identified by Logue et al. (2020) as recovery time, voiced frequency modulation (*voiced FM*), and unvoiced frequency modulation (*unvoiced FM*).

Recovery time comprises a trade-off between the duration of a note (*note duration*) and the duration of the subsequent silent gap (*gap duration*; Figure 6; Logue et al., 2020). When singing a trill, birds typically take a "mini-breath" between each note (Hartley & Suthers, 1989). The trade-off between *note duration* and *gap duration* occurs because longer notes deplete the bird's air supply more than short notes, and thus require long subsequent silent gaps for the bird to replenish its air supply (Cardoso et al., 2007; Suthers & Zollinger, 2004). Similar constraints have been identified in serins (*Serinus serinus*), canaries, and dark-eyed juncos (*Junco hyemalis*; Hartley & Suthers, 1989; Mota & Cardoso, 2001; Cardoso et al., 2007). As silent gaps get shorter relative to preceding notes, the muscles involved in inspiration work harder (Wild et al., 1998). Logue et al. (2020) propose that deviation scores derived from the *note duration* / *gap duration* trade-off characterize the performance of those muscles, making recovery time a form of respiratory performance (discussion of deviation scores below).

A second acoustic trade-off in Adelaide's warbler songs emerges from a note-level comparison of frequency bandwidth (*note BW*) and duration (*note duration*). The deviation score derived from this trade-off is known as *voiced FM.* A third trade-off was observed between the frequency bandwidth in a silent gap (*gap BW*) and the duration of that gap (*gap duration*; Logue et al., 2020). The deviation score from that trade-off is called *unvoiced FM.* Logue et al. (2020) suggest that these two types of trade-offs represent a constraint on the speed that a bird can modulate the fundamental frequency, either due to constraints on the speed of one or more vocal organs, or a constraint on the birds' ability to coordinate these organs.

Performance is measured with the distance of each vocalization to a performance limit. The performance limit is first obtained with a regression line from the relationship between the acoustic traits that trade-off against one another. Podos (1997) estimated the performance limit with upper-boundary regression, but more recent work employs quantile regression for this purpose (Podos, 2001; Ballentine et al., 2004; Logue et al., 2020). Quantile regression models the relationship between an independent variable and a specified quantile of the dependent variable (Hao & Naiman, 2007). Performance is then measured as the orthogonal distance between a vocalization and an observed performance limit. This measure gives us a deviation score. Points on the low-performance side of the limit have positive values, and those on the high-performance side have negative values. Lower scores indicate higher performance (Figure 7; Podos, 1997).

2.1.3 Weather and the dawn chorus

Weather has been shown to impact singing behaviour in birds. Some birds modify the time they start to sing or their singing rates due to variation in weather parameters such as air temperature and precipitation (Bruni et al., 2014; Da Silva et al., 2014; Hasan, 2010; Hasan & Badri, 2016; Stuart et al., 2019; Coomes & Derryberry, 2021; MacDonald & Islam, 2021; Puswal et al., 2021). Most of these studies of weather effects on song focus on temperate or subtropical species. I will analyze the effects of weather on the singing behaviour and song performance in a tropical bird in the Caribbean, the Adelaide's warbler.

2.1.4 Objectives

In this study I attempt to improve on the description of changes in vocal performance during morning singing in male Adelaide's warbler, first discovered by Schraft et al. (2017). I will improve on the earlier description by using a larger dataset, running Bayesian regression models that incorporate splines, and analyzing changes in the fine-grained measures of vocal performance described by Logue et al. (2020; i.e., *recovery time*, *voiced FM*, and *unvoiced FM*).

In addition to describing the improvement in performance over the course of the morning, I test three predictions of the warm-up hypothesis. My first prediction is that vocal performance increases more rapidly during the dawn chorus than it does during normal singing later in the morning. My second prediction is that high song-rates result in higher vocal performance than do low song-rates. My third prediction is based on the relationship between *Ta* and performance. It states that low *Ta* is associated with low performance. Finally, I analyze the effects of day of year and weather parameters on the time that male Adelaide's warblers start to sing.

2.2 Methods

2.2.1 Study system

I studied a population of Adelaide’s warblers at the Cabo Rojo National Wildlife Refuge (US Fish and Wildlife Service), in western Puerto Rico (Latitude: 17.9752° N, Longitude: 67.1686° W; Figure 4). The refuge is a tropical-dry forest, with densely vegetated arroyos in a matrix of savannah-like open spaces. The dominant plants on Adelaide’s warbler territories include mesquite (*Prosopis juliflora*), tamarind (*Tamarindus indica*), Jerusalem thorn (*Parkisonia aculeata*), shrubs (*Lantana, Melochia, Waltheria*), vines (*Stigmaphyllon,* and *Tournefortia*), and epiphytes (*Tillandsia* spp.; Figure 5; Staicer, 1991).

This study adheres to the guidelines from the Institutional Animal Care and Use Committee at the University of Puerto Rico, Mayaguez (September 17, 2010) and the Animal Welfare Committee at the University of Lethbridge (protocol #1605). It also follows the ASAB/ABS Guidelines for the use of animals in research. Field work was conducted with permission from the U.S. Fish and Wildlife Service (permit 2012-01, 41521-2016-11) and the Departamento de Recursos Naturales y Ambientales (permits 2016-IC-068-1). Bird handling was conducted under David M. Logue’s master bird banding licence (no. 23969).

Trained ornithologists used acoustic lures to capture male Adelaide’s warblers in mist nests. Birds were fitted with unique combinations of plastic color bands and Fish and Wildlife Service numbered metal bands on their legs for identification. Subsequently, observers conducted focal recording on banded, mated males. They recorded a total of 29,952 songs during the periods of March 3 – June 19, 2012 (n = 9 males), April 13 – May 6, 2017 (n = 18 males), and March 12 – April 23, 2018 (n = 18 males). Twelve of the eighteen birds recorded in 2018 were the same that were recorded in 2017 and two were recorded in 2012 (see Appendix 1). I used song recordings from 2012, 2017, and 2018 for the dawn chorus start time analysis. Due to logistical constraints, I only used songs from 2012 and 2017 for the song performance analysis (see Appendix 1).

All recording periods were within this population's breeding season (i.e., from March to June; Staicer, 1991; Staicer 1996a; Staicer, 1996b). Recordings were made with Marantz PMD 661 digital recorders and Sennheiser ME67 shotgun microphones (file format = wav, sampling rate = 44.1KHz, bit depth = 16 bits). Birds were recorded continuously from 45 minutes before sunrise until approximately 2.5 hours after sunrise. The recordist visually confirmed the identity of the focal male by the end of each recording session.

In 2012, recordists attempted to record each male for four days with at least four days between recording sessions. Due to logistical constraints, however, some individual males were recorded for two sequential days on two separate occasions. In 2017, recordists attempted to record each male for three days with at least two days in between but on two occasions they recorded individuals on consecutive days and one bird was recorded for one day. In 2018, recordists attempted to record each mated male for at least two days. Seven birds, however, were recorded one day. They also attempted to record the birds with at least two days in between but an individual male was recorded for sequential days on one occasion.

2.2.2 Annotation

Trained observers annotated the field recordings. For the 2012 recordings, annotators visualized recordings in Syrinx PC v.2.6 (J. Burt, Seattle, WA, http://www.syrinxpc.com/; Settings: Blackman window, transform size = 1024 points), and entered data into an Excel (Microsoft Corp., Seattle WA) spreadsheet. Each row represented a five second time bin (e.g., 05:36:00, 05:36:05, 05:36:10, …). Data included all songs from the focal birds, and the song-types to which they belonged. Annotators scored song-types by comparing each song spectrogram to spectrograms of the known repertoire of the focal bird. When an annotator encountered a song-type that did not match any song-types from the focal male's repertoire, a high-quality recording of that song was added to the repertoire file. After all scoring was complete, one observer (David M. Logue) reviewed all song-types and re-scored misclassified songs.

Our lab switched to another software to annotate song recordings, so the recordings from 2017 were annotated in Raven Pro 1.6.1 (Center for Conservation Bioacoustics, 2019). Five trained observers listened to the recording, including comments from the recordist, to detect the focal bird's songs. The initial round of song-type scoring followed the same protocol as in 2012. After all scoring was complete, one observer (Peter C. Mower) reviewed all song-types and re-categorized misclassified songs for the 2017 recordings. The 2018 recordings were scored similarly to the 2017 recordings, but the only data I used from 2018 are the dawn chorus start times.

The times of the songs from 2017 and 2018 are more precise than the ones from 2012 because they show the exact time indicated in the recorder (i.e., with hour, minutes, and seconds). The song times in the 2012 dataset were scored in five second increments (e.g., 05:36:00, 05:36:05, 05:36:10, …). I binned times from the 2017 and 2018 datasets into 5 second intervals so that time data were equally precise in 2012, 2017, and 2018.

2.2.3 Song performance

Song performance analysis was based on subsamples of the 2012 and 2017 song recordings. Other workers in my lab conducted the acoustic analyses on the 2012 recordings (Logue et al. 2020), whereas I personally analyzed the 2017 recordings. Of the 9,499 song recordings from 2012, I used all 2,879 song recordings with signal-to-noise ratios that were sufficiently high to permit detailed acoustic analysis (Logue et al. 2020). I used stratified random sampling to choose songs from the 2017 recordings. I first identified songs with high signal-to-noise ratios from the 18 mated males. Then, I divided the recordings into one-hour bins, beginning at the time of the focal male's first song. For example, if a male started to sing at 5:20:00, his first bin would begin at 5:20:00 and end at 6:20:00, his second bin would span 6:20:00 to 7:20:00, and so on. I then selected a random sample of 10 songs within each time bin. This resulted in a sample of 915 songs from the 2017 data.

There were two birds in 2017 (WTbTb and RbTbL) from which I did not have an adequate sample of high-quality songs (i.e., high signal-to-noise ratios). I excluded those two birds from the sample for the song performance analysis. So, I ended up with 888 songs from 2017.

I used Luscinia v.2.14 to measure songs' acoustic characteristics (Lachlan, 2007; Settings: max. freq. = 10kHz, frame length = 5ms, time step = 1ms, dynamic range = 35 dB, dynamic equalization = 100ms, de-reverberation = 100%, de-reverberation range = 100ms, high pass threshold = 1.0kHz, noise removal = 10dB). To enter songs into Luscinia, a scorer highlights the sound on a spectrogram, using a touchscreen monitor and a stylus (Lachlan, 2015; Logue et al., 2020). Once the sound is highlighted, its acoustic data can be exported to a spreadsheet. For this study, we extracted the time that each note begins and ends, the minimum and maximum peak frequency of the note, and the peak frequency at the beginning and end of each note. We chose peak frequency instead of fundamental frequency because visual inspection indicated that the peak frequency function tracks the note's fundamental frequency more consistently than does the fundamental frequency function.

I used the note structure data from Luscinia to calculate composite acoustic traits and performance variables. The four composite acoustic traits used in this analysis are *note duration*, *gap duration*, *note BW*, and *gap BW*. To calculate *note duration* and *gap duration*, I used the length of notes in milliseconds (ms) and the length of the spaces between notes (ms), respectively (Figure 6).

I used frequency ratios to calculate the *note BW* and *gap BW* (Cardoso, 2013). This approach to calculate frequency bandwidth with a logarithmic scale (i.e., frequency ratio) is a better biological representation of sound production and perception than a linear scale (i.e., frequency average or frequency difference; Cardoso, 2013). I divided the maximum peak frequency by the minimum peak frequency to determine the *note BW*. The *gap BW* was calculated as the ratio of the peak frequency at the end of one note and the beginning of the next note, with the higher value in the numerator (Figure 6). I excluded the final note of each song because it would not be possible to calculate *gap BW* without the start of a subsequent note.

Next, I calculated deviation scores for each of the three performance metrics, for each note. I based my calculation of performance metrics on performance limits described by Logue et al. (2020). That study ran three mixed-quantile regressions: *note duration* versus *gap duration* (*recovery time*), *note BW* versus *note duration* (*voiced FM*), and *gap BW* versus *gap duration* (*unvoiced FM*). For all three comparisons the variable *tau*, which specifies the quantile, was set to 0.10 (i.e., the tenth quantile), and bird identification (*Bird ID*) was included as a random variable. The mixed-quantile regressions resulted in lower boundaries with positive slopes that represent an estimate of a performance limit (Table 2.1; Logue et al. 2020). I used the quantile regression lines from Logue et al. (2020), which were based on the 2012 recordings, to calculate deviation scores for both the 2012 and 2017 recordings.

The deviation score for a given note was calculated as the orthogonal distance between the note and the performance limit. Notes beyond the line have negative deviation scores. Lower deviation scores indicate higher performance (Figure 7; Podos, 2001).

I averaged the composite acoustic traits and deviation scores over all notes within a song to get the average measurements at the song-level. I used these song-level values for all further analyses of acoustic structure because I am interested in understanding how song performance changes over the course of the morning, not within each song.

2.2.4 Dawn chorus start time

Dawn chorus start time (*DCST*) analysis was based on the first song that a focal male sang each morning they were recorded during the breeding season of 2012, 2017, and 2018. I excluded recording sessions where the observers indicated that the focal bird started to sing before they started to record. The dataset for this analysis comprised 29 unique males and 125 dawn chorus start times. The *DCST* is in seconds relative to sunrise. I got sunrise times in Atlantic standard time from http://www.dateandtime.com/.

2.2.5 Weather

My primary source of weather data was a MesoWest weather station located at the field site (Latitude: 17.97 ° N, Longitude: 67.16 ° W; Figure 4; The University of Utah, 2021). It has an elevation of 33 m and a distance from the coast of 1.3 km. That station provided hourly readings of air temperature (*Ta*; °C), relative humidity (*RH*; %), wind speed (m/s), and precipitation (mm).

I used data from a complementary weather station from the National Oceanic and Atmospheric Administration (*NOAA*) in Lajas, Puerto Rico (Latitude: 18.10 ° N, Longitude: 67.10° W (National Oceanic and Atmospheric Administration, n.d.) to test whether the data from the principal weather station were reliable. The distance between the *NOAA* weather station and the coast is 6 km. It has an elevation of 27 m.

I ran a Pearson correlation test to compare the daily mean *Ta* between the two weather stations over the recording periods. *Ta* is the principal weather parameter I am testing and the only one in common between the databases. This test indicated a strong correlation (r = 0.83 in 2012, r = 0.77 in 2017, and r = 0.84 in 2018), between the MesoWest and *NOAA* data. I also visually showed that, apart from a systematic difference of 1-2 °C between the two stations, the daily patterns are the same (see Appendix 2). This difference in 1-2 °C may be because one weather station (*MesoWest*) is closer to the coast than the other one (*NOAA*) and because they are in different elevations. Based on this finding, I conclude that the *MesoWest* weather data are reasonable. I therefore used the *MesoWest* output for all weather data in this thesis.

To analyze the effects of weather on the vocal performance of Adelaide's warblers, I used the main weather parameters that have been shown to affect thermoregulation in endotherms – *Ta* and *RH* (McKechnie et al., 2017; Levesque & Marshall, 2021). The weather station reports *Ta* and *RH* data once per hour, so I weight-smoothed these parameters to increase accuracy between hourly readings. For each song, I calculated the time since the last reading and the time until the next reading and divided both by one hour to generate proportions. I then multiplied the larger of these two proportions by the weather parameter reading that was closer to the time of the song recording, multiplied the smaller proportion by the weather parameter reading that was farther from the time of the song recording, and summed the weighted values.

To analyze the effects of weather on DCST, I used *Ta*, *RH*, *wind speed*, and precipitation accumulation (*Precipitation*). I used their values at the closest previous time they were measured to the time of the song. For example, if a bird started to sing at 05:05:00, I used the weather measure reported at 04:57:00 in the *MesoWest* database.

2.2.6 Analysis software

I used R x64 4.1.0. (R Core Team, 2018) and the packages *ggplot2*, *ggpubr,* and *seewave* to create data visualizations and sound spectrograms (Sueur et al., 2008; Wickham, 2016; Kassambara, 2020; Araya-Salas, 2021). I also conducted the statistical analysis in R x64 4.1.0. (R Core Team, 2018) following Richard McElreath's workflow for Bayesian analysis (McElreath, 2019). I used the packages *mctest* to test multicollinearity (Imdadullah et al., 2016; Imdad & Aslam, 2018), *DAGitty* to create directed acyclic graphs (Textor & Van der Zander, 2016), and *brms* to fit the Bayesian mixed-effects models (Bürkner, 2018).

2.2.7 Statistical analysis: song performance

Directed acyclic graphs (*DAGs*) are networks that represent patterns of causal influence. Their nodes represent variables and their directed edges (arrows) indicate the direction of causal influence (McElreath, 2019). I used *DAGs* to specify my causal hypothesis and avoid statistical confounds (Westreich & Greenland, 2013; McElreath, 2019).

I made *DAGs* for each of the performance metrics (*recovery time, voiced FM,* and *unvoiced FM*), and the acoustic traits that underlie them (*note duration, note BW, gap duration,* and *gap BW*). These *DAGs* included the weather variables *Ta* and *RH*, Julian date (*JD*) to control for variation in performance attributable to time of year, the "cumulative number of songs that the focal bird had sung during the morning" (*Order*), which is known to correlate with performance (Schraft et al., 2017, p. 171), and time relative to sunrise (*Time*).

I designed the *DAG* to reflect my assumptions about the system (Figure 8). A description of all the variables used in this analysis can be found in Table 2.5. I assumed that date would influence air temperature, humidity, the number of songs that the birds sang, and the performance of their songs (visualizing changes of weather across the breeding season; Bruni et al., 2014; Da

Silva et al., 2014; Puswal et al., 2021), so I made *JD* affect *Ta*, *RH*, *Order*, and the dependent variable (*Y*) in the *DAG*. Based on the assumption that air temperature influences humidity, the number of songs that the birds sang, and the performance of their songs (Oke, 1978; United States Environmental Protection Agency, 2012, August; US Department of Commerce, N.O.A.A., 2013, June 21; US Department of Commerce, N.O.A.A., 2015, June 13; Reeping & Hemenway, 2020; MacDonald & Islam, 2021; Puswal et al., 2021), I made *Ta* affect *RH*, *Order*, and the dependent variable (*Y*) in the *DAG*. I assumed that humidity influences the number and performance of the songs that the birds sang (Clink et al., 2020; MacDonald & Islam, 2021), so I made *RH* affect *Order* and *Y*. The number of songs that the birds sang has been shown to influence song performance (Schraft et al., 2017; Dinh et al., 2020), so I made *Order* affect *Y* in the *DAG*. Finally, I assumed that time would influence the air temperature, humidity, the number of songs that the birds sang, and the performance of their songs (visualizing changes of weather during the day; Schraft et al., 2017), so I made *Time* affect *Ta*, *RH*, *Order,* and *Y* in the *DAG*.

I tested for conditional dependencies with *impliedConditionalIndependencies* from *DAGitty*. This function uses *DAGs* to identify conditional independencies. We can use our data to confirm those independencies. The results indicated that *JD* and *Time* were independent from each other (JD _||_Time). I ran a linear mixed-effects model with *JD* as a response variable, *Time* as a predictor variable and *Song-Type* within *Bird ID* as a nested cluster (i.e., random effects) to test that independency. The estimate of the effect from *Time* to *JD* was -4.55 (p = <2e-16). I added the causal influence from *JD* to *Time* in the *DAG*. Then, I used *adjustmentSets* from *DAGitty*. This function helps to determine which independent variables to include in a model to estimate the direct effects of *Ta* on the response variable and avoid statistical confounds. Following the advice of *DAGitty* ensures that every causal path from *Ta* to the response variable is blocked, other than the direct one (i.e., arrow from *Ta* to *Y* in Figure 8*;* Westreich &

Greenland, 2013; McElreath, 2019). The results from *adjustmentSets* indicated that it was appropriate to include all the independent variables in the model: *Ta*, *JD*, *Time*, *Order*, and *RH*.

In summary, I used a *DAG* to represent causal influences between the variables used in the song performance analysis. I, then, checked conditional independencies between the variables and tested them with my data. When the *DAG* included all the causal influences necessary in the system, I used *adjustmentSets* to determine what independent variables to include in my models to avoid confounds.

I used variance inflation factor (*VIF*) analysis to test for multicollinearity among the predictor variables. I fitted a linear model with the functions *lm* and included the performance variables and constituents of their trade-offs as the outcome variables and *Ta*, *JD*, *Order*, *RH*, and *Time* as predictor variables. Then, I used the function *imcdiag* from the package *mctest* to determine the *VIF* (Imdadullah et al., 2016; Imdad & Aslam, 2018). *VIF*'s > 5 indicate high multicollinearity. The *VIF* values for the predictor variables in the song performance analysis were as follows: *Ta* = 3.65, *JD* = 2.64, *Order* = 2.03, *RH* = 2.13, *Time* = 4.57. These values indicate that the predictor variables exhibit low-to-moderate correlation, so I retained them in the analysis.

2.2.8 Constructing and selecting the model for song performance

Bayesian models require users to input prior estimates of model parameters. I mean-centred and standardized all variables with the *scale* function in R x64 4.1.0. (R Core Team, 2018) and used weakly informative priors centered on 0. These priors set the models' initial state to assume that the independent variable has no effect on the dependent variable*;* Appendix 3; McElreath, 2019).

I constructed models at different levels of complexity (i.e., linear mixed-effects model, quadratic mixed-effects model, and spline mixed-effects-model) with a Gaussian distribution using the *brms* package (Bürkner, 2018). I used the Markov Chain Monte Carlo (*MCMC*) conditioning engine with four chains and 3,500 iterations to estimate posterior distributions. Based on the results of my analysis of the song performance *DAGs*, I included *Ta*, *RH*, *JD*, *Time*, and *Order* as fixed effects. I treated *Year* and song-type (*ST*) nested within *Bird ID* as clusters (i.e., random variables) to control for the non-independence of performance within year, song-type, and individual. *ST* were nested within individuals because the acoustic structure of shared song-types often varies consistently among individuals.

