New horizons for female birdsong: evolution, culture and analysis tools

A thesis presented in
partial fulfilment of the
requirements for the degree of
Doctor of Philosophy in Ecology
at Massey University, Auckland,
New Zealand

Wesley Webb

2020

“The bellbird may be said
to sing matins and vespers for the
warblers of the woods, as it is at the grey break of
dawn, and in the still hour that closes in the day, that
its chimes strike clearest on the ear… a few
individuals meet on a tree or shrub that offers, it
may be, a tempting show of honey-bearing blossoms,
a note or two is briefly sounded, the numbers rapidly
increase, after much noisy fluttering of wings a gush
of clanging melody bursts forth from a score of
quivering throats, forming a concert of
inharmonious, yet most pleasing sounds. Nor is the
ear alone gratified, for the actions and postures of
the melodists are extremely quaint and
droll during this performance…”

Johannes Carl Andersen, 1926

Bird-song and New Zealand song birds.

Auckland, NZ: Whitcombe & Tombs.

Thesis abstract

As a result of male-centric, northern-hemisphere-biased sexual selection theory, elaborate female traits in songbirds have been largely overlooked as unusual or non-functional by-products of male evolution. However, recent research has revealed that female song is present in *most* surveyed songbirds and was in fact the ancestral condition to the clade. Additionally, a high proportion of songbird species have colourful females, and both song and showy colours have demonstrated female-specific functions in a growing number of species. We have much to learn about the evolution and functions of elaborate female traits in general, and female song in particular. This thesis extends the horizons of female birdsong research in three ways: (1) by revealing the broad-scale evolutionary relationship of female song and plumage elaboration across the songbirds, (2) by developing new accessible tools for the measurement and analysis of song complexity, and (3) by showing—through a detailed field study on a large natural metapopulation—how vocal culture operates differentially in males and females.

First, to understand the drivers of elaborate female traits, I tested the evolutionary relationship between female song presence and plumage colouration across the songbirds. I found strong support for a positive evolutionary correlation between traits, with female song more prevalent amongst species with elaborated female plumage. These results suggest that contrary to the idea of trade-off between showy traits, female plumage colouration and female song likely evolved together under similar selection pressures and that their respective functions are reinforcing.

Second, I introduce new bioacoustics software, *Koe*, designed to meet the need for detailed classification and analysis of song complexity. The program enables visualisation, segmentation, rapid classification and analysis of song structure. I demonstrate *Koe* with a case study of New Zealand bellbird *Anthornis melanura* song, showcasing the capabilities for large-scale bioacoustics research and its application to female song.

Third, I conducted one of the first detailed field-based analyses of female song culture, studying an archipelago metapopulation of New Zealand bellbirds. Comparing between male and female sectors of each population, I found equal syllable diversity, largely separate repertoires, and contrasting patterns of sharing between sites—revealing female dialects and pronounced sex differences in cultural evolution.

By combining broad-scale evolutionary approaches, novel song analysis tools, and a detailed field study, this thesis demonstrates that female song can be as much an elaborate signal as male song. I describe how future work can build on these findings to expand understanding of elaborate female traits.

Acknowledgements

After a conjoint BSc/BMus (biology and music composition) at the University of Auckland, I had little idea how to combine these disparate interests. And as I stood printing my Honours thesis on chaetognath reproduction, my future was still a mystery—until a fateful conversation with Dr Clive Evans there at the printer. Clive alerted me to a remarkable opportunity to combine my love of biology and music: a PhD on birdsong diversity at Massey University!

And so it was that I met my supervisor, Prof Dianne Brunton, and embarked on the greatest adventure of my life thus far. What a privilege it has been. Thank you, Dianne, for choosing me for this project, for believing in me and investing the time to guide me despite a formidable workload as Head of School. Not only are you an excellent supervisor, you care for your students on a personal level. Thank you for getting me to conferences and workshops, for extending my skillset and bringing employment opportunities to my attention. I am grateful for all you've done for me.