I identified four biologically realistic two-way interactions, which I included in the initial performance models. I included the interaction between *Ta* and *RH* because temperature and humidity interact to affect thermoregulation in endotherms (Gerson et al, 2014; Levesque & Marshall, 2021). I included the interaction between *Time* and *Order* to account for a possible effect of song-rate on performance. The third interaction is between *Ta* and *Time*. That interaction addresses the possibility that performance changes over time depend on the temperature (e.g., birds may warm-up faster when the air is warm). Similarly, I included the interaction of *Ta* and *Order* to account for the possibility that temperature mediates the relationship between the amount of recent practice and performance (Déaux et al. 2020). Finally, I considered a three-way interaction between *Ta*, *Time*, and *Order* following the same assumptions as in the two-way interaction between *Time* & *Order*, *Ta* & *Time*, and *Ta* & *Order*.

I chose the kind of model (i.e., linear, quadratic, and spline) that best fit the data by comparing the Watanabe-Akaike information criterion (*WAIC*), where lower values mean better fit (Table 2.2). The spline mixed-effects-model had the lowest *WAIC* by far for all models. This model has higher degrees of freedom than linear or quadratic mixed-effects models. It creates data bins and fit a regression within each bin (Mackenzie et al., 2005). I also used *WAIC* to choose which interactions to include in the final models (shown in Table 2.3). I compared the *WAIC* of the combination of interactions for each of the performance variables (i.e., *recovery time*, *voiced FM*, and *unvoiced FM*) as outcome variables. The best model fit was the same for the three outcome variables (Table 2.3). I also applied this model to the each of the acoustic traits that underlie trade-offs for consistency in the analysis.

The following equation represents the best fit models, where "t2" represents the interaction between predictor variables with splines and "s" represents individual predictor variables with splines:

$$Y \sim \text{t2}(\text{Ta.scaled}, \text{RH.scaled}) + \text{t2}(\text{Order.scaled}, \text{Time.scaled}) + \text{t2}(\text{Ta.scaled}, \text{Time.scaled}) + \text{t2}(\text{Ta.scaled}, \text{Order.scaled}) + \text{s}(\text{JD.scaled}) + (1|\text{Year}) + (1+\text{BirdID}/\text{ST})$$

Another method that I used to test if the models fit the data well is the posterior predictive check with the function *pp_check* from the *brms* package (Bürkner, 2018). This function compares the distribution of the observed data (*Y* in Appendix 5) with samples of the posterior predictive distribution (i.e., replicated data; *Yrep* in Appendix 5). If *Y* and *Yrep* show a similar distribution, it means that the model fitted the observed data well (Gelman & Hill, 2018). The posterior predictive distributions for all the performance variables and acoustic traits showed a good model fit (see example in Appendix 5).

Finally, I examined *R-hat* values to determine if the *MCMC* chains converged. As a rule of thumb, values lower than 1.05 indicate sufficient convergence (Gelman & Shirley, 2011). The *R-hat* values for each of the predictor variables and interactions in the selected model (i.e., with the lowest WAIC) were lower than 1.05, so I concluded that the chains converged.

2.2.9 Interpreting the results in song performance

I used two methods to interpret the results. First, I used *conditional_effects* from the package *brms* (Bürkner, 2018) to examine the model parameters. Second, I used the conditioned models to make predictions based on simulated data. In complex models with splines and interactions, the model parameters are difficult to interpret. The predictions paint a clearer picture of the consequences of varying the independent variables.

I made simulated data with different song-rates and ran it through the song performance models to better understand how rate affects performance. I made a dataframe in which simulated "birds" sang at one of three constant song-rates: one song every 30 seconds, every 60 seconds, or every 600 seconds in the time range of the whole real dataset (i.e., from 2,460 seconds before sunrise to 12,860 after sunrise). These song-rates were arbitrary but fall within the species' natural range (unpublished data). I set the other variables (i.e., *Ta*, *RH*, and *JD*) to their average values and averaged over all levels of the random variables. I then used the *fitted* function to generate model predictions from the simulated dataframes. I visualized those predictions in *ggplot2* (Wickham, 2016).

Similarly, I simulated data with different *Ta* and ran it through the models to better understand how variation in *Ta* affects performance. Low *Ta* was defined as the 5th percentile of *Ta*'s in the whole dataset (19.44 °C), median was the 50th percentile (24.49 °C), and high was the 95th percentile (27.78 °C). I set *RH* and *JD* to their average values. I made a dataframe with real values of *Time* and *Order* for each of six different birds (i.e., three from 2012 and three from 2017). I selected those birds based on a minimum, median, and maximum number of songs (i.e., sample size) in each year. All six birds showed a similar pattern on the effects of *Ta* on performance. I, then, only showed visualizations with real values of *Time* and *Order* from one of the six birds. Finally, I used the *fitted* function to generate model predictions, averaging over all levels of the random variables, and visualized those predictions with *ggplot2* (Wickham, 2016).

2.2.10 R-squared and variance in song performance

I evaluated the variance (σ^2) associated with each of the random effects as well as the model and parameter R-squared (R^2) values to better understand the influence of the parameters on the model. I used conditional and marginal R^2 values to indicate the proportion of variance explained by the whole model and by predictor variables, respectively.

2.2.11 Statistical analysis: dawn chorus start time

The *DAG* for dawn chorus start time included *Ta*, *RH*, *wind speed*, *Precipitation, and JD* (Figure 9). A description of all the variables used in this analysis can be found in Table 2.5. I assumed Julian date would influence air temperature, humidity, precipitation, wind speed, and the time the birds start to sing (visualizing changes of weather across the breeding season; Bruni et al., 2014; Da Silva et al., 2014; Puswal et al., 2021), so I made *JD* affect *Ta*, *RH*, precipitation accumulation (*Precipitation*), wind speed, and *DCST*. Based on the assumption that the air temperature influences humidity and the time the birds start to sing (Oke, 1978; United States Environmental Protection Agency, 2012, August; US Department of Commerce, N.O.A.A., 2013, June 21; US Department of Commerce, N.O.A.A., 2015, June 13; Reeping & Hemenway, 2020; MacDonald & Islam, 2021; Puswal et al., 2021), I made *Ta* affect *RH* and *DCST*. Precipitation accumulation has been shown to affect air temperature, humidity, and the time the birds start to sing (Oke, 1978; Hasan, 2010; Da Silva et al., 2014; Bruni et al., 2014; Clink et al., 2020; MacDonald & Islam, 2021), so I assumed that *Precipitation* affects *Ta*, *RH*, and *DCST*. I assumed that relative humidity would affect the dawn chorus start time and precipitation accumulation (Oke, 1978; Hasan, 2010; Da Silva et al., 2014; Bruni et al., 2014; Clink et al., 2020; MacDonald & Islam, 2021), so I made *RH* affect *DCST* and *Precipitation*. Finally, I made *wind speed* affect *DCST* based on the assumption that it affects the time birds start to sing (Puswal et al., 2021).

The output of *impliedConditionalIndependencies* indicated that *precipitation accumulation* (*Precipitation*) and *wind speed* were independent given *JD (*Precipitation _||_ wind speed | JD*)*. I ran a linear mixed-effects model with *Precipitation* as the response variable, *wind speed* and *JD* as predictor variables and *Bird ID* as a random effect. The effect from *wind speed* to *Precipitation* was low and non-significant (0.02, p = 0.53), so my assumption was correct based on the dataset. The output of *impliedConditionalIndependencies* also showed that *RH* is independent from *wind speed* given *JD* (RH _||_ wind speed | JD). I followed the same method to test this independence. The effect from *wind speed* to *RH* given *JD,* was negative and significant (-0.2, p < 0.0001). Finally, *Ta* and *wind speed* seem to be independent from each other given *JD* (Ta _||_ wind speed | JD). The effect from *wind speed* to *Ta* is positive and significant (0.5, p < 0.0001). However, the last two relationships are not meteorologically realistic for this system so I will not include a causal influence from *wind speed* to *Ta* in the *DAG,* nor from *wind speed* to *RH* (Oke 1978).

The results from the function *adjustmentSets* from *DAGitty* (Textor & van der Zander, 2016) indicated that it was appropriate to include *JD, Precipitation,* and *RH* as covariates, to avoid confounds when estimating the effect from *Ta* on *DCST*.

The *VIF* values for the predictor variables in the *DCST* analysis were as follows: *Ta* =1.60, *JD* =15.01, *Precipitation* =13.84, *RH* =1.11. There was a high correlation between *JD* and *Precipitation*, so I tested if excluding one of these predictor variables changed any of the effect estimates. Including both did not have a major effect on any of the predictor variables, so I kept both variables in the model.

2.2.12 Constructing and selecting the model for dawn chorus start time

I mean-centred and standardized all predictor variables with the *scale* function in R x64 4.1.0. (R Core Team, 2018) and used weakly informative priors centered on 0. These priors set the models' initial state to assume that the predictor variables have no effect on *DCST* (Appendix 4; McElreath, 2019).

For the analysis of *DCST*, I fitted a spline mixed-effects model with a Gaussian distribution (Bürkner, 2018). I used the Markov Chain Monte Carlo (*MCMC*) conditioning engine with four chains and 3,500 iterations to generate posterior distributions. Based on the results of the *DAG* analysis and the multicollinearity test, I included *Ta*, *JD*, *Precipitation*, and *RH* as fixed effects. I also included *Year* and *Bird ID* as clusters (i.e., random effects) to control for the potential non-independence of data within individuals and years. Finally, I included random slopes to test the effect of *Ta* within *Bird ID*.

Because the effect of weather on dawn chorus start times may depend on date (Bruni et al., 2014; Puswal et al., 2014), I considered all possible combinations of two-way interactions between *Ta* and *RH, Precipitation,* and *JD,* a three-way interaction between *Ta*, *RH,* and *Precipitation* and a four-way interaction between *Ta*, *RH*, *Precipitation,* and *JD.*

I chose the model that fitted the data better comparing the *WAIC*, where lower values mean better fit (Table 2.4). The following equation represents the best fit model, where "t2" represents the interaction between predictor variables with splines and "s" represents individual predictor variables with splines:

$$\text{DCST} \sim \text{t2(Ta.scaled, Humid.scaled)} + \text{s(Precipitation.scaled)} + \text{s(JD.scaled)} + \text{(1+Ta.scaled|BirdID)} + \text{(1|Year)}$$

The distributions from the *pp_check* in the *DCST* analysis showed that, even though the samples of the posterior predictive distribution (*Yrep*) do not perfectly fit the observed data (*Y*), they follow almost a same pattern (Appendix 6). The model fits the data well.

The *R-hat* values for each of the predictor variables and interactions in *DCST* were lower than 1.05, so I concluded that the MCMC chains converged (Gelman & Shirley, 2011).

2.2.13 Interpreting the results in dawn chorus start time

I interpreted results in *DCST* with the same two methods that I used in the song performance analysis. I used plots of conditional effects to describe the independent effects of predictor variables on *DCST*. I also used the conditioned models to make predictions based on simulated data.

I simulated data with different temperatures and ran it through the models to better understand how variation in *Ta* influences the time birds start to sing. For these dataframes, I included the mean *RH*, mean *Precipitation* and the values of *Julian* from the whole *DCST* dataset. I used three *Ta's* based on the 5th percentile (20.00°C), 50th percentile (23.33°C), and 95th percentile (25.56°C) temperatures. Finally, I used the *fitted* function, to generate model predictions with an average over all random variables, and visualized those predictions in *ggplot2* (Wickham, 2016).

2.3 Results

2.3.1 Song performance analysis

The descriptive statistics for the performance dataset are shown in Table 2.6. The dataset includes 22,574 notes from the 2017 recordings in addition to the 68,602 notes from the 2012 dataset, which were analyzed previously by Logue et al. (2020). Scatterplots of the 2012 and 2017 data showed similar patterns. Both the 2012 and 2017 data indicate linear lower boundaries, except for a prominent bulge in the lower left part of the boundary on *recovery time* (Figure 10 (a-c); Logue et al., 2020).

The distribution of the deviation scores were all moderately right-skewed (Figure 11 (a-c)).

The estimated conditional effects from the Bayesian mixed-effects models on song performance variables indicated that *recovery time* and *unvoiced FM* improved during pre-dawn singing and then worsened after sunrise. *Voiced FM* improved during the pre-dawn period, held steady after sunrise, and then improved again beginning at 5,000 seconds after sunrise (Figure 12 (a-c)). Estimated conditional effects also showed that male Adelaide's warblers' vocal performance was at its worst first thing in the morning but improved as the birds elevated their singing rate before sunrise (Figure 12 (a-c)). The model estimates indicated that birds attained their maximum levels of performance with respect to *recovery time* and *unvoiced FM* at intermediate *Ta's* (Figure 13 (a, c)). In contrast, *voiced FM* performance was worst at intermediate temperatures and appeared to improve at higher temperatures, however data are sparse at high *Ta* (Figure 13 b). *Recovery time* and *unvoiced FM* improved with recent practice *(Order)*. Conversely, *voiced FM* decreased with *Order*. However, the model was not as confident about the effect of *Order* on *voiced FM* compared to its effect on the other two performance variables, as indicated by the wider confidence interval (Figure 14 (a-c)). The effects of *RH* and *JD* on the performance variables were weak, so I do not discuss them further (Figures 15 & 16 (a-c)).

The models on the acoustic traits that trade-off against each other indicated that the duration of silent gaps (*gap duration*) decreased during the dawn chorus, then increased over the next few hours (Figure 17 (a)). The estimated effect of *Time* on *note duration* was negative during the dawn chorus, flattened out or increased slightly after sunrise and decreased again beginning at 5,000 seconds after sunrise (Figure 17 (b)). The frequency bandwidth of silent gaps (*gap BW*) narrowed slightly during the dawn chorus, then broadened after sunrise (Figure 17 (c)). *Note BW* narrowed during pre-dawn singing, then broadened after sunrise (Figure 17 (d)). The conditional effects of *Ta* on the acoustic traits are not strong. *Note duration, gap BW*, and *note BW* decreased with high *Ta* but there is a relatively low sample size around this temperature (Figure 18 (a-d)). The conditional effects of song order on acoustic traits were mostly weak. *Gap duration*, however, decreased moderately with *Order*. *Order* did not appear to exert strong, independent effects on the other acoustic traits (Figure 19 (a-d)). The model estimates indicated that *RH* does not exert strong, independent effects on the acoustic traits (Figure 20 (a-d)). The estimated independent effects of *Julian* date on acoustic traits were subtle, but there were positive trends with both *Note BW* and *Gap BW* (Figure 21 (a-d)).

I used the conditioned models to predict performance given each of three constant song-rates. High song-rates predicted the best *recovery time* and *unvoiced FM* performance, whereas low song-rates predicted the poorest performance (Figure 22 (a, c)). These patterns broke down about 10,000 seconds (2.8 hours) after sunrise, but data were sparse that late in the morning. In contrast, the model predicted slightly better *voiced FM* performance when song-rates were low, compared to when they were medium or high (Figure 22 (b)).

My models predicted the best *recovery time* and *unvoiced FM* performance at the median *Ta*, and inferior performance at both low and high temperatures (Figure 23 (a, c)). The effect of varying temperature on *recovery time* was particularly strong. The results from the *voiced FM* analysis showed the opposite effect: performance was predicted to be better at low and (especially) high temperatures than at the median *Ta* (Figure 23 (b)).

The *conditional* R^2 and *marginal* R^2 showed that for all performance variables and acoustic traits, the percent of the variance explained by the whole model was higher than the percent explained by the predictor variables alone (Table 2.7). The random variables greatly improved the explanatory power of the model.

The variance (σ^2) of the random effects in the performance analysis showed that for most of the models, song-type within *Bird ID* (*Bird ID / ST*) accounted for the most variance, followed by Year, then *Bird ID* (Table 2.8). The confidence intervals from the variance of Year were very broad, so the model was not as confident about this variance.

2.3.2 Dawn chorus start time analysis

The descriptive statistics for the *DCST* dataset are shown in Table 2.9. Neither the parameter estimates, nor the model predictions indicate that *Ta* has a substantial independent effect on *DCST* (Figure 24 (a) & 25). There is, however, moderate evidence that *RH* has a positive influence on *DCST* (Figure 24 (b)). *Precipitation* has a negative influence on dawn chorus start time. Birds start to sing earlier if it rained during the previous day of the dawn chorus (Figure 24 (c)). Finally, average *DCST* was earliest in mid-April, and later at more extreme dates (Figure 24 (d)). However, the 95% confidence interval is broad in this analysis. The model is not confident about the effects of weather and day of year on dawn chorus start time.

The *conditional* R^2 was higher than the *marginal* R^2 in the *DCST* analysis. The whole model explained 43% of the variance and the predictor variables alone explained 8% of the variance. There was variation in the data that the predictor variables did not explain. *Year* accounted for most of the variance **(**σ^2) in the grouping levels but the confidence intervals of this estimates are very broad (Table 2.10).

2.4 Discussion

In this section I summarize and interpret my results with respect to vocal performance and the warm-up hypothesis in male Adelaide's warblers. For a broader interpretation, including implications for other species, please refer to Chapter 3.

2.4.1 Describing the variables

I found evidence of constrained performance in all three acoustic comparisons. The 2017 data appeared to be subject to the same constraints as the 2012 data analyzed in Logue et al. (2020; Figure 10). The deviation scores of *recovery time, voiced FM*, and *unvoiced FM,* averaged over notes within a song, were all moderately skewed to the right (Figure 11 (a-c)). Here, a right skew implies a high density of songs near the putative performance constraint. Logue et al. (2020) propose that this kind distribution implies that birds are attempting to approach the performance limit.

High *recovery time* performance occurs when songs have relatively short silent gaps and relatively long notes, resulting in high sound density. Sound density during trilled songs is constrained by respiratory performance, so long as birds take mini-breaths between notes (Hartley & Suthers, 1989; Cardoso et al., 2007; Suthers & Zollinger, 2004; Logue et al., 2020; Goller, 2021). There is evidence that Adelaide's warblers do take mini-breaths between notes most, but not all, of the time (Logue et al. 2020). Low deviation scores of *voiced FM* and *unvoiced FM* indicate rapid frequency modulation during notes and gaps, respectively. Frequency modulation speed can be constrained by the speed and coordination of the syrinx and upper vocal tract (see description of the syrinx below; Logue et al., 2020; Goller, 2021).

2.4.2 Song performance analysis

I found that vocal performance in male Adelaide’s warblers improved rapidly during the dawn chorus. These performance variables diminished or held steady after sunrise. Only one of the three performance variables (i.e., *voiced FM*) improves again after sunrise. Two of the performance metrics (i.e., *recovery time* and *unvoiced FM*) improved with the cumulative number of songs that a bird sang during the morning and with higher song-rates. *Recovery time* and *unvoiced FM* attained their maximum performance at intermediate *Ta’s*, compared to low and high *Ta*. The third performance variable, *voiced FM,* did not seem to be affected by *Ta*, *Order*, and song-rate in the same way that *recovery time* and *unvoiced FM* were.

My lab previously showed that male Adelaide’s warblers warm-up their voices during morning singing (Schraft et al., 2017). I attempt to advance this finding by testing two predictions of the hypothesis that the dawn chorus facilitates vocal warm-up, and one prediction of the hypothesis that physiological warming underlies the improvement in performance

2.4.3 Prediction 1: Performance increases more rapidly during the dawn chorus

I attempted to improve on the description of the change in vocal performance during the morning singing in male Adelaide’s warblers by using a larger dataset, Bayesian regression models with splines, and a fine-scale approach to measure vocal performance. Schraft et al. (2017) showed that performance decreases with time relative to sunrise and increases with song order. Their linear models did not allow them to characterize fine scale changes in performance over the course of the morning. I addressed that limitation by modeling performance over time with splines, which have higher degrees of freedom than linear or quadratic models (Mackenzie et al., 2005).

I used the spline-based models to test the prediction of the warm-up hypothesis that vocal performance increases more rapidly during the dawn chorus than it does the rest of the morning. Spline mixed-effects models from my analysis did characterize the change of performance variables over the course of the morning singing better than the linear models used in Schraft et al. (2017). The models with splines showed a strong pattern of a rapid improvement during the dawn chorus and a diminution after sunrise of the performance variables (i.e., *recovery time*, *voiced FM*, and *unvoiced FM*; Figure 12 (a-c)). This result supported the first prediction of the warm-up hypothesis. Results from this analysis also showed that performance is at its worst first thing in the morning and improves during the dawn chorus, as expected (Figure 12 (a-c)). It seems likely the birds' voices "cool off" overnight when they do not sing, and warm-up during the dawn chorus as they use the muscles involved in sound production (Déaux et al., 2020). The precise physiological mechanisms underlying behavioural warm-up, however, remain unknown.

The change of deviation scores over *Time* is primarily driven by the variation in one of the raw acoustic traits that comprise their trade-offs (Figure 12 (a-c) & 17 (a-d)). Specifically, *gap duration* tracks changes in *recovery time* better than *note duration; note duration* tracks with *voiced FM* better than *note BW*, and the trend in *gap duration* is more like that of *unvoiced FM* than *gap BW*. This uneven influence occurs because the slopes of the quantile regression curves used to calculate deviation scores are very different from 1 (Table 2.1), meaning that one of the constituent acoustic traits is weighted more heavily that the other (Logue et al. 2020).

To summarize, this analysis improved the characterization of vocal performance over the course of the morning singing and showed that the improvement in performance described by Schraft et al. (2017) occurs during the dawn chorus, when birds sing at higher rates than the rest of the morning. This finding gives support to the hypothesis that the dawn chorus is an adaptation to warm-up the voice.

2.4.4 Prediction 2: Song-rate positively influences vocal performance

Results from this study showed that *recovery time* and *unvoiced FM* improved with *Order* (Figure 14 (a-c)). This result lends evidence to the finding that birds improve their song performance with recent practice (Schraft et al., 2017; Dinh et al., 2020). However, I predicted that performance may not simply increase with the cumulative number of songs that a bird sang during the morning. Rather, based on the hypothesis that the dawn chorus is an adaptation for rapid vocal warm-up, I predicted that song-rate (songs / minute) would exert a positive influence on vocal performance. Predictions based on my conditioned models indicated that high song-rates led to much better *recovery time* and *unvoiced FM* than did median or low song-rates (Figure 22 (a-c)). This finding suggests that song order (the cumulative number of songs) and time relative to sunrise interact to influence performance and that, as predicted, vocal performance improves in response to high song-rates. Thus, I found support for two predictions of the hypothesis that the dawn chorus is an adaptation to warm-up the voice (Schraft et al., 2017). My third prediction relates to the hypothesis that *Ta* influences *Tb*, which in turn influences vocal performance.

2.4.5 Prediction 3: Performance is lower at low Ta

My third prediction stated that low *Ta* is associated with low performance. Spline mixed-effects models supported this third prediction in two of the performance variables (i.e., *recovery time* and *unvoiced FM*, Figure 13 (a-c)). This result may be explained by the hypothesis that the improvement in performance is a consequence of physical warming (Bishop, 2003a, 2003b). *Ta* and *Tb* are positively correlated in endotherms (Merola-Zwartjes & Ligon, 2000; Nord et al., 2009; McKechnie et al., 2017; McWhorter et al., 2018). Changes in *Tb* affect the performance of the peripheral nervous system and muscles involved in breathing and sound production, which can ultimately affect acoustic traits (Wood et al., 2013; Wu et al., 2021). The hypothesis that a physical warm-up leads to the improvement in performance, and the relationship between *Ta* and *Tb* may explain why birds perform better at intermediate Ta's compared to low temperatures. There is a range of temperatures where any kind of performance peaks. It starts to decrease after that limit in temperature (Schulte et al., 2011 Cunningham et al., 2013; Brandt et al., 2018). *Recovery time* and *unvoiced FM* in male Adelaide's warblers are not an exception to this pattern. Both performance variables also diminished at high *Ta* (Figure 13 (a-c)). The link between *Ta* and performance may be mediated by overall *Tb*, or by the temperature of the muscles involved in sound production.

To summarize, the maximum vocal performance of male Adelaide's warblers occurred at intermediate *Ta*'s, suggesting a relationship between *Ta* and *Tb* and a behavioural warm-up due to a physical warm-up.

2.4.6 Voiced FM versus recovery time and unvoiced FM

The deviation score of *voiced FM* behaved differently than the other two deviation scores. It diminished with recent practice and high song-rates. It was also worst at intermediate *Ta's* but improved at higher *Ta*. The key to understanding these differences may be *note duration*, which is one of the acoustic variables that trade-off at the limit of *voiced FM*. Critically, *note duration* influences *voiced FM* and *recovery time* in different ways: longer *note durations* lead to worse *voiced FM* performance, but better *recovery time* performance (Figure 10 (a, b)). Thus, improvement in *recovery time* may come at the cost of worse performance in *voiced FM*.

A respiratory constraint leads to the trade-offs between *note duration* and *gap duration* (Hartley & Suthers, 1989; Wild et al., 1998; Mota & Cardoso, 2001; Suthers & Zollinger, 2004; Cardoso et al., 2007; Logue et al., 2020). I hypothesize that singers are attempting to maximize trill speed while minimizing silent gaps. Tests of receiver responses to songs with varying *voiced FM* performance and *recovery time* could shed light on this hypothesis.

Imprecise measurement may also have contributed to the relative lack of pattern in *voiced FM*. The ratio between the maximum and minimum frequency in a note is the standard method to measure *note BW*, one of the variables used to calculate *voiced FM* (Cardoso, 2013). However, this approach is blind to patterns of within-note *FM* that do not affect *note BW*. Two notes with the same minimum and maximum frequencies but different rates of *FM* would yield the same *note BW*, so our measure of *voiced FM* does not fully capture the bird's ability to rapidly modulate note frequency (Goller, 2021). An analytic method that traces the contour of each note, like FEX, but also allows users to separate *voiced FM* from *unvoiced FM* would be ideal.

A third possible explanation for the different patterns in *voiced FM* is that the trade-off between *note BW* and *note duration* ignores an important constraint on *voiced FM*. The syrinx is the organ that birds use to generate the sound waves that comprise song. It includes two sound sources – the left and right sides – that specialize in different sound frequencies with different bandwidths (Suthers & Zollinger, 2008). The left side produces lower frequency sounds with wider bandwidths than does the right side (Suthers & Zollinger, 2008). My measure of voiced *FM* does not consider how *FM* may be constrained differently depending on whether the bird is using the right side of the syrinx, the left side, or both sides (Goller, 2021). Interestingly, neither of the issues discussed in the last two paragraphs should affect *unvoiced FM*, which neither contains non-linear *FM*, nor involves syringeal sound production.

2.4.7 Random effects

The random effects had a major impact on the explanatory power of the models in the song performance analysis (Table 2.7). *Recovery time*, *voiced FM*, and *unvoiced FM* were highly variable among song-types within *Bird ID* (Table 2.8). This result supports Logue et al.'s (2020) finding that the different song-types in males' repertoires have different levels of performance. Song-types that are particularly difficult to produce may function in mate attraction or aggressive contexts. For example, swamp sparrows have song-types in their repertoires with different levels of performance and females make more copulation solicitation displays in response to playback of higher performance songs (Ballentine et al., 2004). Dark-eyed juncos differ the performance of their song-types in aggressive contexts (Cardoso et al., 2009). Male red-winged blackbirds also have song-types with different levels of performance in their repertoires and respond more – higher song-rate – to low performance songs (Cramer & Price, 2007). Further investigation is required to know if different song-types with different levels of performance function in mate attraction or in aggressive encounters with other males in Adelaide's warblers.

It is also possible that song-types that were low performance according to the three measures used in this study, require high levels of performance in a metric that I did not consider. Above, I discussed how the two-source nature of the syrinx could constrain sound production in ways that were not considered in this study. Another possible constraint that I did not measure directly involves the mechanism the bird uses to produce trilled notes (i.e., mini-breath or pulsatile trill; Goller, 2021). Mini-breaths, which birds use to replenish the air supply between notes, permit lower trill rates than pulsatile trills, which are defined as a "sustained expiratory pulse" (Logue et al. 2020; Goller, 2021, p. 3). The duration of pulsatile trills, but not mini-breath trills, is subject to a respiratory constraint because birds do not inhale during pulsatile trills. I did not separate mini-breath trills from trills produced by pulsatile expiration in Adelaide's warblers' songs in this analysis. These systems may induce constraints in sound production that affect certain song-types in the repertoires of male Adelaide's warblers, currently defined as lower performance songs. There was also variance among years in the performance analysis. Each individual in this analysis contributed only one year of data, so age is unlikely to explain this variance. Variance among years may be explained by a difference in habitat quality. For example, one year could have had higher food availability, which influenced birds' condition and vocal performance (Strain & Mumme, 1988; Rashotte et al., 2001; Clarkson, 2007). The strong influence of random effects in this analysis emphasizes the importance of including song-types and year as random effects in analyses of vocal performance.

2.4.8 Dawn chorus start time analysis

Despite a robust sample size, my model did not show that weather and day of year affect dawn chorus start time in male Adelaide's warblers. I conclude that the measured parameters probably have little to no effect on dawn chorus start time in this population. However, if I were to assume that the weak results, I observed proved to be repeatable, I would interpret them as follows.

Results from this model provided no evidence that *Ta* affected the time that male Adelaide's warblers start to sing (Figure 24 (a) & 25). There is little variance in temperature at my study site, and the lowest observed temperatures are still quite warm. In the absence of extreme temperatures, *DCST* may be canalized (Puswal et al., 2020).

My *DCST* model showed a weak positive relationship between dawn chorus start time and relative humidity (*RH*). Birds, however, started to sing earlier with higher precipitation accumulation. The influence of precipitation accumulation before the dawn chorus might be explained by the sound propagation hypothesis (Henwood & Fabrick, 1979). Maybe they could not sing as much during the previous day because the noise of precipitation impaired the propagation of their songs. Similarly, they may start earlier after rainy days to re-establish territorial claims or mate attraction after a day in which rain interrupted normal singing behavior.

They might also avoid getting wet because humidity impairs their thermoregulatory methods such as evaporative heat loss (Gerson et al, 2014; Levesque & Marshall, 2021). Consistent with these idea, male Adelaide's warblers do not sing when it was raining (D.M. Logue, pers. obs.).

Finally, while the overall trend of the effects of *JD* on *DCST* was fairly flat, all the earliest *DCST's* were observed in mid-April (Figure 24 (d)). A possible explanation for the change in the time that male Adelaide's warblers start to sing as the days progress might be the different stages of their mating season, changes in testosterone levels, changes in territorial intrusions, etc. (Mace, 1987; Welling et al., 1995; Hunt et al., 1997; Foerster et al., 2002; Catchpole & Slater, 2003; Da Silva et al., 2014; Puswal et al., 2021). It may be possible, for example, that the period in the breeding season when males start to sing earlier is due to higher levels of testosterone or because the females are more fertile (Mace, 1987; Hunt et al., 1997; Foerster et al., 2002; Catchpole & Slater, 2003).

2.4.9 Summary

This study improved the characterization of performance over the course of the morning using models with splines, a larger dataset, and a fine-scale approach to measure performance. This characterization showed a pattern of vocal performance over time that strongly supports the first prediction of the hypothesis that the dawn chorus is an adaptation to warm-up the voice: vocal performance increases more rapidly during the dawn chorus than it does during the rest of the morning. The second prediction of the warm-up hypothesis, that high song-rates drive improvement in vocal performance, was also supported with respect to *recovery time* and *unvoiced FM*. This result suggests the importance of considering the interaction between time and measures of recent practice song (e.g., song order) in the analysis of the warm-up hypothesis. I also found that *recovery time* and *unvoiced FM* peak at intermediate *Ta's*. This result supports my third prediction – low *Ta* is associated with low performance. In addition, this analysis indicated that high *Ta* is also associated with low performance.

Voiced FM did not show the same patterns of *recovery time* and *unvoiced FM*, possibly because it inherently trades off with other, more important performance variables, or because the way that I calculated *voiced FM* fails to capture the key.

I found no evidence that ambient temperature exerts a strong effect on *DCST* in Adelaide's warblers. *DCST* may, however, increase slightly with relative humidity and decrease slightly with precipitation accumulation. Although there was little evidence that date affected the mean *DCST* over the study period, the earliest *DCST's* all occurred in mid-April.

2.5 Tables and figures

2.5.1 Tables

Table 2.1. *Results of the mixed-quantile regression analyses (tau=0.1) used to estimate note-level performance limits for recovery time, voiced frequency modulation, and unvoiced frequency modulation. These intercepts and slopes were obtained from Logue et al. (2020).*

Performance variable	x	y	n (notes)	Intercept (raw)	Slope	p-value (slope)
Recovery time	Note duration	Gap duration	68,602	13.85	0.24	< 0.0001
Voiced FM	Note BW	Note duration	68,602	8.74	9.34	< 0.0001
Unvoiced FM	Gap BW	Gap duration	68,602	17.06	3.43	< 0.0001

Table 2.2. *WAIC and standard error (SE) to compare the different levels of complexity of the model for song performance analysis (i.e., linear, quadratic, spline). The model with the lowest WAIC is in bold font.*

Mixed-effects model	WAIC/ SE
Linear	7877.1 / 119.5
Quadratic	7852.6 / 119.7
Spline	**7646.4 / 124.0**

Table 2.3. *WAIC and standard error (SE) for each model with two-way interactions and three-way interactions between predictor variables in the performance analysis. (*) represents the interaction between predictor variables and (**) means that R-hat values are equal to or higher than 1.05 (see description of R-hat below). The model with the lowest WAIC is in bold font.*

Model	Fixed effects	WAIC/ SE
1	Ta, RH, Time, Order, JD	7647.0 / 124.1
2	Ta * RH, Time, Order, JD	7630.5 / 124.8
3	Ta, RH, Time * Order, JD	7636.3 / 122.7
4	Ta * Time, RH, Order, JD	7649.5 / 121.8**
5	Ta * Order, RH, Time, JD	7658.5 / 123.5
6	Ta * RH, Order * Time, JD	7617.9 / 123.3
7	Ta * Time, Ta *RH, Order, JD	7638.5 / 123.2
8	Ta * Order, Ta * RH, Time, JD	7629.1 / 124.1

9	Ta * Time, Order * Time, RH, JD	7628.2 / 123.1**
10	Ta * Order, Order * Time, RH, JD	7643.5 / 123.0
11	Ta * Time, Ta * Order, RH, JD	7666.0 / 122.6
12	Ta * Time, Ta * RH, Order * Time, JD	7608.6 / 123.2
13	Ta * Order, Ta * RH, Order * Time, JD	7614.3 / 123.5
14	Ta * Order, Ta * Time, Order * Time, JD, *RH*	7618.7 / 122.9
15	Ta * RH, Ta * Order, Ta * Time, JD	7634.4 / 123.2
16	**Ta * Time, Ta * RH, Order * Time, Ta * Order, JD**	**7599.6 / 123.5**
17	Ta * Time * Order, Ta * RH, JD	7606.5 / 123.4
18	Ta * Order * Time, RH, JD	7626.6 / 122.8
19	Ta * Time * RH, Order, JD	7659.6 / 123.3
20	Ta * RH * Order * Time, JD	7647.2 / 123.3**

Table 2.4. *WAIC and standard error (SE) for each model with two-way interactions and three-way interactions between the predictor variables in the dawn chorus start time analysis. (*) represents the interaction between predictor variables. The model with the lowest WAIC is in bold font.*

Model	Fixed effects	WAIC/ SE
1	Ta, RH, Precipitation, JD	1655.1 / 17.4
2	**Ta * RH, Precipitation, JD**	**1651.9 / 17.7**
3	Ta, RH, Precipitation * JD	1659.9 / 16.5
4	Ta * RH, Precipitation * JD	1660.6 / 17.0
5	Ta * RH * Precipitation, JD	1654.1 / 17.7
6	Ta * RH * Precipitation * JD	1665.2 / 16.2

Table 2.5. *Description of the response and predictor variables used in the analysis of song performance and dawn chorus start time.*

Variable	Description	Analysis
Recovery time	deviation score of the trade-off between note duration and gap duration	Performance
Voiced FM	deviation score of the trade-off between note BW and note duration	Performance

Unvoiced FM	deviation score of the trade-off between gap BW and gap duration	Performance
Gap duration	duration of silent gaps in milliseconds	Performance
Note duration	duration of notes in milliseconds	Performance
Gap BW	frequency bandwidth in the silent gap	Performance
Note BW	frequency bandwidth in the note	Performance
DCST	time of the first song in a recording session	DCST
Air temperature (*Ta*)	temperature in degrees Celsius	Performance and DCST
Relative Humidity (*RH*)	percent of humidity relative to air temperature	Performance and *DCST*
Precipitation accumulation	precipitation accumulation in millimeters	*DCST*
Julian date (*JD*)	number of the day since January 1st	Performance and *DCST*

Song order (*Order*)	cumulative number of songs that a bird had sung during the morning	Performance
Time relative to sunrise (*Time*)	time of the song relative to the time of sunrise	Performance

Table 2.6. *Mean and standard deviation of song performance variables, acoustic traits, and predictor variables used in the analysis of song performance. Measures of performance and acoustic traits were taken at the note level and then averaged for each song.*

Variable	**Mean ± sd**
Recovery time	9.08 ± 4.10
Voiced FM	2.52 ± 0.88
Unvoiced FM	3.32 ± 1.32
Gap duration (ms)	35.57 ± 4.88
Note duration (ms)	51.60 ± 8.91
Note BW	2.06 ± 0.24
Gap BW	1.94 ± 0.25
Ta (°C)	24.14 ± 2.40
RH (%)	80.24 ± 7.17
JD (day of year)	114.80 (Apr. 24) ± 26.18
Time (s)	1,336.13 ± 3,576.24
Order	151.52 ± 102.28

Table 2.7. *Conditional R^2 and marginal R^2 for the model of each performance variable and acoustic trait.*

Model	Conditional R^2	Marginal R^2
Recovery time	0.57	0.13
Voiced FM	0.86	0.03
Unvoiced FM	0.71	0.12
Gap duration	0.73	0.12
Note duration	0.86	0.04
Gap BW	0.86	0.02
Note BW	0.87	0.02

Table 2.8. *Variance and lower and upper 95% confidence intervals of random effects from the spline mixed-effects model for each performance variable and acoustic trait.*

Random effect	Variance (σ^2)	l-95% CI	u-95% CI
	Recovery time		
Bird ID	0.18	0.09	0.36
Year	0.70	0.00	10.24
Bird ID / ST	0.36	0.29	0.45
	Voiced FM		
Bird ID	0.07	0.01	0.21
Year	0.56	0.00	8.88
Bird ID / ST	0.94	0.79	1.10
	Unvoiced FM		

Bird ID	0.15	0.06	0.32
Year	0.62	0.00	8.82
Bird ID / ST	0.52	0.44	0.62
Gap duration			
Bird ID	0.15	0.06	0.32
Year	0.61	0.00	9.61
Bird ID / ST	0.58	0.49	0.69
Note duration			
Bird ID	0.06	0.01	0.18
Year	0.46	0.00	8.82
Bird ID / ST	0.88	0.76	1.02
Gap BW			
Bird ID	0.02	0.00	0.10
Year	0.28	0.00	6.50
Bird ID / ST	0.85	0.72	1.00
Note BW			
Bird ID	0.02	0.00	0.11
Year	0.27	0.00	6.00
Bird ID / ST	0.94	0.81	1.10

Table 2.9. *Mean and standard deviation of male Adelaide's warbler dawn chorus start times, weather variables, and Julian date during the sampling periods in 2012, 2017, and 2018.*

Variable	Mean ± sd
DCST	-2,024.56 ± 508.34
Ta (°C)	23.20 ± 1.70
RH (%)	82.00 ± 6.04
Precipitation (mm)	130.76 ± 58.74
Julian (day of year)	107.92 (Apr. 18) ± 21.63

Table 2.10. *Variance and lower and upper 95% confidence intervals of random effects from the spline mixed-effects model in DCST analysis.*

Random effect	Variance (σ^2)	l-95% CI	u-95% CI
Bird ID (intercept)	4,033.52	7.78	29,241.00
Year (intercept)	177,055.81	23,596.03	966,819.89
1 + Ta.scaled (slope)	3,875.06	10.05	26,902.56

2.5.2 Figures

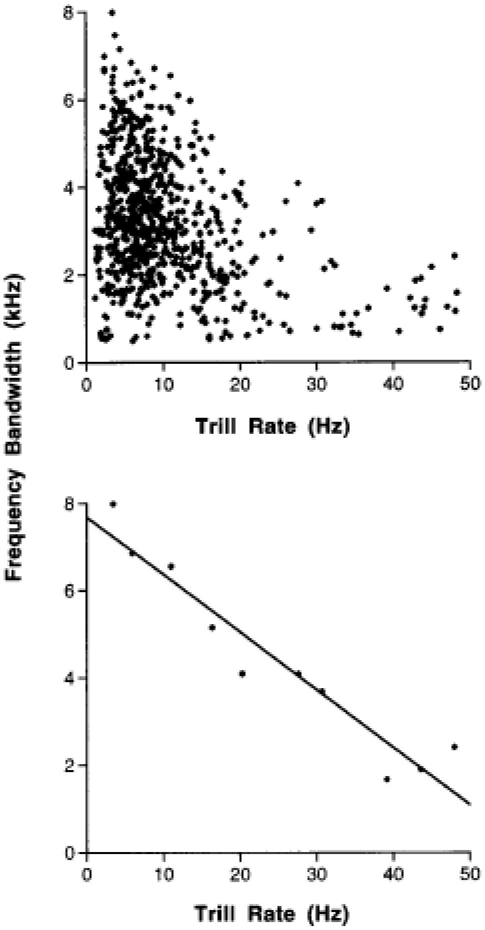


Figure 1. *Plots showing the trade-off between frequency bandwidth and trill rate. The upper panel shows all of the data. The lower panel shows the upper-bound regression (black line) that represents a performance limit. Copied from Podos. (1997) with the author's permission.*

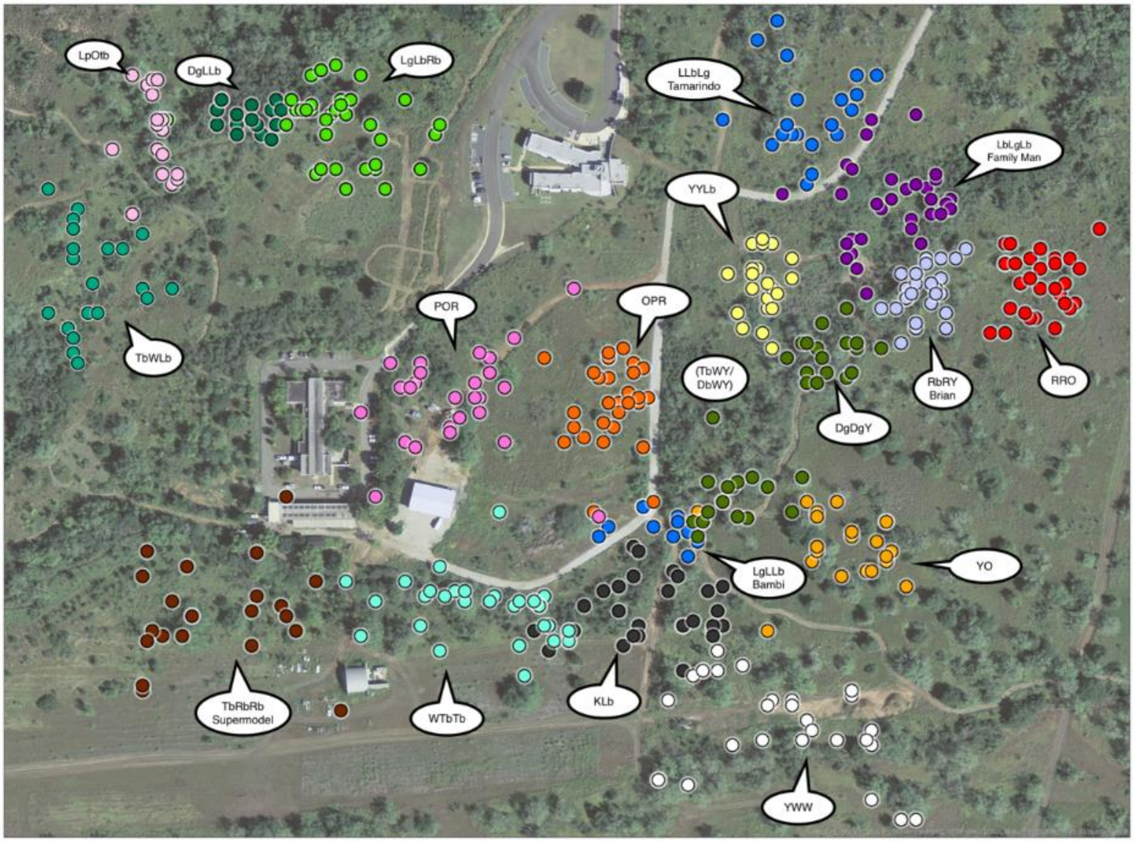


Figure 2. *Territories of focal male Adelaide's warblers recorded in 2017 at the Cabo Rojo National Wildlife Refuge. Dots with different colors represent the location of individuals. Speech bubbles indicate a bird identification for each color. Peter C. Mower created this map using ArcGIS® software by Esri.*