When I'm an eccentric old man, I intend to regale wide-eyed children with tales of fieldwork; how we paddled with dolphins and bunked with Dauvacel geckos; how I was bitten by a Buller's shearwater and chased down the track by a feisty tuatara (www.youtube.com/watch?v=KZnLxRkEhzw). Glorious dawn choruses, cliff-scrambling, al fresco dining, boat landings in rough seas—it's been high adventure, and I've relished it. Thank you to all who were involved in the field work, especially Dianne Brunton, Michelle Roper, Barbara Evans, Aaron Harmer, Jessica Patiño Pérez, Mehrnaz Tavasoli, and William Pearman. Thank you to the Department of Conservation and iwi who granted permits for working in these special places.

A major component of the PhD has been computer work, and here too I'm privileged to have had excellent, helpful colleagues. To my 'personal computer scientist' and good friend Yukio Fukuzawa—a hearty thanks for all the time you've spent developing *Koe* bioacoustics software; our camaraderie working on this tool together has been a real highlight. You are an immensely talented (and patient) coder, and by demanding payment in Zumba attendance you have introduced me to things I would never have tried otherwise… Thanks also to all who have given R help, stats advice, and other computer tips, particularly my co-supervisors Prof James Dale, Dr Aaron Harmer, Dr Daniel Thomas, and Dr Matthew Pawley, as well as Dr Dave Aguirre, who was always willing to lend a statistical ear, Dr Daniel Playne, who helped with some coding, and Dr Robert Lachlan, who hosted me at Queen Mary College in London and imparted much wisdom on animal culture.

Throughout the project I have worked closely with Michelle Roper (now Dr Michelle Roper), whose thesis work has made my own project possible. My PhD would not have been anywhere near as rewarding without our shared bird-nerding, scientific discussions and conference trips. You are an excellent friend: good-humoured, dependable, supportive, patient, unflappable, and a jolly good birder and scientist. I am grateful!

My colleagues in Ecology and Conservation have done an exceptional job of building community. I commend you all for a highly enjoyable work environment full of kindness and collegiality. I can't name you all, but there are some I would particularly like to thank. Professor Weihong Ji—thank you for the music jam sessions, course tutoring opportunities and beautiful Christmas cards. Jenny Laycock, Luca Butikofer, and Sarah Scott—thank you for the exciting volunteering opportunities (frogs, skinks and geckos) which led to some great stories. Jacques de Satge—my dear chap, thanks for the pun fests and manly emotional support. Laureline Maynier—I have enjoyed our witty banter. Kevin Parker—thank you for the sage career advice. Anne Wignall, Marleen Baling, Mehrnaz Tavasoli, Jessica Patiño Pérez, Cheryl Cross—it's been a pleasure working alongside you. Annette Warbrooke—thanks for your fantastic admin help.

And of course, to my family and friends who supported me throughout—I really could not have done this without you. My deep thanks to my parents, Howard and Lynette, for their love, wisdom and perspective. Thank you, Dylan, my genius and kind-hearted brother, for your comic relief and help with *Excel* macros which were key to my song-cataloguing endeavours. And to my sister Catherine, who has embarked on a PhD herself—it is great fun nerding with you, and I am watching expectantly to see how you are going to change the world (no pressure). Dr Zachary Ardern, Dr Greg Liston, Giorgi Meladze, Steve Worsley, and many others who I do not have space to name here: your friendship and interest has meant a great deal.

This study was funded by a Royal Society of New Zealand Marsden Fund grant (13-MAU-004) to my primary supervisor Professor Dianne Brunton. Contributions to conference travel were made by a Claude McCarthy Travel Fellowship and an Association for the Study of Animal Behaviour Student Award.

Permits and ethics

Bellbird song recording, mist-netting and blood-sampling at my six field sites (Hauturu, Tawharanui, Tiritiri Matangi, Lady Alice, Tawhiti Rahi, and Repunga) was undertaken with permission from Ngati Wai and Ngati Manuhiri iwi, the New Zealand Department of Conservation (47948-FAU, 34833–FAU, 41756-FAU, 48000-FAU), and Massey University Animal Ethics Committee (15/21).

Table of contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendices

List of Tables

Chapter 3

Chapter 4

Chapter 5

Appendices

1 Introduction

Birdsong is one of the best-studied areas in behavioural ecology

Acoustic communication is a fundamental aspect of behaviour for many species across the animal kingdom, used in contexts such as territorial defence, mate attraction, group coordination, and recognition of individuals and species (Peters et al., 1980; Catchpole and Slater, 2008; Soltis, 2009; Janik and Sayigh, 2013; Starnberger et al., 2014; Sperber et al., 2017). But the vocalisations of songbirds (oscine passerines)—particularly those vocalisations we call 'songs'—are some of the most complex and well-studied (Marler and Slabbekoorn, 2004; Birkhead, 2008; Rothenberg et al., 2014).

The advent of the sound spectrogram in the 1950s elevated bird song research to a science (reviewed in Marler, 2004); what previously could only be listened to and described in words or musical notation could now be visualised and measured. This new measurability, together with the appealing diversity, ubiquity and tractability of bird song, launched the field into what is now one of the best-studied areas of behavioural ecology (Catchpole and Slater, 2008). For example, laboratory studies have mapped song-learning neural pathways of the brain and the influences of hormones on singing behaviour (e.g. Nottebohm, 1977, 1980). Endoscopic imaging of singing birds and studies utilizing electrodes in syringeal and respiratory musculature have greatly progressed our understanding of syrinx function (e.g.

Goller and Suthers, 1996; Larsen and Goller, 1999). Song-tutoring and cross-fostering studies—where eggs or chicks from two nests are swapped so that they are raised by non-biological parents—have shed light on the ontogeny of song development and learning modes (e.g. Slagsvold et al., 2002; Holveck and Riebel, 2014; Riebel et al., 2015). Playback studies have helped to decipher the role of male song in species recognition, mate choice, and resource defence (e.g. Slater, 2003; Brunton et al., 2008a, 2008b).

Such research revealed that songbird songs are learnt through vocal imitation of other individuals, rather than innate (Kroodsma, 1982). Vocal learning is a sparsely documented trait in the animal kingdom, known only in humans, cetaceans, bats, elephants, seals, songbirds, hummingbirds, and parrots (Paton et al., 1981; Baptista and Schuchmann, 1990; Janik and Slater, 1997; Poole et al., 2005; Sanvito et al., 2007; Catchpole and Slater, 2008). While the sub-oscines (Passeriformes, suborder Tyranni) can have elaborate vocalisations, these vocalisations are generally considered stereotyped and non-learnt, developing without reference to any auditory model (Nottebohm, 1972; Baptista and Kroodsma, 2001; but see Kroodsma et al., 2013; Liu et al., 2013). By contrast, the songs of oscines are a clear example of animal *culture* (reviewed by Aplin, 2019).

Birdsong is culturally transmitted, leading to geographic variation

Culture is shared information or behaviour acquired through social learning from conspecifics (Dawkins, 1976; Aplin, 2019). It involves the transmission of cultural units—termed *memes* (Dawkins, 1976)—through populations. A meme is essentially an idea or concept, such as a skill, melody, image, gesture, word or vocalisation, that replicates through imitation by other individuals (McNamara, 2011).

In the context of vocally learnt birdsong, syllable types, phrase types or song types may be considered acoustic memes. A meme, Dawkins argues, has many similarities to the gene (transmissible unit of genetic information) in that both are replicated (i.e. have fecundity), can spread through a population, both are subject to mutation and are operated on by natural selection. Thus, the three prerequisites for a Darwinian evolutionary system—replication, variance and selection—are met.