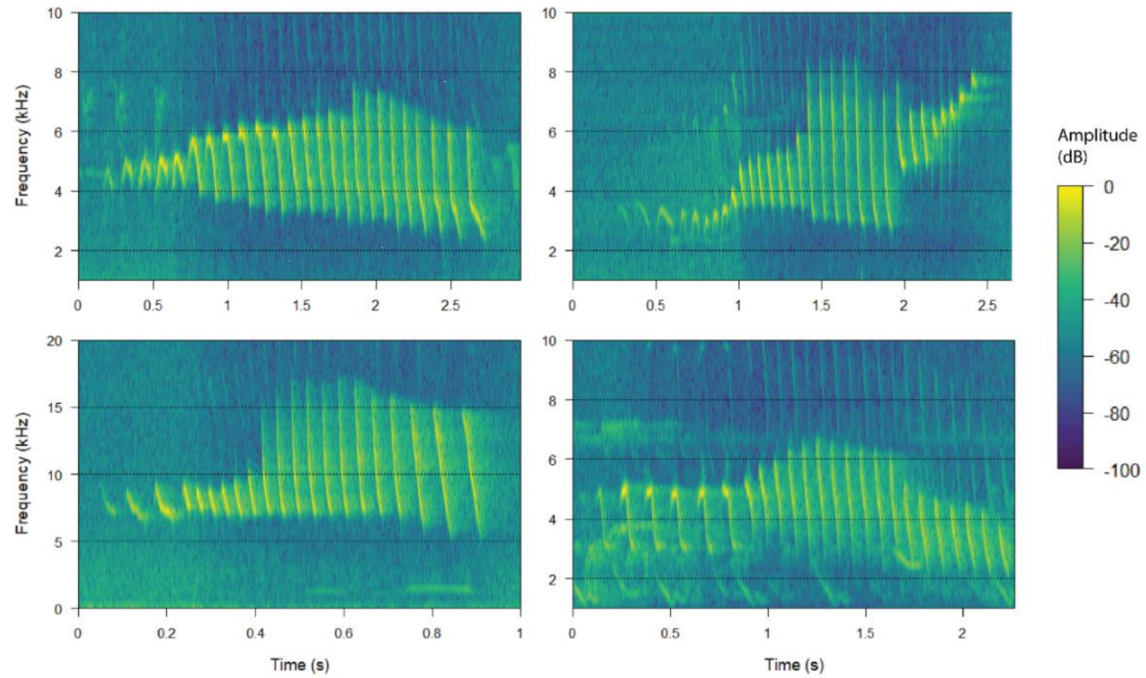


Figure 3. *Sound spectrograms of four song-types from male Adelaide's warblers. The color range represents sound amplitude in decibels (dB).*

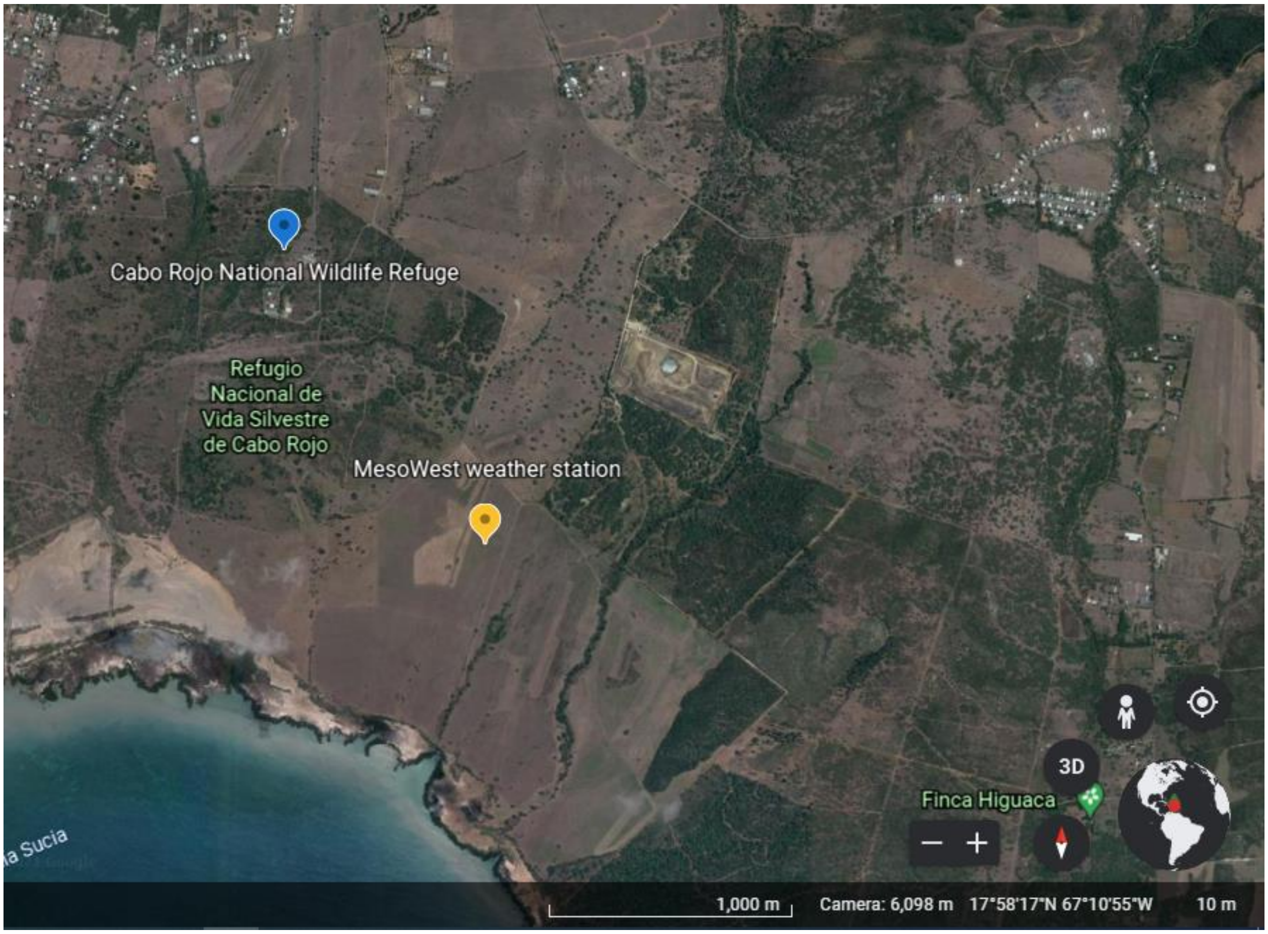


Figure 4. *Location of the Cabo Rojo National Wildlife Refuge, where the population of Adelaide's warblers were recorded (blue location symbol) and location of the MesoWest weather station, where I got weather parameters for my analysis (yellow location symbol; Google Earth, 2021).*

Figure 5. *Habitat of the territories of Adelaide's warblers at the Cabo Rojo Wildlife Refuge. Photo credits to Garry Tucker, Peter C. Mower, and David M. Logue.*

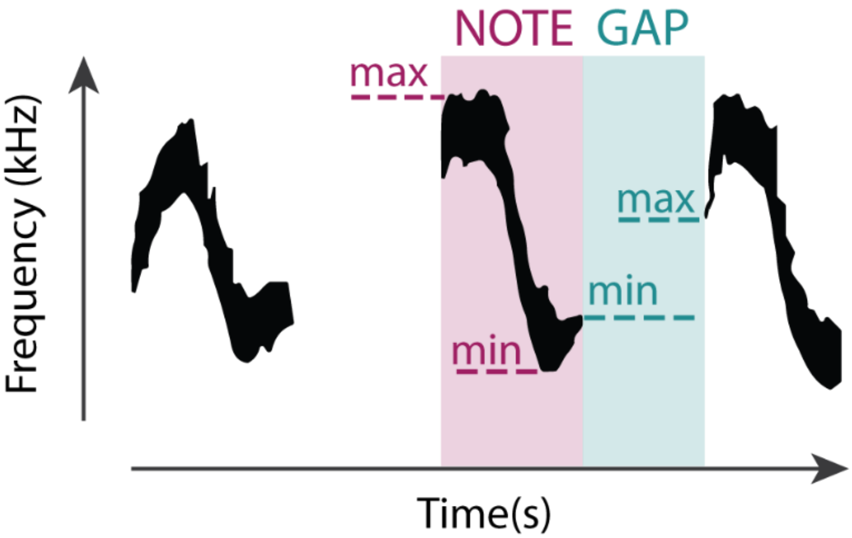


Figure 6. *Note-level measurements used to calculate acoustic traits for male Adelaide's warbler songs. The length of the purple shaded region represents note duration, and the length of the blue shaded region represents the duration of the silent gap. The purple max represents the maximum frequency of the note, and the purple min represents the minimum frequency of the note. These two measures are used to calculate note BW. The blue min represents the frequency of the end of a note and the blue max is the frequency of the beginning of the following note. Blue min and max are used to calculate gap BW (Logue et al., 2020).*

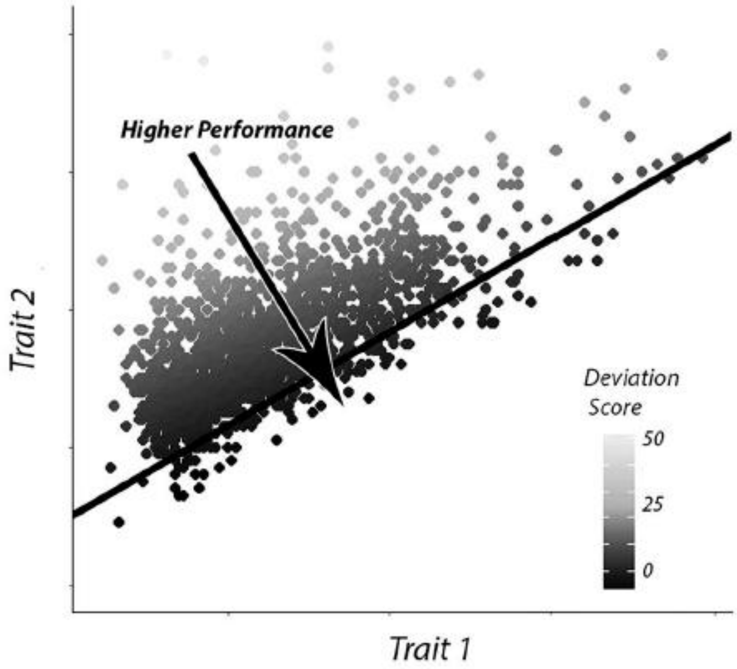


Figure 7. *Scatterplot of a hypothetical trade-off between two traits. The black line is a 10th quantile regression line used to estimate a performance limit. The closer each dot to the performance limit, the lower its deviation score and the higher its performance. Copied from Logue et al. (2020) with the author's permission.*

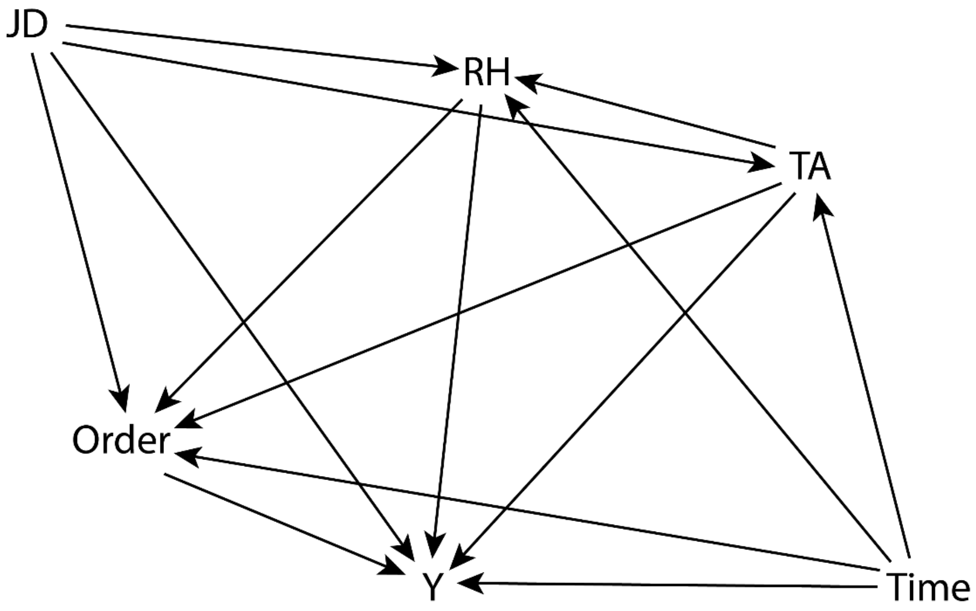


Figure 8. *A directed acyclic graph (DAG) describing hypothesized patterns of causality for variation in song performance. The words represent nodes, and the arrows represent directed edges with the direction of causal influence. JD is the day of year, Order is the cumulative number of songs, Time is time of the song relative to the time of sunrise, Ta is air temperature, RH is relative humidity. Y represents the performance variables (i.e., recovery time, voiced frequency modulation, and unvoiced frequency modulation) and the raw acoustic traits used to calculate the performance variables (i.e., note duration, gap duration, note bandwidth, and gap bandwidth).*

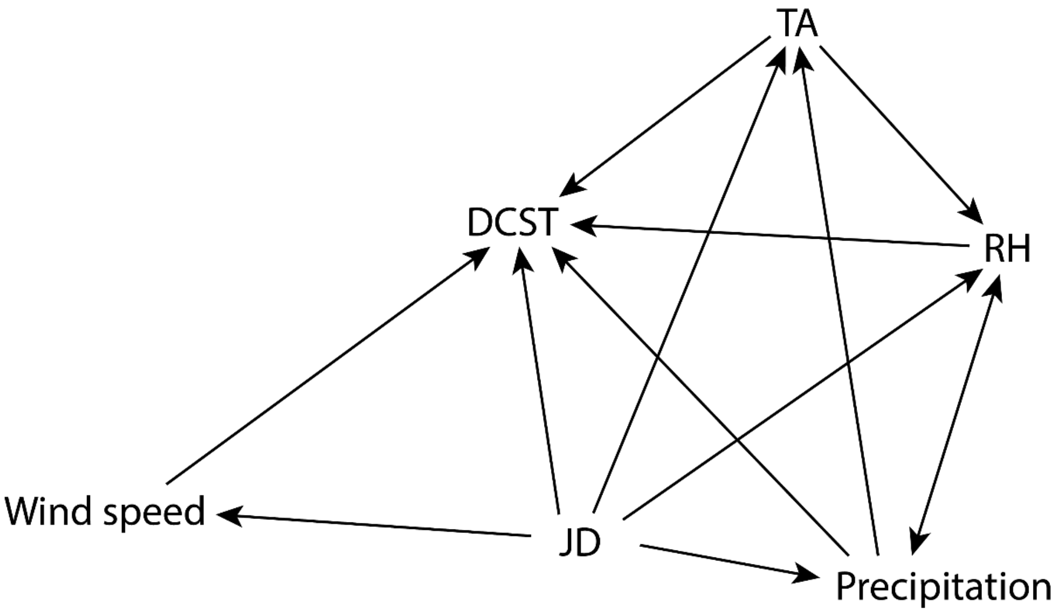


Figure 9. *A directed acyclic graph (DAG) describing hypothesized patterns of causality for variation in dawn chorus start time. The words represent nodes, and the arrows represent directed edges with the direction of causal influence. JD is the day of year, Ta is air temperature, RH is relative humidity, wind speed is the speed of wind, and Precipitation is precipitation accumulation. DCST is the time that the dawn chorus starts relative to the time of sunrise.*

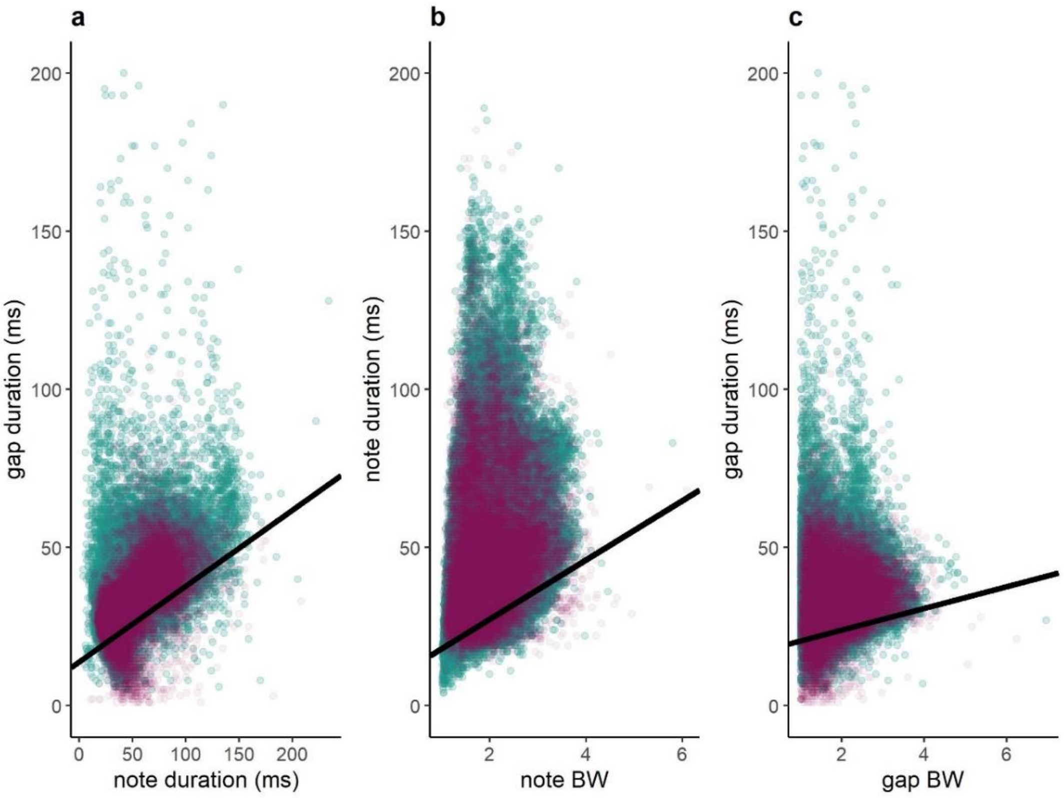


Figure 10. *Mixed-quantile regression (tau = 0.10) for the trade-offs between (a) note duration and gap duration, (b) note bandwidth and note duration, and (c) gap bandwidth and gap duration, that underlie the performance variables (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. Semi-transparent turquoise dots show the density of notes within songs from 2012. Purple dots indicate the density of notes within songs from 2017. 10th quantile regression lines (black) represent the estimates of performance limits with the intercepts and slopes that Logue et al. (2020) reported.*

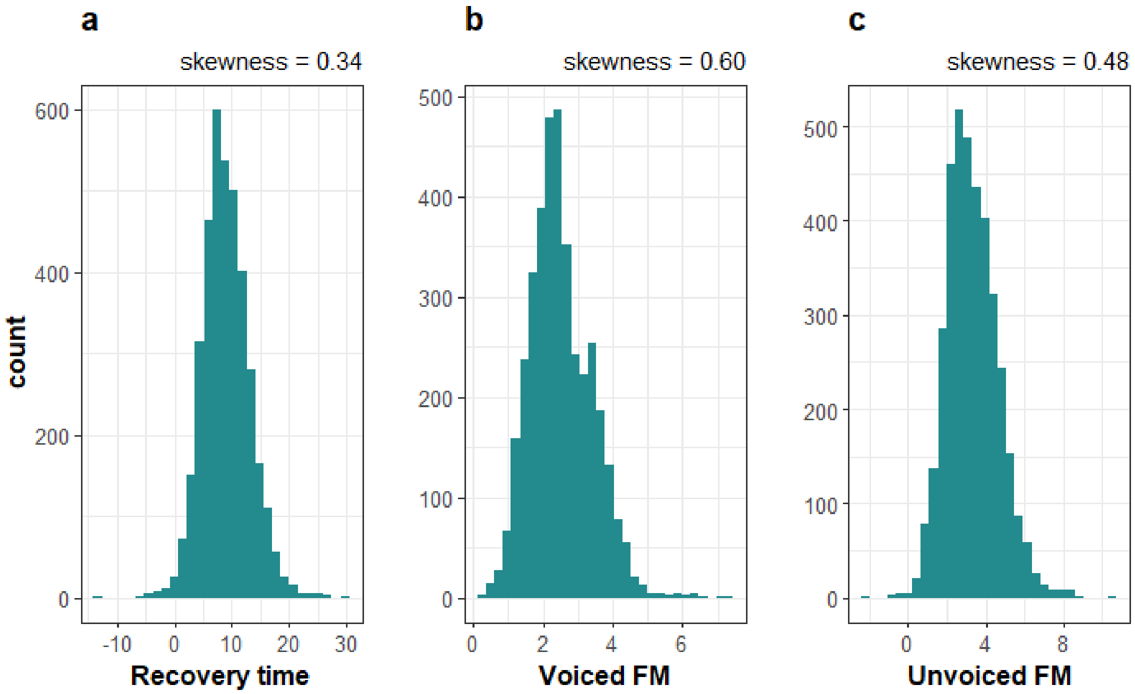


Figure 11. *Distribution of the deviation scores for (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation from 2012 and 2017. Deviation scores were measured at the note level, and then averaged for each song.*

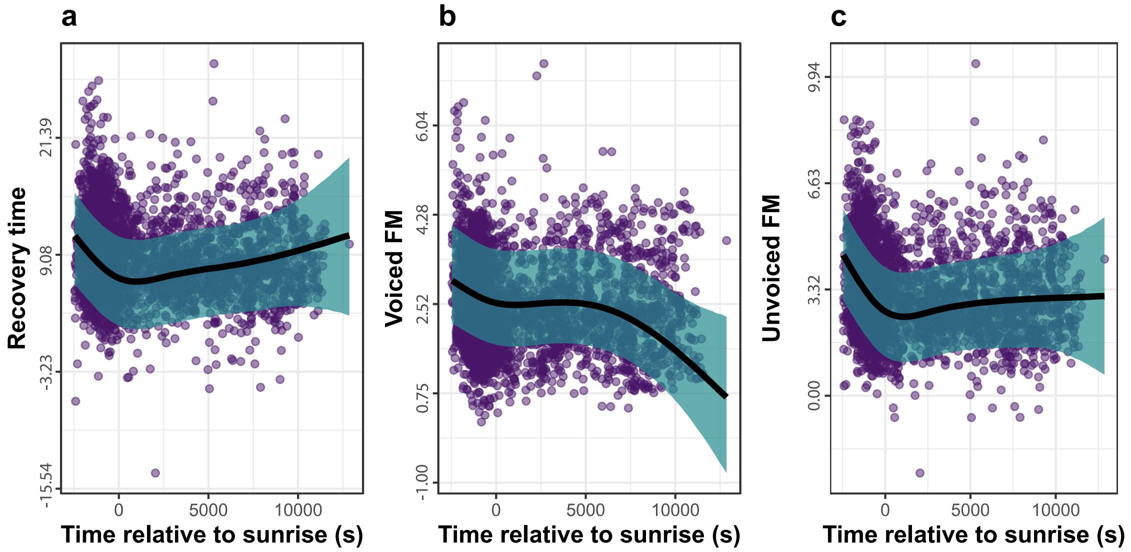


Figure 12. *Estimates of the effects of time relative to sunrise on the acoustic performance metrics (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation in*

male Adelaide's warbler song. Lower values of performance metrics indicate better performance. The black lines represent the estimate of the effect on the performance variable, the turquoise shaded areas represent the 95% confidence intervals of the estimate, and the semi-transparent purple dots represent songs. Estimates and confidence intervals were derived from Bayesian mixed models (see text).