However, there are several notable differences between genes and memes (Claidière and André, 2011; Oren et al., 2018). Firstly, while the genetic transmission mode is highly stable and independent of the content transmitted, cultural transmission modes are highly unstable and variable, potentially depending on meme content. To illustrate, genetic transmission in animals is vertical (parent to offspring), involving bi-parental transmission with meiosis for chromosomal DNA, regardless of fitness conferred to the receiver. By contrast, cultural transmission can theoretically take place between any individuals; it may be vertical, horizontal (between individuals of the same age class) or oblique (different age classes) (Payne et al., 1988; Lynch et al., 1989; Baker and Gammon, 2008). Memes are also far more prone to copy errors, and are subject to deliberate modifications and learning biases (Shennan, 2011). Receivers may choose to adopt a meme or reject it depending on the fitness benefits it confers (Laland, 2004). Because a behavioural innovation may quickly spread through an entire population, memes frequently have much higher fecundity than genes (Whiten et al., 2011).

Thus, cultural evolution is complex, and can be a rapid and powerful means of adaptation to environments, sometimes more so than genetic evolution (Perreault, 2012). It has been proposed as a main reason for the global dominance of the human species (Richerson et al., 2010) and as a key factor in speciation through behavioural isolation. For example, in birds, the cultural evolution of local song dialects and mating preferences for the local dialect can form mating barriers leading to speciation (Wilkins et al., 2013). This dynamic is thought to have been a driving force in the diversification of the passerines (Podos and Warren, 2007). Because cultural evolution can profoundly influence the dynamics of population change and survival through time, understanding the dynamics of memes is of great importance for ecology and evolutionary biology (Parker et al., 2012).

One interesting outcome of learning songs by cultural transmission is the phenomenon of geographic song variation. The song profile of a songbird population changes as new memes are introduced through immigration, copying errors, and innovation, while other memes fall out of use. This leads to a diversification of the meme pool across the species, much like the diversification of human language. There are three types of geographic variation in song, summarised in Figure 1.1: dialects, clinal variation, and random variation (Podos and Warren, 2007). Of these, dialects are by far the best-studied in songbirds.

Dialects are characterised by uniformity of signal forms *within* given localities, with sharp geographic boundaries and discrete differences in signal forms *between* localities (Podos and Warren, 2007). Geographic dialects have been documented not only in songbirds (Jenkins, 1978; Aplin, 2019) but also in other vocally learning taxa (Cetaceans: Winn et al., 1981; Janik and Sayigh, 2013; Filatova et al., 2015. Bats:

Esser and Schubert, 1998; Prat et al., 2017. Parrots: Wright and Dahlin, 2018. Hummingbirds: Araya-Salas and Wright, 2013; Kapoor, 2016). While some non-learning avian taxa show gradual, clinal variation in temporal or spectral parameters across their ranges (Baptista et al., 1983; Isler et al., 2005), only learning appears capable of producing dialects.

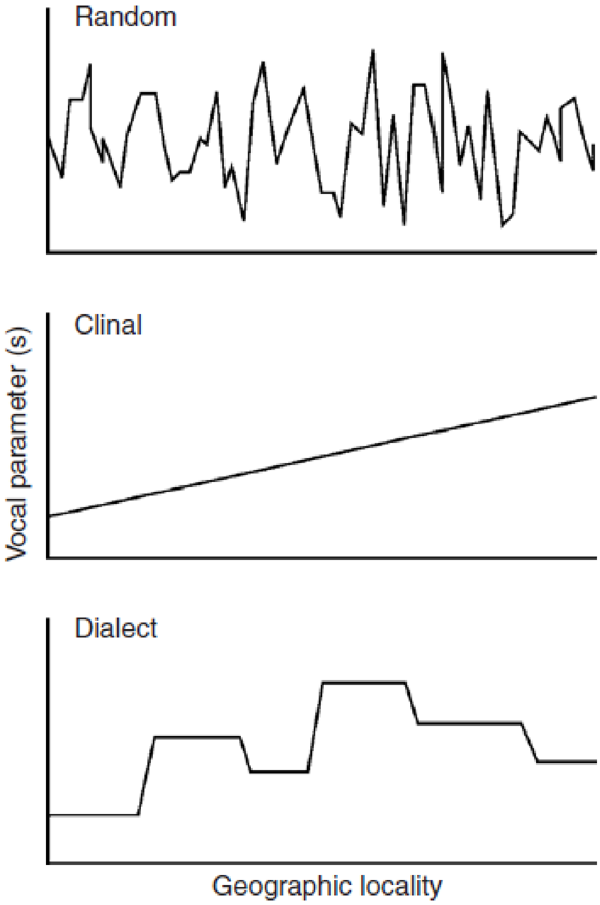


Figure 1.1—Schematic representation of three patterns of geographic variation in song.