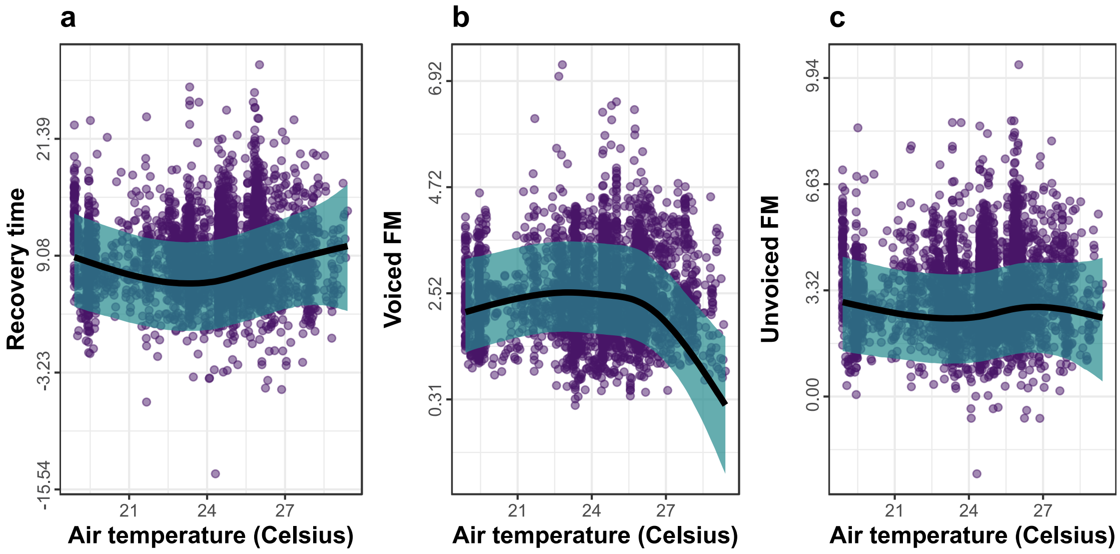


Figure 13. *Effects of air temperature on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

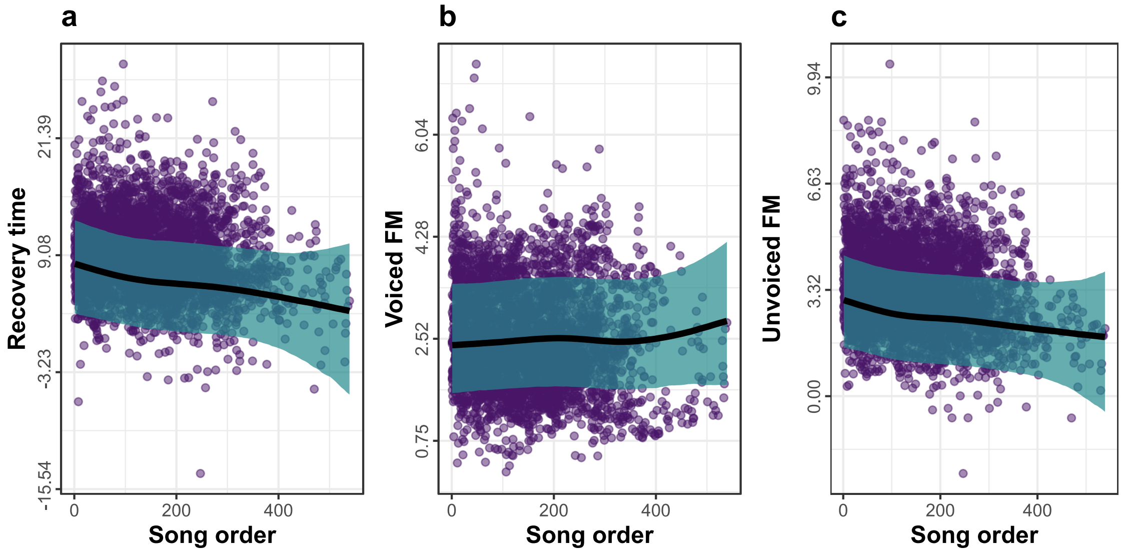


Figure 14. *Song order effects on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

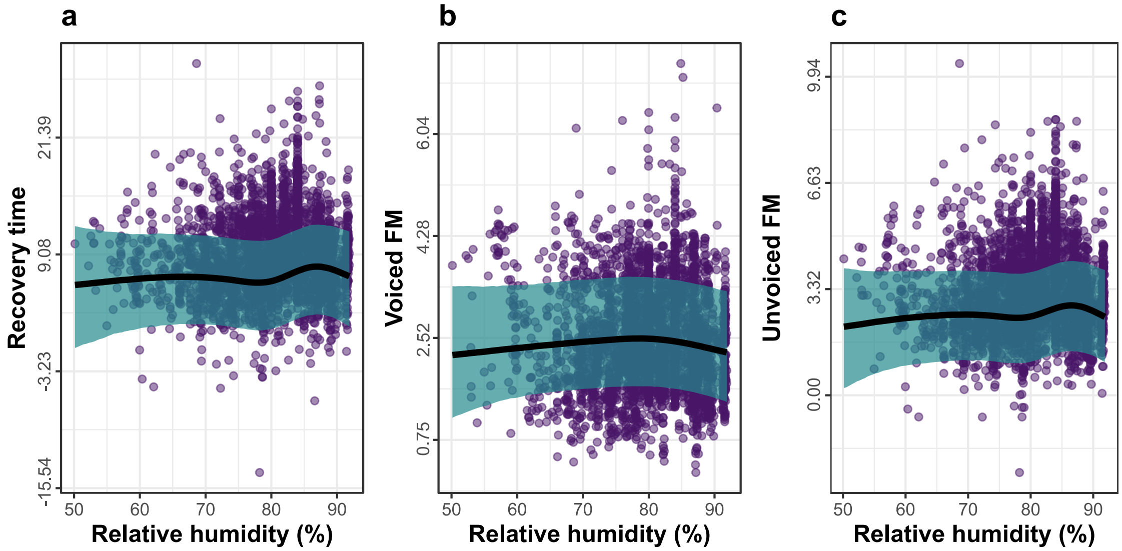


Figure 15. *Effects of relative humidity on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

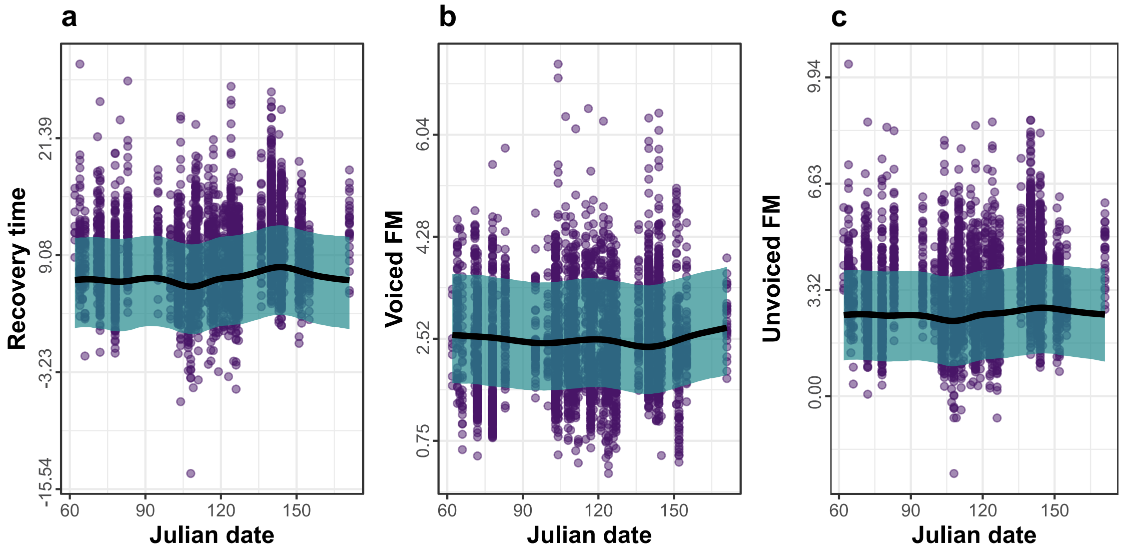


Figure 16. *Effects of Julian date on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. See Figure 12 caption for details.*

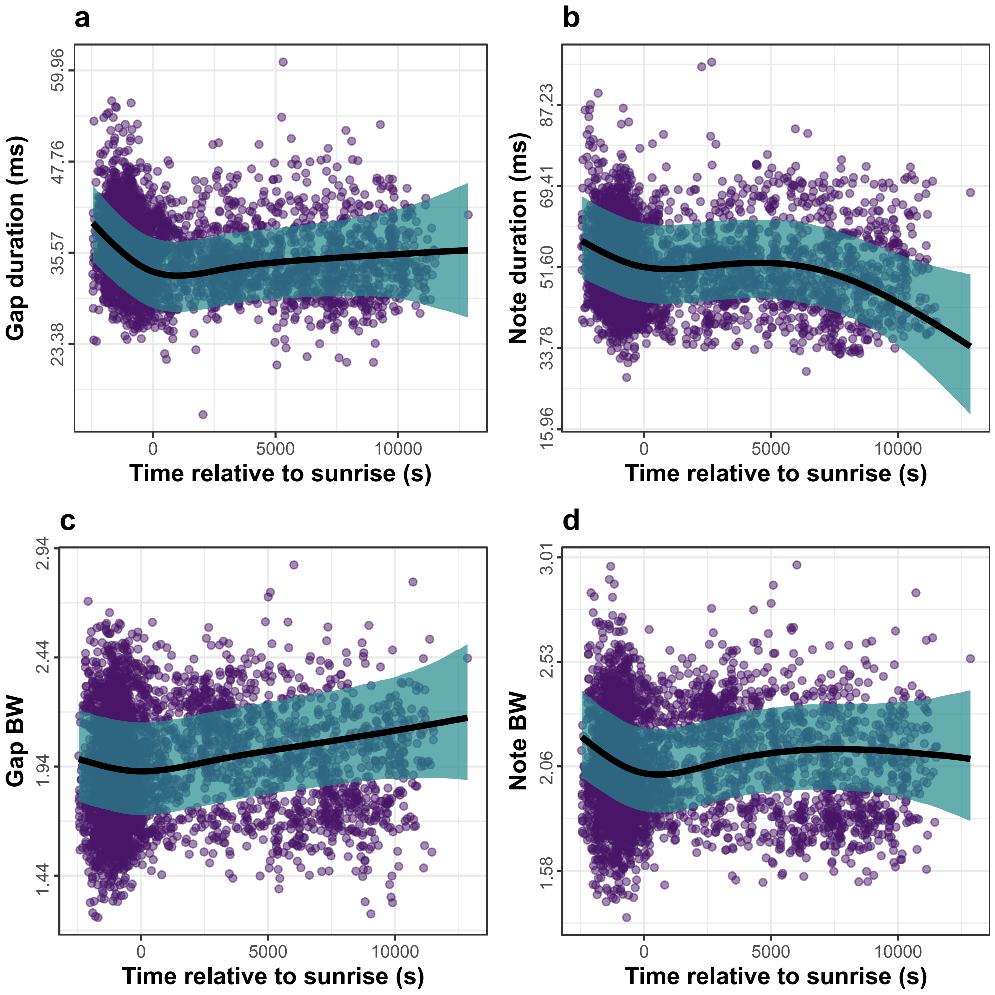


Figure 17. *Effects of time relative to sunrise on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. The black line represents the estimate of the effect on each acoustic trait, the turquoise shade represents the 95% confidence interval of the estimate, and the semi-transparent purple dots are songs. Estimates and confidence intervals were derived from Bayesian mixed models (see text).*

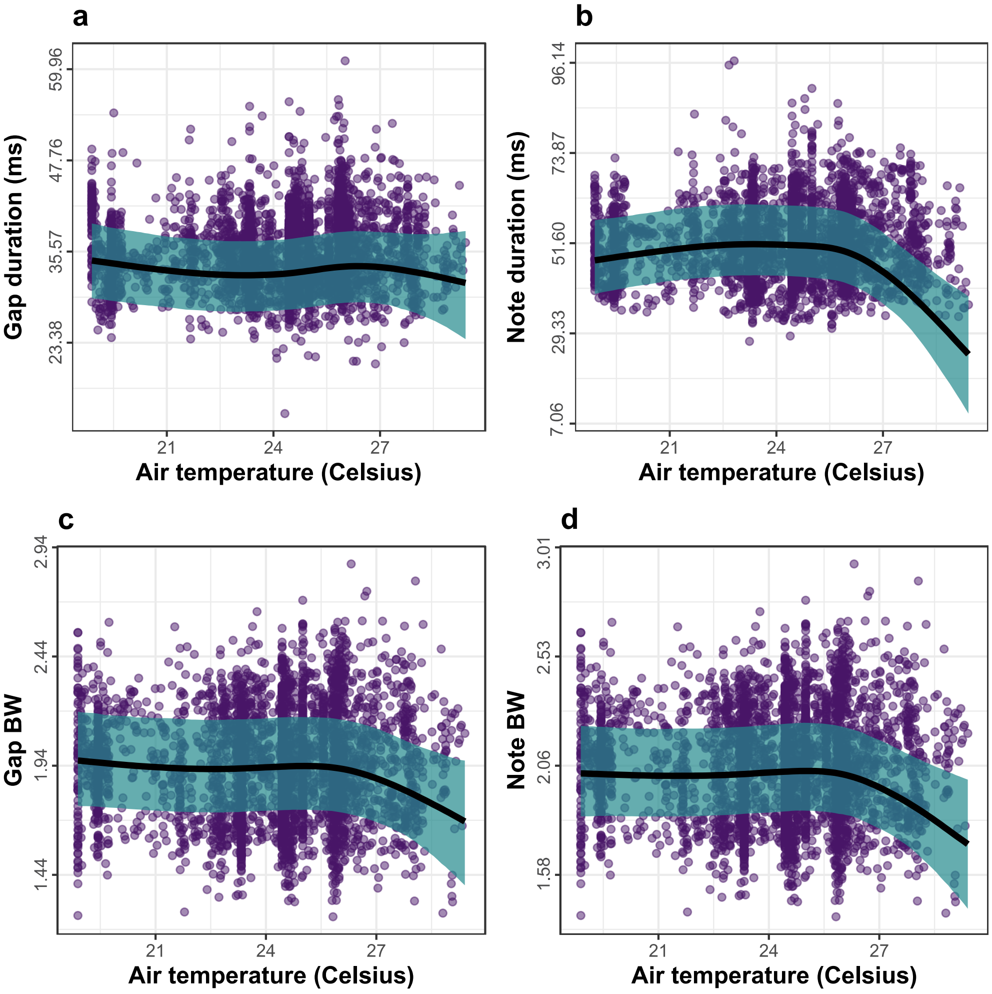


Figure 18. *Air temperature effects on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

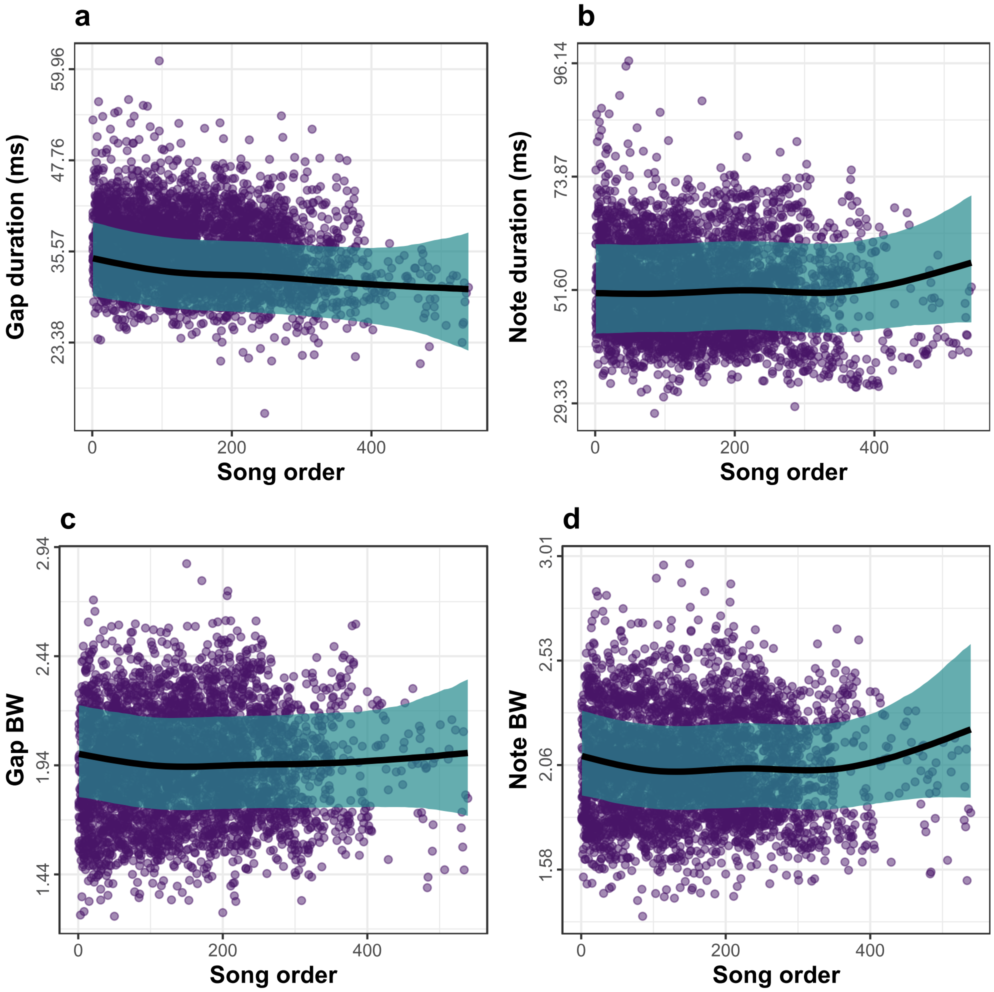


Figure 19. *Song order effects on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

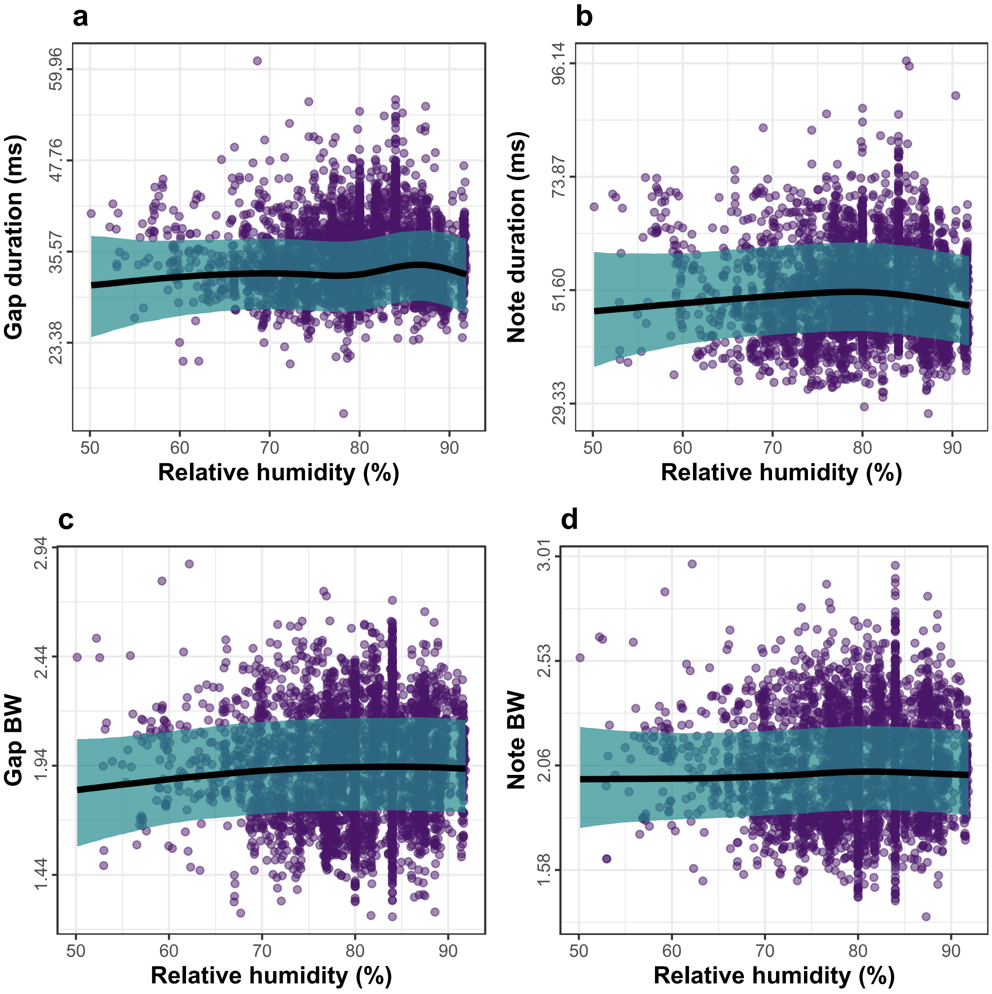


Figure 20. *Effects of relative humidity on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

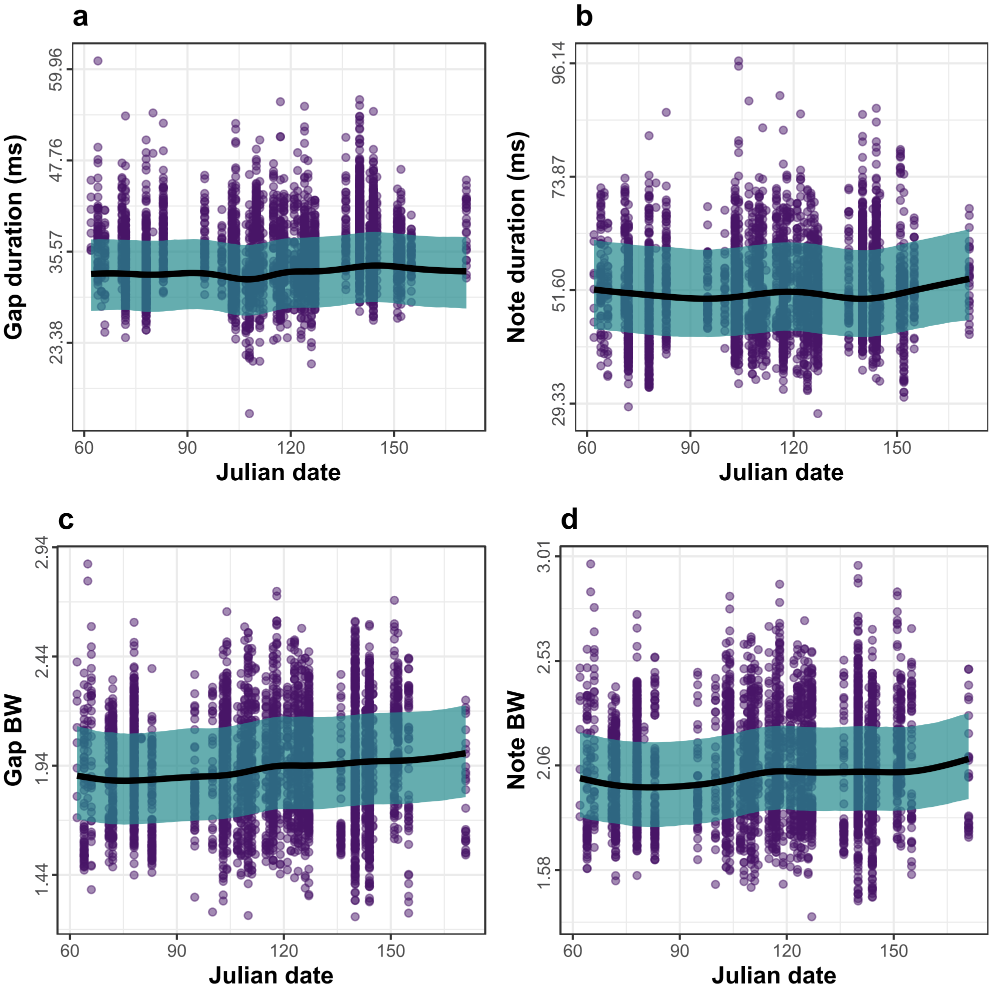


Figure 21. *Effects of Julian date on (a) gap duration, (b) note duration, (c) gap bandwidth, and (d) note bandwidth. See Figure 17 caption for details.*

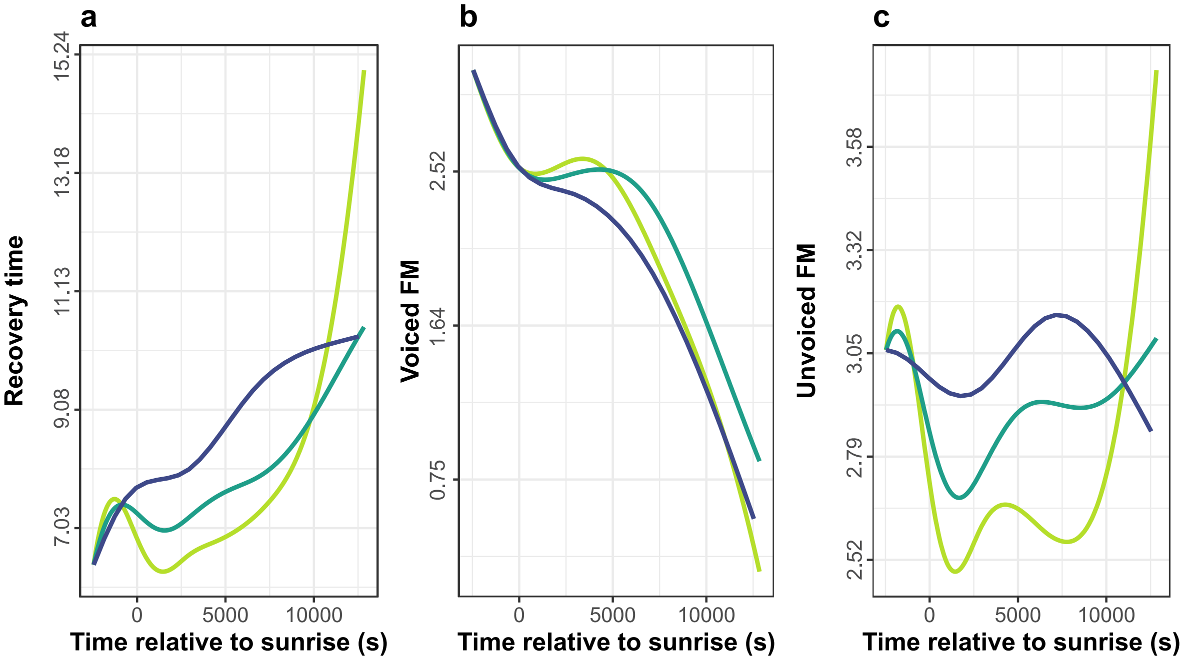


Figure 22. *Plots of simulated data comparing the effects of a low song-rate – a song every 600 seconds – (blue), median song-rate – a song every 60 seconds – (turquoise), and a high song-rate – a song every 30 seconds – (green) to (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation.*

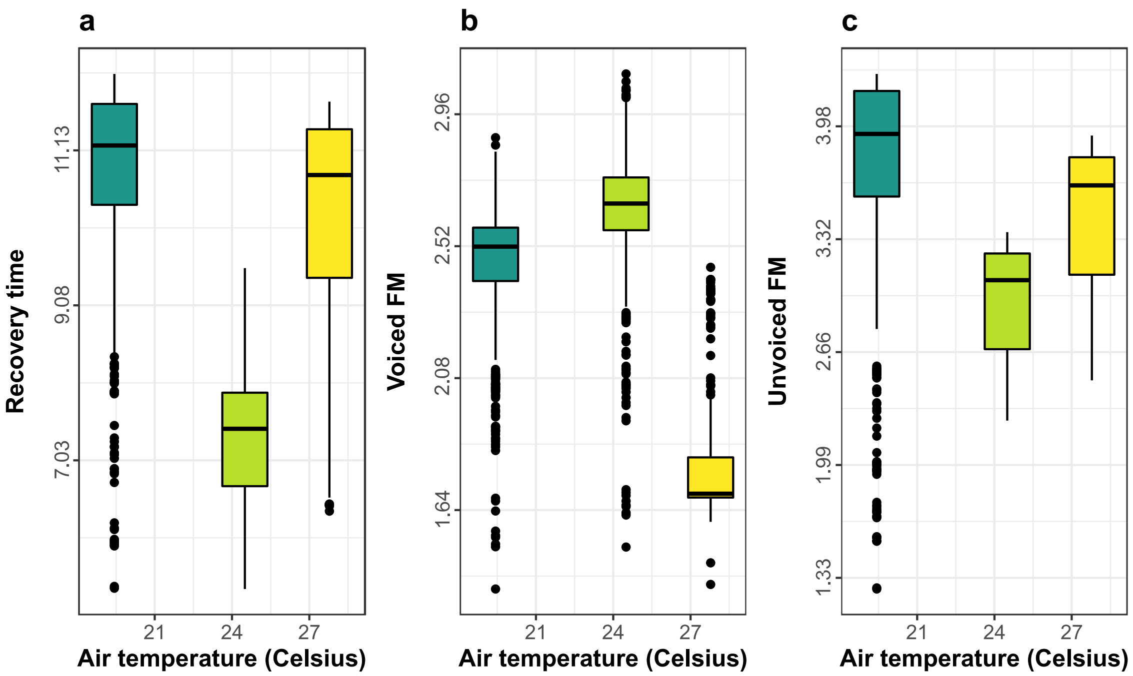


Figure 23. *Boxplot of simulated data comparing the effects of different air temperatures on (a) recovery time, (b) voiced frequency modulation, and (c) unvoiced frequency modulation. The turquoise box represents the 5*th *percentile of Ta (19.44 °C), the green box represents the median Ta (24.49 °C), and the yellow box represents the 95*th *percentile (27.78 °C). Dots show outliers.*

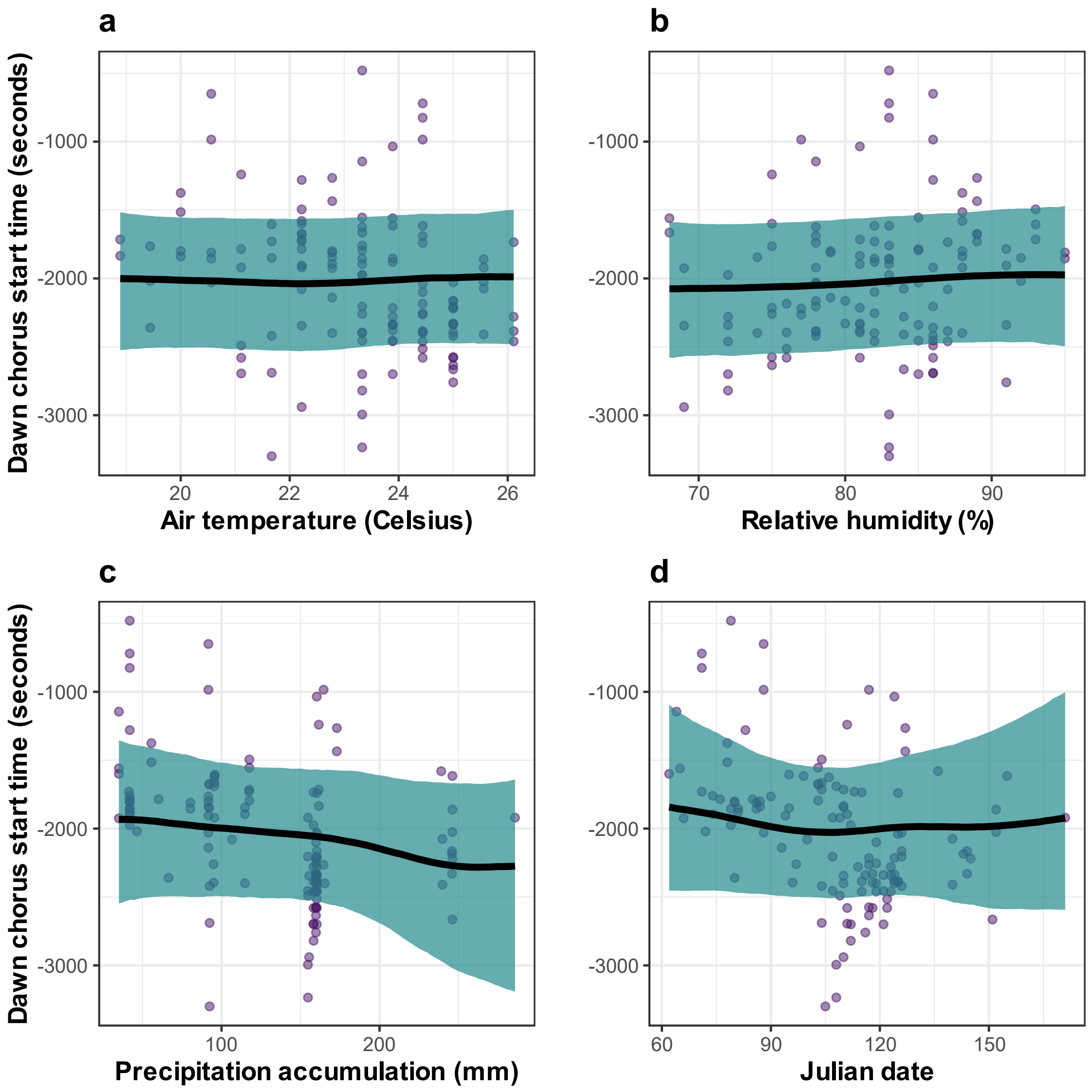


Figure 24. *Effects of (a) air temperature, (b) relative humidity, (c) precipitation accumulation, and (d) Julian date on the dawn chorus start time relative to sunrise. The black line represents the estimate of the effect from each predictor variable to dawn chorus start time. The turquoise shade represents the 95% confidence interval of the estimate, and the semi-transparent purple dots represent the first song for each recording session in 2012, 2017, and 2018. Estimates and confidence intervals were derived from Bayesian mixed models (see text).*

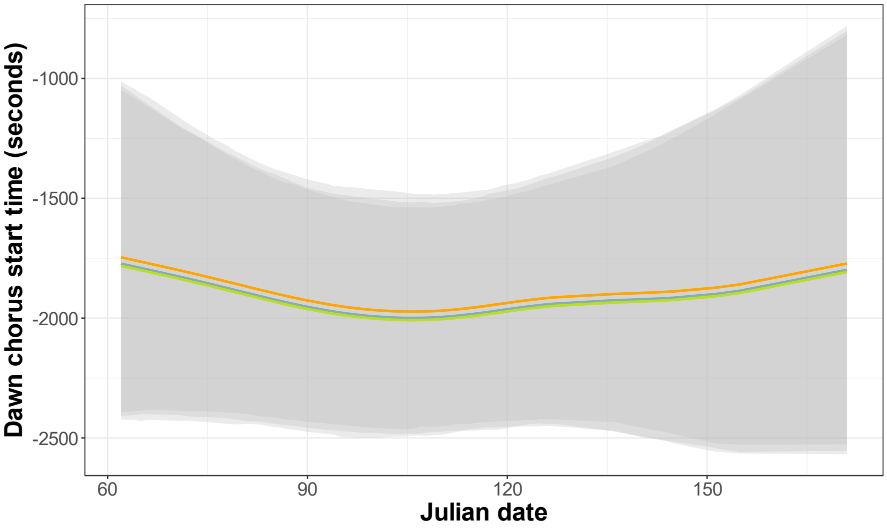


Figure 25. *Plot of simulated data comparing the effects of low (20.00°C; turquoise), median (23.33°C; green), and high (25.56°C; orange) air temperature on the dawn chorus start time relative to sunrise.*

CHAPTER 3: Interpretation, discussion, and conclusions

This chapter includes a broad interpretation of my results. I compare methods from this and other studies and suggest the use of song-rate in future studies that attempt to test the warm-up hypothesis. I explain alternative hypotheses for the evolution of the dawn chorus and how they may interact with the warm-up hypothesis. I also talk about the implications of climate change on song performance. Finally, I compare results from the effects of weather and day of year on the dawn start time of Adelaide's warblers and other species, and present future directions to test predictions of the warm-up hypothesis.

3.1 Methodological considerations

This study showed the value of fine-scale, data-driven vocal performance metrics. Schraft et al. (2017) quantified performance with *FEX*, a theory-driven metric that measures the speed of frequency modulation during both notes and silent gaps (Podos et al. 2016). The three performance measures that Logue et al. (2020) identified in Adelaide's warblers and that I used in this analysis allowed me to examine changes in the speed of frequency modulation separately for notes and silent gaps. This approach revealed that *voiced FM* varies over time differently than does *unvoiced FM*, raising interesting questions about trade-offs between performance metrics and the role of *voiced FM* in Adelaide's warblers' signalling system. The use of fine-scale metrics also revealed important changes in a third performance metric, *recovery time*, that is not accounted for in *FEX*. I conclude that fine scale performance metrics derived from patterns in the data have the potential to provide a more detailed picture of avian vocal performance than do wholistic measures, like *FEX*. Nevertheless, *FEX* has considerable value as an "off the shelf" tool for measuring performance.

I would not have been able to show that the improvement in vocal performance occurs during the dawn chorus without the use of splines, because linear and polynomial models are not able to capture this nuanced pattern of change over time, whereas splines do by making curves in non-linear relationships (Mackenzie et al., 2005; Schraft et al. 2017, Dinh et al. 2020). The only study I am aware of with a similar method to describe changes of acoustic traits over time of day, used Loess smoothing curves to show changes of fundamental frequency and amplitude of zebra finches' songs during the day (Wood et al., 2013). I conclude that Bayesian models with splines are useful tools for studying non-linear change of performance over the course of morning singing.

3.2 Alternative hypotheses for the evolution of the dawn chorus

I failed to reject all three predictions of the warm-up hypothesis that I tested, raising my confidence that the dawn chorus may be an adaptation for vocal warming-up. In the pure form of this hypothesis, females preferentially selected males that sing with high performance songs or performance could have served as a signal to neighbors when re-establishing territories (Cardoso et al., 2007; Catchpole & Slater, 2008; DuBois et al., 2008; Irschick et al., 2008; Janicke et al., 2008; Cardoso et al., 2009; Byers et al., 2010; Bradbury & Vehrencamp, 2011; Botero & de Kort, 2013; Dinh et al., 2020). The dawn chorus may have evolved during an evolutionary arms race among males to warm-up their voices before their neighbors had a chance to do so (Schraft et al., 2017). An important untested assumption of this hypothesis is that high performance songs are attractive to potential mates or repellant to male rivals. The function of song performance in Adelaide's warblers requires further investigation.

It is possible that the dawn chorus may have more than one function. For example, the ancestors of Adelaide's warbler may have evolved early morning singing because their songs traveled farther in the early morning acoustic environment (sound propagation hypothesis; Henwood & Fabrick, 1979; Bradbury & Vehrencamp, 2011); because they still had energy to spend at that time (i.e., surplus of energy; McNamara et al., 1987); or because the temperatures early in the morning were not optimal for foraging – perhaps because insects were not as active (Kacelnik, 1979). Any of these phenomena may have provided the initial selection pressure for the evolution of early morning singing. Once birds started singing earlier in the morning, their vocal performance would have improved due to the warm-up effect, potentially driving further exaggeration of the dawn chorus (Schraft et al., 2017; Dinh et al., 2020).

Two measures of vocal performance diminished or held steady after sunrise. Other hypotheses that attempt to explain the origin of the dawn chorus may help to explain this phenomenon. Birds may have spent all their reserved energy singing at elevated rates during the dawn chorus (Ward et al., 2003). It is also possible that insects were more active when the temperatures started to increase after sunrise, so birds switched to foraging (Kacelnik, 1979). It may also be possible that males re-establish their territories or signal their condition to females during the dawn chorus, so they no longer need to signal at elevated rates after dawn (Mace, 1987a; Krebs & Kacelnik, 1983). Another explanation is that birds experienced fatigue after a high rate of displays. This last hypothesis has been shown in other species. The sonic muscles of oyster toadfish (*Opsanus tau*) experience fatigue after 5 minutes of advertisement calls (Mitchell et al., 2008). Fatigue also affects the locomotor performance in the repetitive courtship display of Cuban burrowing cockroaches (*Byrsotria furnigata*; Mowles & Jepson, 2015). In birds, the ruffed grouse's drumming rate increases with the cumulative number of displays but starts to decrease after a certain number, suggesting fatigue (Déaux et al., 2020). Finally, swamp sparrows have an inconsistent improvement in performance with the cumulative number of songs among different song-types (Dinh et al., 2020). The authors of that study conclude that this inconsistency is likely due to fatigue. There are several hypotheses that have not come to a consensus to explain the origin of the dawn chorus. However, it seems clear that male Adelaide's warblers improve the performance of their songs when they sing at elevated rates during the dawn chorus.

3.3 Song-rate as a driver of warming-up

Schraft et al. (2017) showed a positive relationship between song performance and song order and a negative relationship with time in male Adelaide's warblers. The models in this study included time and order as covariates and focused on the cumulative number of displays to represent warm-up (Schraft et al., 2017). However, results from my analysis suggested that song-rate, a combination of *Time* and *Order*, is the true driver of the vocal warm-up phenomenon.

I suggest including song-rate as a predictor variable in future tests of the warm-up hypothesis. This suggestion is based on a mechanistic consideration of warming-up. It makes sense that performance does not simply increase with *Order*, independent of the time between displays, but rather with the rate at which the individuals display. Support this approach comes from the finding that ruffed grouse drumming performance decreases after animals take a break from displaying (Déaux et al., 2020). Similarly, human soccer players that re-warm-up during the half time of a game have higher sprint performance than players that rest during this period (Mohr et al., 2004).

3.4 Air temperature and performance

Male Adelaide's warblers attained maximum vocal performance at intermediate *Ta's*. The same relationship has been shown in the performance of both ectotherms and endotherms, including signalling behaviour (Schulte et al., 2011 Cunningham et al., 2013; Brandt et al., 2018; Levesque & Marshall, 2021). Jumping spider sexual displays are another example of an animal display with a peaked performance vs. *Ta* relationships (Brandt et al., 2018). In humans, exercise performance is also negatively affected by cold or hot conditions (No & Kwak, 2016). Results from the current study give more evidence to the curvilinear relationship between performance and temperature, which is well known in ectotherms but not as well in endotherms. The effects of *Ta* on the vocal performance of male Adelaide's warblers suggest the importance to understanding the physiological mechanisms of warming-up and how they relate to extrinsic factors such as *Ta*, to better understand if there is a link between a physical warm-up and a behavioural warm-up.

While my dataset does not allow me to speak about performance under extreme temperatures, the finding that male Adelaide's warbler's vocal performance diminishes at "normal" high *Ta,* suggests that continued global warming may have detrimental effects on mean vocal performance. In birds, climate change is associated with changes in migratory patterns, altered reproductive phenology, community shifts to habitats with better resources, among others (Marra et al., 2004; Yang & Rudolf, 2010; Harmon & Barton, 2013). Little is known about its effect on signalling behaviours. It is necessary to better understand the plasticity in thermoregulation of tropical birds, to estimate to what degree an increase in air temperature will impact their communication systems.

3.5 Dawn chorus start time

Ta seems to have no effect on the time male Adelaide's warblers start to sing. There are mixed results in terms of the effects of *Ta* on the dawn chorus start time of other species. Like Adelaide's warblers, Chinese Hwamei (*Garrulax canorus*) do not seem to change dawn chorus start time based on *Ta* (Puswal et al., 2021). In contrast, song thrushes (*Turdus philomelos*), cerulean warblers (*Setophaga cerulea*), house wrens (*Troglodytes aedon*), Alström's warblers (*Phylloscopus soror*), streak-breasted scimitar babblers

(*Pomatorhinus ruficolli*), and brownish-flanked bush warblers (*Cettia fortipes*) start to sing later as *Ta* increases (Da Silva et al., 2014; MacDonald & Islam, 2021; Stuart et al., 2019; Puswal et al., 2021). Seven temperate species birds exhibit the opposite effect, singing earlier with increasing temperatures (Bruni et al. 2014; Hasan & Badri, 2016). The highly variable effects of *Ta* on dawn chorus start time may be explained by the difference in temperatures that these species are exposed to (e.g., temperate versus tropical birds) and the methods they use to thermoregulate. A comparative analysis on the effects of air temperature on singing behaviour would be useful.

The effect of Julian date on dawn chorus start time also varies among species. Black-capped chickadees (*Poecile atricapillus*) start their dawn chorus earlier during the middle of the period of the time they were recorded (from April to July 2012; Bruni et al., 2014). Song thrushes, cerulean warblers, house wrens, Alström's warblers, streak-breasted scimitar babblers, and brownish-flanked bush warblers start to sing earlier as the days progress during the period they were recorded (from February to April; Da Silva et al., 2014). Conversely, dawn chorus start time did not covary with Julian date in Alström's warbler, streak-breasted scimitar babbler, brownish-flanked bush warbler, and Chinese Hwamei (Puswal et al. 2021). Finally, song sparrows (*Melospiza melodia*), white-throated sparrows (*Zonotrichia albicollis*), and Eastern phoebes (*Sayournis phoebe*) start to sing later as the season progress (from April to July 2012; Bruni et al., 2014). This contrast in the effect of *JD* on the singing behaviour of several species may be explained by different methods and sample sizes when recording the birds. It may also be explained by their different mating systems or territory defense periods (Puswal et al., 2021). Again, an interspecific comparative analysis could clarify this variation in the effects of date on different bird species. I am not aware of many studies incorporating the effects of humidity and precipitation accumulation on the dawn chorus start time of other species, although some studies consider precipitation at the time that singing behaviour was recorded (Bruni et al., 2014; MacDonald & Islam, 2021).

Understanding the effects of weather and other extrinsic factors on singing behaviour may also clarify their effects on vocal performance. If the dawn chorus is, in fact, the outcome of an evolutionary arms race to behaviourally warm-up (Schraft et al., 2017), and birds start to sing late due to some external factor, that may affect the performance of their communication. *Ta* seems to have no effect on the time birds start to sing, but there was some evidence they started to sing later with higher *RH* and they avoided singing when it was raining. It would be interesting to test whether birds that start to sing earlier also perform better.

3.6 Future directions

The function of vocal performance requires further investigation. Future playback studies with Adelaide's warblers should examine if performance functions in mate attraction or territory defense. A top priority is to test female preference for songs with different durations of silent gaps relative to notes to determine if the shortness of these silent gaps serves to signal male condition and if females pay attention to those signals. Further investigation is also required to test the function of performance variation among song-types. We could test, for example, the hypothesis that birds with higher performance songs have a benefit in song-type matches (Logue & Forstmeier, 2008).

The physiological mechanisms of vocal warm-up remain largely unknown. A key question here is the degree to which physical warming underlies behavioural warm-up. One valuable approach would be to measure changes in body temperature when birds are singing at different rates. Such a study would have implications for the mechanisms and benefits of warming-up in other species, including humans.

It may be useful to incorporate other variables that may affect the time birds start to sing such as solar radiation. It may also be necessary to measure weather parameters at the territories of the birds recorded, to control for the variation of temperature, humidity, etc. due to different habitat conditions such as vegetation cover.

3.7 Summary

In this thesis, I described variation in male Adelaide's warblers' vocal performance over time and across different weather conditions. I failed to reject three important predictions of the warm-up hypothesis and concluded that vocal warm-up is a probable function of the dawn chorus in this species. My findings suggest that display-rate is a useful measure of recent practice. I found that one of the acoustic traits that underlies vocal performance, *gap duration*, may be particularly important in this species. My study emphasizes the value of considering the mechanism of vocal performance, and suggests further refinement of this approach, such as considering which side of the syrinx produces each sound. Finally, this thesis helped to better understand the effects of extrinsic factors on singing behaviour and performance in a tropical bird, and warns that global climate change has the potential to affect display performance.

REFERENCES

Amir, O., Amir, N., & Michaeli, O. (2005). Evaluating the influence of warmup on singing voice quality using acoustic measures. *Journal of Voice, 19*(2), 252–260.

Araya-Salas, M. (2021). dynaSpec: dynamic spectrogram visualizations. R package version 1.0.1. https://CRAN.R-project.org/package=dynaSpec

Ballentine, B., Hyman, J., & Nowicki, S. (2004). Vocal performance influences female response to male bird song: an experimental test. *Behavioral Ecology*, *15*(1), 163–168.

Barnett, C. A., & Briskie, J. V. (2006). Energetic state and the performance of dawn chorus in silvereyes (Zosterops lateralis). *Behavioral Ecology and Sociobiology*, *61*(4), 579–587.

Bishop, D. (2003a). Warm Up I. Potential mechanisms and the effects of passive warm up on exercise performance. *Sports Medicine*, *33*, 439–454.

Bishop, D. (2003b). Warm Up II. Performance changes following active warm up and how to structure the warm up. *Sports Medicine*, *33*, 483–498.

Botero, C. A., & de Kort, S. R. (2013). Learned signals and consistency of delivery: a case against receiver manipulation in animal communication. *Animal Communication Theory,* U. 281–296. Cambridge: Cambridge University Press.

Botero, C. A., Rossman, R. J., Caro, L. M., Stenzler, L. M., Lovette, I. J., de Kort, S. R., & Vehrencamp, S. L. (2009). Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Animal Behaviour*, *77*(3), 701–706.

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication*. Sunderland, MA: Sinauer Associates.

Brandt, E. E., Kelley, J. P., & Elias, D. O. (2018). Temperature alters multimodal signaling and mating success in an ectotherm. *Behavioral Ecology and Sociobiology*, *72*(12). https://doi.org/10.1007/s00265-018-2620-5

Bruni, A., Mennill, D. J., & Foote, J. R. (2014). Dawn chorus start time variation in a temperate bird community: relationships with seasonality, weather, and ambient light. *Journal of Ornithology*, 155(4), 877-890.

Bürkner, P. C. (2018). Advanced Bayesian Multilevel modeling with the R Package brms. *The R Journal*, 10(1), 395-411.

Byers, B. E. (2007). Extrapair paternity in chestnut-sided warblers is correlated with consistent vocal performance. *Behavioral Ecology*, *18*(1), 130–136.

Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour*, 79(4), 771-778.

Cambridge University Press. (n.d.). Performance. In *Cambridge dictionary*. Retrieved April 13, 2021 from https://dictionary.cambridge.org/dictionary/english/performance

Cardoso, G. C. (2013). Using frequency ratios to study vocal communication. *Animal Behaviour, 85*(6), 1529–1532. https://doi.org/10.1016/j.anbehav.2013.03.044

Cardoso, G. C. (2017). Advancing the inference of performance in birdsong. *Animal Behaviour, 125*, e29- e32.

Cardoso, G. C., Atwell, J. W., Ketterson, E. D., & Price, T. D. (2007). Inferring performance in the songs of dark-eyed juncos (Junco hyemalis). *Behavioral Ecology, 18*(6), 1051-1057.

Cardoso, G. C., Atwell, J. W., Ketterson, E. D., & Price, T. D. (2009). Song types, song performance, and the use of repertoires in dark-eyed juncos (Junco hyemalis). *Behavioral Ecology*, *20*(4), 901–907.

Catchpole, C. K., & Slater, P. J. B. (2008). *Bird song: Biological themes and variations*. Cambridge: Cambridge University Press.

Center for Conservation Bioacoustics. (2019). Raven Pro: Interactive Sound Analysis Software (Version 1.6.1) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. Available from http://ravensoundsoftware.com/

Clarke, A., & Rothery, P. (2008). Scaling of body temperature in mammals and birds. *Functional Ecology*, 22(1), 58-67.

Clarkson, C. E. (2007). Food supplementation, territory establishment, and song in the prothonotary warbler. *The Wilson Journal of Ornithology*, *119*(3), 342–349. https://doi.org/10.1676/04-073.1

Clink, D. J., Hamid Ahmad, A., & Klinck, H. (2020). Gibbons aren't singing in the rain: Presence and amount of rainfall influences ape calling behavior in Sabah, Malaysia. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-57976-x

Coomes, C. M., & Derryberry, E. P. (2021). High temperatures reduce song production and alter signal salience in songbirds. *Animal Behaviour, 180*, 13–22.

Costanzo, J. P. (2011) Extreme Cold Hardiness in Ectotherms. *Nature Education Knowledge* 3(10):3

Cramer, E. R., & Jordan Price, J. (2007). Red-winged blackbirds Ageliaus phoeniceus respond differently to song types with different performance levels. *Journal of Avian Biology*, *38*(1), 122–127.

Cunningham, S. J., Martin, R. O., Hojem, C. L., & Hockey, P. A. (2013). Temperatures in excess of critical thresholds threaten nestling growth and survival in a rapidly-warming arid savanna: A study of common fiscals. *PLoS ONE, 8*(9). https://doi.org/10.1371/journal.pone.0074613

Da Silva, A., Samplonius, J. M., Schlicht, E., Valcu, M., & Kempenaers, B. (2014). Artificial night lighting rather than traffic noise affects the daily timing of dawn and dusk singing in common European songbirds. *Behavioral Ecology, 25*(5), 1037–1047.

De Kort, S. R., Eldermire, E. R. B., Valderrama, S., Botero, C. A., & Vehrencamp, S. L. (2009). Trill consistency is an age-related assessment signal in banded wrens. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1665), 2315–2321.

Déaux, E. C., O'Neil, N. P., Jensen, A. M., Charrier, I., & Iwaniuk, A. N. (2020). Courtship display speed varies daily and with body size in the Ruffed Grouse (Bonasa umbellus). *Ethology,* 126(5), 528- 539.

Demery, A.-J. C., Burns, K. J., & Mason, N. A. (2021). Bill size, Bill shape, and body size constrain bird song evolution on a macroevolutionary scale. *Ornithology*, *138*(2). https://doi.org/10.1093/ornithology/ukab011

Dinh, J. P., Peters, S., & Nowicki, S. (2020). Song performance improves with continued singing across the morning in a songbird. *Animal Behaviour*, *167,* 127–137.

Drăgănoiu, T. I., Nagle, L., & Kreutzer, M. (2002). Directional female preference for an exaggerated male trait in canary (Serinus canaria) song. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 269*(1509), 2525–2531.

DuBois, A. L., Nowicki, S., & Searcy, W. A. (2008). Swamp sparrows modulate vocal performance in an aggressive context. *Biology Letters, 5*(2), 163–165.

DuBois, A. L., Nowicki, S., & Searcy, W. A. (2011). Discrimination of vocal performance by male swamp sparrows. *Behavioral Ecology and Sociobiology*, *65*(4), 717–726.

Elliot, N., Sundberg, J., & Gramming, P. (1995). What happens during vocal warm-up? *Journal of Voice*, *9*(1), 37–44.

Foerster, K., Poesel, A., Kunc, H., & Kempenaers, B. (2002). The natural plasma testosterone profile of male blue tits during the breeding season and its relation to song output. *Journal of Avian Biology*, *33*(3), 269–275.

Forstmeier, W., Kempenaers, B., Meyer, A., & Leisler, B. (2002). A novel song parameter correlates with extra-pair paternity and reflects male longevity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1499), 1479–1485.

Fradkin, A. J., Zazryn, T. R., & Smoliga, J. M. (2010). Effects of warming-up on physical performance: A systematic review with meta-analysis. *Journal of Strength and Conditioning Research*, *24*(1), 140–148.

Geberzahn, N., & Aubin, T. (2014). How a songbird with a continuous singing style modulates its song when territorially challenged. *Behavioral Ecology and Sociobiology*, *68*(1), 1–12.

Gelman, A. B., & Hill, J. (2018). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., & Shirley, K. (2011). Inference from simulations and monitoring convergence. In S. Brooks, A. Gelman, G. L. Jones, & X.-L., Meng (Eds.), Handbook of Markov Chain Monte Carlo (p. 434). Chapman and Hall/CRC. https://doi.org/10.1201/b10905-7

Gerson, A. R., Smith, E. K., Smit, B., McKechnie, A. E., & Wolf, B. O. (2014). The impact of humidity on evaporative cooling in small desert birds exposed to high air temperatures. *Physiological and Biochemical Zoology*, *87*(6), 782–795.

Gil, D., & Gahr, M. (2002). The honesty of bird song: multiple constraints for multiple traits. *Trends in Ecology & Evolution*, *17*(3), 133–141.

Gil, D., & Llusia, D. (2020). The Bird Dawn Chorus Revisited. *Coding Strategies in Vertebrate Acoustic Communication*, 45–90. https://doi.org/10.1007/978-3-030-39200-0_3

Goller, F. (2021). Vocal Athletics – from birdsong production mechanisms to sexy songs. *Animal Behaviour*. https://doi.org/10.1016/j.anbehav.2021.04.009

Google Earth (2021). [Location of the Cabo Rojo National Wildlife Refuge and the MesoWest weather station]. Retrieved November 25, 2021, from https://earth.google.com/web/search/17.969997434293855,+-67.16000318527222/@17.97308973,-67.16718521,20.07055548a,6077.98462868d,35y,0h,0t,0r/data=CigiJgokCZkYkLxl-zFAEfOIEnJE-TFAGdvrhA7tyVDAIXWwMwwXy1DA

Hao, L., & D. Q. Naiman (2007). *Quantile regression*. Thousand Oaks, Calif., Sage Publications.

Harmon, J. P., & Barton, B. T. (2013). On their best behavior: How animal behavior can help determine the combined effects of species interactions and climate change. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.12192

Hartley R., & Suthers R. A. (1989). Airflow and pressure during canary song: direct evidence for mini-breaths. *Journal of Comparative Physiology A*, 165(1), 15-26.

Hasan, N. (2010). The effect of environmental conditions on start of dawn singing of blackbirds (*Turdus merula*) and bulbuls (*Pyconotidae*). *Jordan Journal of Biological Sciences,* 3, 13,16.

Hasan, N. M., & Badri, M. (2016). Effect of ambient temperature on dawn chorus of house sparrows. *Environment and Ecology Research*, 4(3), 161-168.

Henwood, K., & Fabrick, A. (1979). A Quantitative Analysis of the Dawn Chorus: Temporal Selection for Communicatory Optimization. *The American Naturalist*, *114*(2), 260–274.

Hoese, W. J., Podos, J., Boetticher, N. C., & Nowicki, S. (2000). Vocal tract function in birdsong production: experimental manipulation of beak movements. *Journal of Experimental Biology*, *203*(12), 1845–1855.

Hunt, K. E., Hahn, T. P., & Wingfield, J. C. (1997). Testosterone implants increase song but not aggression in male Lapland Longspurs. *Animal Behaviour*, *54*(5), 1177 -1192.

Imdad, M. U., & Aslam, M. (2018). Mctest: multicollinearity diagnostic measures. https://CRAN.R-project.org/package=mctest, R package version 1.3.1

Imdadullah, M., Aslam, M., & Altaf, S., (2016). Mctest: an R package for detection of collinearity among regressors. *The R Journal*, 8(2), 499-509.

Irschick, D.J., Meyers, J.J., Husak, J. F., & Le Galliard, J.-F. (2008). How does selection operate on whole-organism functional performance capacities? A review and synthesis. *Evolutionary Ecology Research*, 10, 177–196.

Janicke, T., Hahn, S., Ritz, M. S., & Peter, H.-U. (2008). Vocal performance reflects individual quality in a nonpasserine. *Animal Behaviour*, 75(1), 91-98.

Kacelnik, A. (1979). The foraging efficiency of great tits (Parus major L.) in relation to light intensity. *Animal Behaviour*, *27*, 237–241.

Kaluthota, C.D., Medina, O. J., Logue, D. M. (2019). Quantifying song categories in Adelaide's warbler (Setophaga adelaidae). *Journal of Ornithology*. 160(2), 305-315.

Kassambara, A. (2020). Ggpubr: 'ggplot2' based publication ready plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr

Krebs, J. R., & Kacelnik, A. (1983). The dawn chorus in the great tit (Parus Major): Proximate and ultimate causes. *Behaviour*, *83*(3-4), 287–308.

Kunc, H. P., Amrhein, V., & Naguib, M. (2006). Vocal interactions in nightingales, *Luscinia megarhynchos*: More aggressive males have higher pairing success. *Animal Behaviour*, *72*(1), 25–30.

Lachlan, R. (2007). Luscinia: a bioacoustics analysis computer program: Version 2.14.

Lachlan, R. (2015). *Rflachlan/Luscinia*. GitHub. https://github.com/rflachlan/Luscinia/wiki/Measurements

Levesque, D. L., & Marshall, K. E. (2021). Do endotherms have thermal performance curves? *Journal of Experimental Biology, 224*(3). https://doi.org/10.1242/jeb.141309

Logue, D. M., & Forstmeier, W. (2008). Constrained performance in a communication network: Implications for the function of song-type matching and for the evolution of multiple ornaments. *The American Naturalist, 172*(1), 34–41.

Logue, D. M., Sheppard, J. A., Walton, B., Brinkman, B. E., & Medina, O. J. (2020). An analysis of avian vocal performance at the note and song levels. *Bioacoustics*. https://doi.org/10.1101/664896

MacDonald, G. J., & Islam, K. (2021). Do social factors explain seasonal variation in dawn song characteristics of paired male Cerulean Warblers (Setophaga cerulea)? *Bioacoustics, 30*(1), 1–16.

Mace, R. (1987). The dawn chorus in the great tit Paras major is directly related to female fertility. *Nature, 330*(6150), 745–746.

Mackenzie, M. L., Donovan, C. R., & McArdle, B. H. (2005). Regression spline mixed models: A forestry example. *Journal of Agricultural, Biological, andEnvironmental Statistics, 10*(4), 394–410.

Marra, P. P., Francis, C. M., Mulvihill, R. S., & Moore, F. R. (2004). The influence of climate on the timing and rate of Spring Bird Migration. *Oecologia, 142*(2), 307-315.

McElreath, R. (2019). *Statistical Rethinking*. CRC Press.

McKechnie, A. E., Gerson, A. R., McWhorter, T. J., Smith, E. K., Talbot, W. A., & Wolf, B. O. (2017). Avian thermoregulation in the heat: Evaporative cooling in five Australian passerines reveals within-order biogeographic variation in heat tolerance. *Journal of Experimental Biology*. https://doi.org/10.1242/jeb.155507

McNamara, J. M., Mace, R. H., & Houston, A. I. (1987). Optimal daily routines of singing and foraging in a bird singing to attract a mate. *Behavioral Ecology and Sociobiology, 20*(6), 399–405.

McWhorter T. J., Gerson, A. R., Talbot, W. A., Smith, E. K., McKechnie, A. E. & Wolf, B. O. (2018). Avian thermoregulation in the heat_evaporative cooling capacity and thermal tolerance in two Australian parrots. *Journal of Experimental Biology,* 221, jeb168930.

Merola-Zwartjes, M., & Ligon, J. D. (2000). Ecological energetics of the Puerto Rican Tody: heterothermy, torpor, and intra-island variation. *Ecology*, *81*(4), 990–1003.

Mitchell, S., Poland, J., & Fine, M. L. (2008). Does muscle fatigue limit advertisement calling in the oyster toadfish opsanus tau? *Animal Behaviour, 76*(3), 1011–1016.

Mohr, M., Krustrup, P., Nybo, L., Nielsen, J. J., & Bangsbo, J. (2004). Muscle temperature and sprint performance during soccer matches – beneficial effect of re-warm-up at half-time. *Scandinavian Journal of Medicine and Science in Sports, 14*(3), 156–162.

Mota, P. G., & Cardoso, G. C. (2001). Song organisation and patterns of variation in the serin (Serinus serinus). *Acta Ethologica,* 3(2), 141-150.

Motel, T., Fisher, K. V., & Leydon, C. (2003). Vocal Warm-up Increases Phonation Threshold Pressure in Soprano Singers at High Pitch. *Journal of Voice*, *17*(2), 160–167.

Mowles, S. L., & Jepson, N. M. (2015). Physiological costs of repetitive courtship displays in cockroaches handicap locomotor performance. *PLOS ONE*, *10*(11). https://doi.org/10.1371/journal.pone.0143664

National Oceanic and Atmospheric Administration. (n.d.). *Global Historical Climatology Network daily (GHCNd).* https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily

No, M., & Kwak, H.-B. (2016). Effects of environmental temperature on physiological responses during submaximal and maximal exercises in soccer players. *Integrative Medicine Research*, 5(3), 216- 222.

Nord, A., Nilsson, J. F., Sandell, M. I., & Nilsson, J. (2009). Patterns and dynamics of rest-phase hypothermia in wild and captive blue tits during winter. *Journal of Comparative Physiology B*, 179, 737–745.

Nordt, A., & Klenke, R. (2013). Sleepless in town – drivers of the temporal shift in dawn song in Urban European Blackbirds. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071476

Nowicki, S., Westneat, M., & Hoese, W. (1992). Birdsong: motor function and the evolution of communication. *Seminars in the Neurosciences, 4*(6), 385-390.

Oke, T. R. (1978). *Boundary layer climates.* Methuen; Wiley.

Podos, J. (1996). Motor constraints on vocal development in a songbird. *Animal Behaviour*, *51*(5), 1061–1070.

Podos, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution*, *51*(2), 537. https://doi.org/10.2307/2411126

Podos, J. (2001). Correlated evolution of morphology and vocal signal structure in Darwins finches. *Nature,* 409(6817), 185-188.

Podos, J. (2017). Birdsong performance studies: reports of their death have been greatly exaggerated. *Animal Behaviour, 125,* e17-e24.

Podos, J., & Nowicki, S. (2004). Beaks, adaptation, and vocal evolution in Darwin's finches. *BioScience, 54*(6), 501. https://doi.org/10.1641/0006-3568(2004)054[0501:baavei]2.0.co;2

Podos, J., & Sung, H.-C. (2020). Vocal Performance in Songbirds: From Mechanisms to Evolution. *The Neuroethology of Birdsong*, 245–268.

Podos, J., Moseley, D. L., Goodwin, S. E., McClure, J., Taft, B. N., Strauss, A. V. H., Rega-Brodsky, C., & Lahti, D. C. (2016). A fine-scale, broadly applicable index of vocal performance: frequency excursion. *Animal Behaviour, 116*, 203–212.

Puswal, S. M., Jinjun, M., & Liu, F. (2021). Effects of temperature and season on birds' dawn singing behavior in a forest of eastern China. *Journal of Ornithology, 162*(2)*,* 447–459.

Racinais, S., Cocking, S., & Périard, J. D. (2017). Sports and environmental temperature: From warming-up to heating-up. *Temperature*, *4*(3), 227–257.

Rashotte, M. E., Sedunova, E. V., Johnson, F., & Pastukhov, I. F. (2001). Influence of food and water availability on undirected singing and energetic status in adult male zebra finches (Taeniopygia guttata). *Physiology & Behavior*, *74*(4-5), 533- 541.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Reeping, P. M., & Hemenway, D. (2020). The association between weather and the number of daily shootings in Chicago (2012–2016). *Injury Epidemiology*, *7*(1). https://doi.org/10.1186/s40621-020-00260-3

Schraft, H. A., Medina, O. J., McClure, J., Pereira, D. A., & Logue, D. M. (2017). Within-day improvement in a behavioural display: wild birds 'warm up'. *Animal Behaviour, 124*, 167-174.

Schulte, P. M., Healy, T. M., & Fangue, N. A. (2011). Thermal Performance Curves, Phenotypic Plasticity, and the Time Scales of Temperature Exposure. *Integrative and Comparative Biology*, *51*(5), 691–702.

Singh, R., Prathibha, P., & Jain, M. (2020). Effect of temperature on life-history traits and mating calls of a field cricket, Acanthogryllus Asiaticus. *Journal of Thermal Biology, 93,* 102740. https://doi.org/10.1016/j.jtherbio.2020.102740

Staicer, C. A. (1991). *The role of male song in the socioecology of the tropical resident Adelaide's warbler (Dendroica adelaidae)*. Amherst (MA): University of Massachusetts

Staicer, C. A., Spector, D. A., & Horn, A. G. (1996). The dawn chorus and other diel patterns in acoustic signaling. *Ecology and Evolution of Acoustic Communication in Birds*, 426–453. https://doi.org/10.7591/9781501736957-033

Strain, J. G., & Mumme, R. L. (1988). Effects of food supplementation, song playback, and temperature on vocal territorial behavior of carolina wrens. *The Auk*, 105(1), 11-16.

Stuart, C. J., Grabarczyk, E. E., Vonhof, M. J., & Gill, S. A. (2019). Social factors, not anthropogenic noise or artificial light, influence onset of dawn singing in a common songbird. *The Auk*, *136*(3). https://doi.org/10.1093/auk/ukz045

Sueur, J., Aubin T., & Simonis C. (2008). Seewave: a free modular tool for sound analysis and synthesis. *Bioacoustics*, 18: 213-226.

Suthers, R. A., & Zollinger, S. A. (2004). Producing song: the vocal apparatus. *Annals of the New York Academy of Sciences,* 1016(1), 109-129.

Suthers, R. A., & Zollinger, S. A. (2008). From brain to song: The vocal organ and vocal tract. In H. P. Zeigler, & P. Marler (Eds.), Neuroscience of birdsong (pp. 78e98). Cambridge, U.K.: Cambridge University Press.

Tan, C. L., & Knight, Z. A. (2018). Regulation of body temperature by the nervous system. *Neuron*, *98*(1), 31–48.

Textor, J., & Van der Zander, B. (2016). Dagitty: Graphical Analysis of Structural Causal Models. R package version 0.2-2. https://CRAN.R-project.org/package=dagitty

The University of Utah. (2021, June 18). *Weather conditions for CRRP4*. Meso West. https://mesowest.utah.edu/cgi-bin/droman/meso_base_dyn.cgi?stn=CRRP4

Toms, J. D. (2020). Adelaide's warbler (Setophaga Adelaidae). *Birds of the World*. https://doi.org/10.2173/bow.adewar1.01

Trillo, P. A., & Vehrencamp, S. L. (2005). Song types and their structural features are associated with specific contexts in the banded wren. *Animal Behaviour, 70*(4), 921–935.

United States Environmental Protection Agency (2012, August). What climate change means for Puerto Rico. https://19january2017snapshot.epa.gov/sites/production/files/2016-09/documents/climate-change-pr.pdf

US Department of Commerce, N. O. A. A. (2013, June 21). *How does the ocean affect climate and weather on land?* Ocean Exploration Facts: NOAA Office of Ocean Exploration and Research. https://oceanexplorer.noaa.gov/facts/climate.html

US Department of Commerce, N. O. A. A. (2015, June 13). *Discussion on humidity*. National Weather Service. https://www.weather.gov/lmk/humidity

Vallet, E., & Kreutzer, M. (1995). Female canaries are sexually responsive to special song phrases. *Animal Behaviour*, *49*(6), 1603–1610.

Vehrencamp, S. L., Yantachka, J., Hall, M. L., & de Kort, S. R. (2013). Trill performance components vary with age, season, and motivation in the banded wren. *Behavioral Ecology and Sociobiology*, *67*(3), 409–419.

Wada, H. (2010). The development of birdsong. *Nature Education Knowledge*, 3(10):86.

Ward, S., Speakman, J. R., & Slater, P. J. B. (2003). The energy cost of song in the canary, serinus canaria. *Animal Behaviour*, *66*(5), 893–902.

Welling, P., Koivula, K., & Lahti, K. (1995). The dawn chorus is linked with female fertility in the Willow Tit Parus montanus. *Journal of Avian Biology*, *26*(3), 241.

Westreich, D., & Greenland, S. (2013). The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, *177*(4), 292-298.

Wickham, H. (2016). Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wild, J., Goller, F., & Suthers, R. (1998). Inspiratory muscle activity during bird song. *Journal of Neurobiology,* 36(3), 441-453.

Wood, W. E., Osseward, P. J., Roseberry, T. K., & Perkel, D. J. (2013). A daily oscillation in the fundamental frequency and amplitude of harmonic syllables of zebra finch song. *PloS ONE*, *8*, e82327.

Wu, H., Gong, L., Jiang, T., Feng, J., & Lin, A. (2021). Echolocation call frequencies of bats vary with body temperature and weather conditions. *Animal Behaviour, 180,* 51–61. https://doi.org/10.1016/j.anbehav.2021.08.005

Yang, L. H., & Rudolf, V. H. (2010). Phenology, ontogeny and the effects of climate change on the timing of species interactions. *Ecology Letters*, *13*(1), 1–10.

Zipple, M. N., Nowicki, S., Searcy, W. A., & Peters, S. (2019). Full life course analysis of birdsong reveals maturation and senescence of highly repeatable song characteristics. *Behavioral Ecology*, *30*(6), 1761–1768.

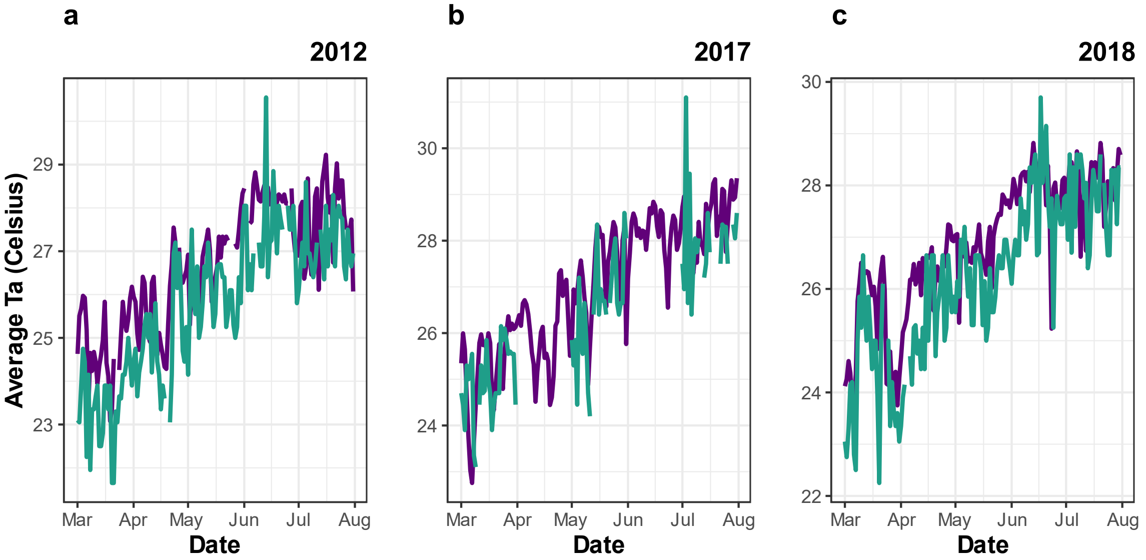
APPENDIX 1: Number of songs in each analysis by bird identification

Appendix 1: *Bird identification, year the bird was recorded, and # of songs from that bird used in each of the analyses (i.e., song performance and dawn chorus start time).*

Bird ID	Year	Songs in performance	Songs in DCST
DbWY	2017	32	3
DDLb	2012	605	4
DgDgY	2017, 2018	62 (2017)	6
DgLLb	2017, 2018	70	4
KLb	2017	62	3
KYK	2012	110	4
LbLgLb	2017, 2018	23	4
LgLbRb	2017, 2018	40 (2017)	6
LgLLb	2018	0	3
LgRLg	2012	255	4
LgWV	2012	254	3
LLbLg	2017, 2018	79 (2017)	5
LpOTb	2017, 2018	44	4
OKO	2018	0	4
OPR	2017, 2018	83	7
OWO	2012, 2018	461	5
PDP	2012	139	4
POR	2017, 2018	80	4

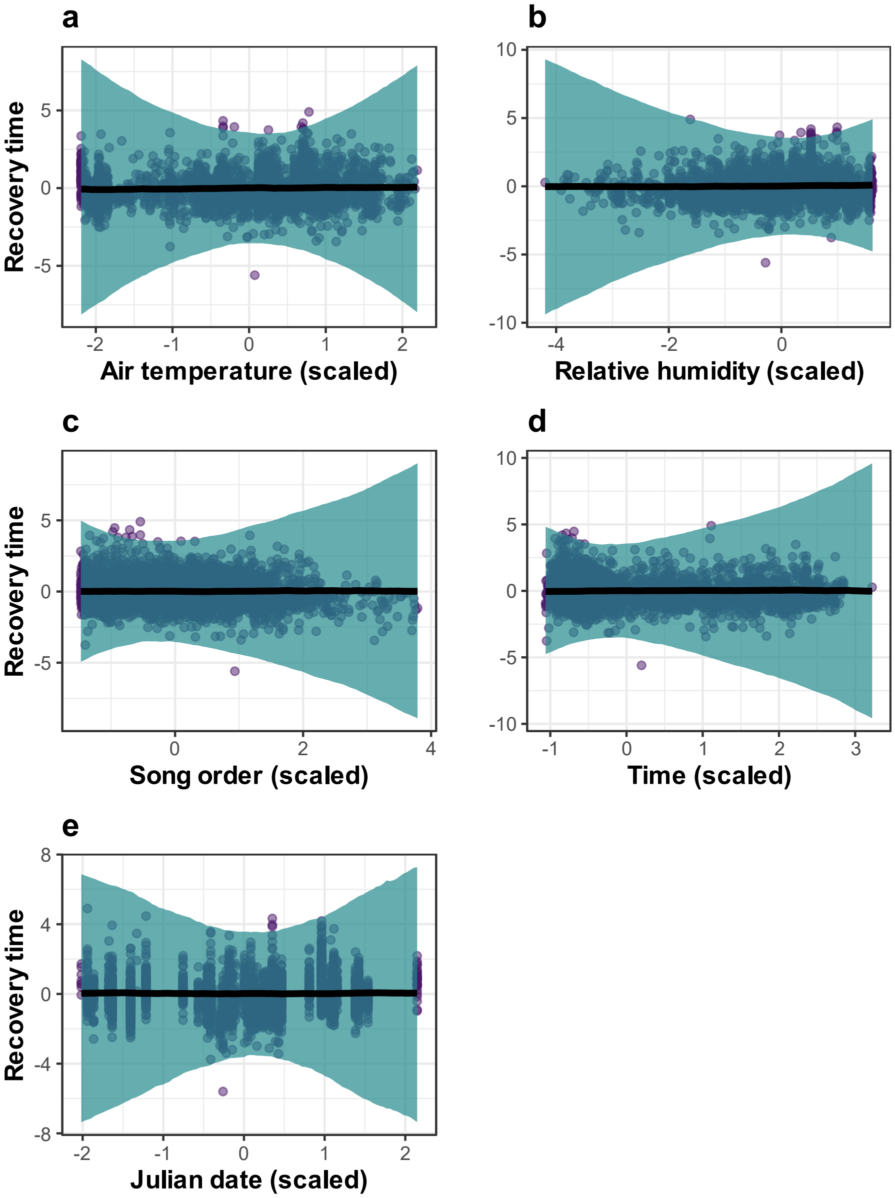
RbRbO	2012, 2018	425	6
RbTbL	2017	0	1
RDY	2012	509	4
ROLb	2012	121	4
RRO	2017, 2018	30 (2017)	8
TbRbRb	2017	39	3
TbWLb	2017	66	3
WTbTb	2017, 2018	0	5
YO	2017	50	3
YWW	2017, 2018	35	6
YYLb	2017, 2018	93	5

APPENDIX 2: Comparison of air temperature measures between weather stations



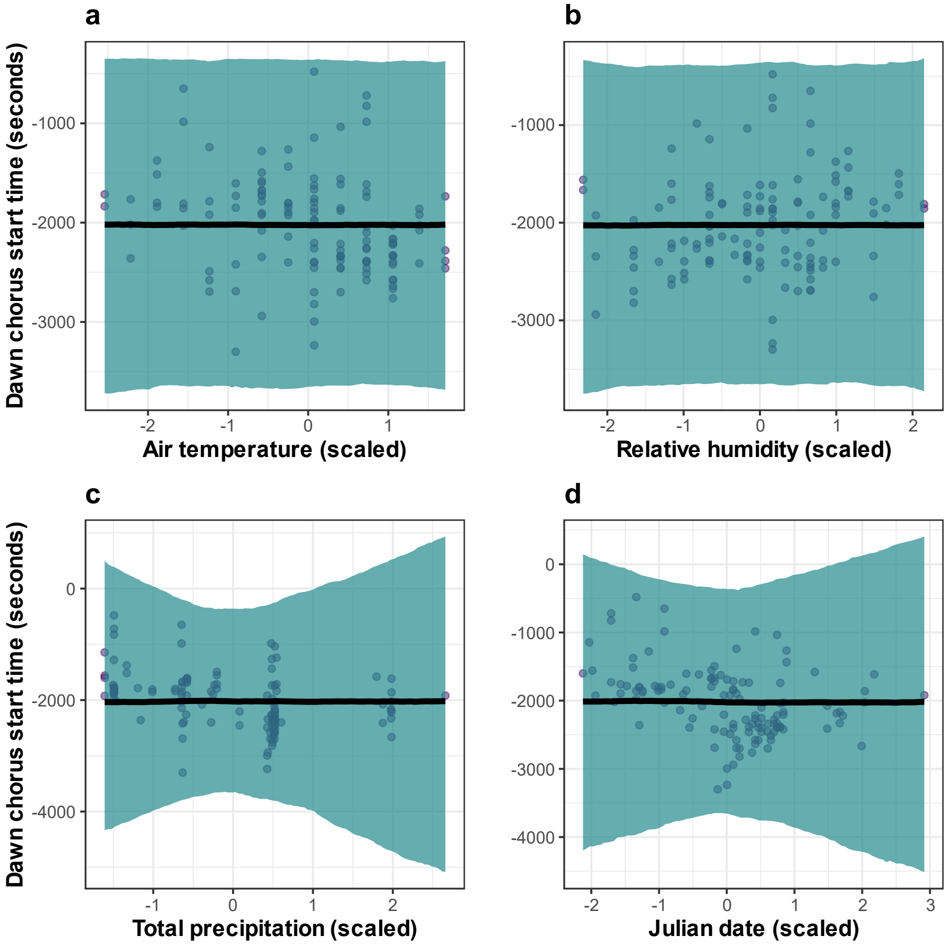
Appendix 2. *Plot comparing the measures for daily average air temperature (degrees Celsius) from the NOAA weather station (turquoise line) and MesoWest weather station (purple line) in (a) 2012, (b) 2017, and (c) 2018.*

APPENDIX 3: Weakly informative priors in the song performance analysis



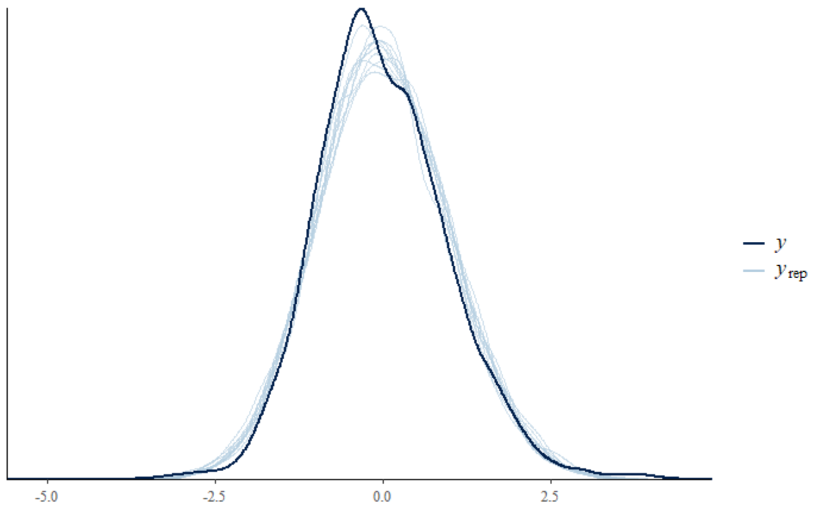
Appendix 3. *Model prior predictions for the effects of each predictor variable (a) air temperature, (b) relative humidity, (c) song order, (d) time relative to sunrise, and (e) Julian date on recovery time. The black line is the model estimate curve, the turquoise shade indicates the 95% confidence intervals, and the purple dots represent datapoints. The other two performance variables and the acoustic traits that underlie their trade-offs produced similar patterns. None of the predictor variables influences the response variable in the priors.*

APPENDIX 4: Weakly informative priors in the dawn chorus start time analysis



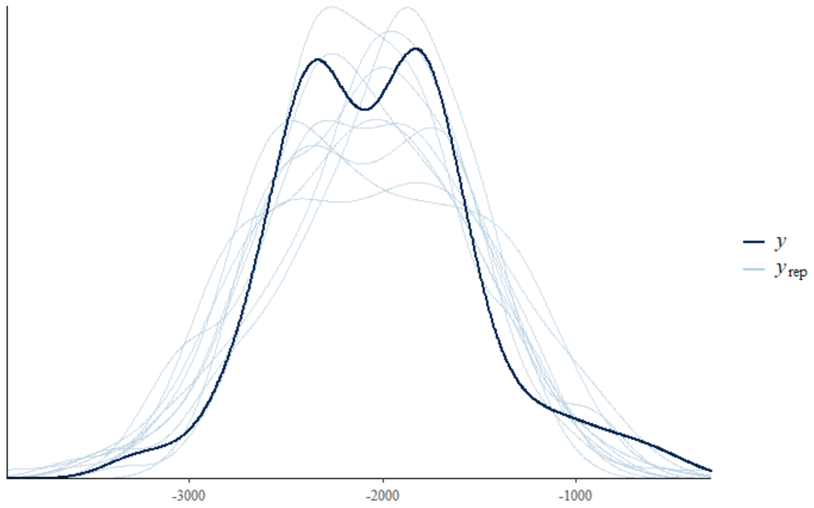
Appendix 4. *Model prior predictions for the effects of (a) air temperature, (b) relative humidity, (c) precipitation accumulation, and (d) Julian date on the dawn chorus start time. The black line is the model estimate curve while the turquoise shade indicates the 95% confidence intervals. The purple dots represent the first song in each recording session.*

APPENDIX 5: Posterior predictive checks in the song performance analysis



Appendix 5. *Example of a posterior predictive check. This is the output from the model of recovery time. The light blue lines (Yrep) represent generated samples of posterior predictive distribution and the dark blue line represents the distribution of the observed data (Y).*

APPENDIX 6: Posterior predictive checks in the dawn chorus start time analysis



Appendix 6. *Posterior predictive check in the dawn chorus start time analysis. The light blue lines (Yrep) represent generated samples of the posterior predictive distribution and the dark blue line represents the distribution of the observed data (Y).*