Geographic variation can be manifested at varying scales. Dialect variation features sharp transitions in vocal parameters between localities, and consistency in vocal parameters within localities (from Podos and Warren, 2007). Other forms of variation include 'clinal' variation, in which vocal parameters vary continuously rather than discretely from one place to another, and random variation, where there is no discernible pattern to vocal parameter changes over distance.

Three hypotheses dominate the literature regarding why song dialects form: the local adaptation hypothesis, the social adaptation hypothesis, and the epiphenomenon hypothesis.

The **local adaptation hypothesis** (Marler and Tamura, 1964) centres on sexual selection by females. Females gain fitness by selecting males from their natal area, because these males will generally be better adapted to the local conditions than foreign males and will produce offspring better adapted to the local conditions. Assuming song is an accurate indicator of natal locality, females will choose males that sing the local song, resulting in distinct geographical dialects. This hypothesis requires that males learn songs *pre-dispersal*, and do not subsequently modify their repertoire, for song to be an accurate indicator of natal region.

By contrast, the **social adaptation hypothesis** focuses on the role of aggressive social interactions in the formation of local dialects. This hypothesis has several variants. One version is the 'deceptive mimicry hypothesis' (Payne, 1981) which envisages subordinate males mimicking dominant males, to appear more dominant and thus reduce probabilities of aggressive encounters with other males. This would predict temporally unstable dialects, with the prevalence of acoustic memes depending on which males were dominant at the time. Other versions include the 'honest convergence' (Rothstein and Fleischer, 1987) and 'colony password' (Feekes, 1977) hypotheses, which posit that local song types provide accurate indicators of long-term residence. The longer a male inhabits a neighbourhood, the better he will be at producing local song types. Foreign males (that sing non-local songs) suffer social penalties such as increased aggression, because they represent more of a threat (Hardouin et al., 2006), being 'inherently expansionist'. Unlike the local adaptation hypothesis, social adaptation hypotheses require *post-dispersal* song learning in order for males to adopt songs of their neighbourhood.

A third option, the **epiphenomenon hypothesis**, proposes that dialects arise not for any functional reason, but rather as a spurious by-product of certain song-learning and dispersal modes. For example, populations with limited dispersal may diverge in song parameters by chance processes such as cultural drift, or as a result of acoustic adaptation to differing environmental conditions (Wilkins et al., 2013; Potvin and Clegg, 2015). So even without social or sexual selection for males to sound like other males of the area *per se*, if song is learnt and dispersal after learning is restricted, "some sort of variation in both time and space seems inevitable" (Slater, 1986).

There does not appear to be a single, universal explanation for how dialects form, which is perhaps unsurprising considering the variety of habitats, social structures, song learning and dispersal modes in vocally-learning birds (Podos and Warren, 2007). The epiphenomenon hypothesis may have greater explanatory power than the two functional hypotheses, neither of which can account for the majority of examples of dialects in the literature (Podos and Warren, 2007).

How is dialect stability maintained?

Why is it that white-crowned sparrow *Zonotrichia leucophrys* dialects can remain stable acoustically and geographically for over 40 years (Nelson et al., 2004), while yellow-rumped cacique *Cacicus cela* dialects change from year to year (Trainer, 1989)? The processes promoting dialect stability are far from clear (Salinas-Melgoza and Wright, 2012; Planqué et al., 2013). Typically, the hypotheses discussed above regarding the origin of dialects are assumed to apply to dialect maintenance, too. For example, dispersal across dialect boundaries may be limited due to fitness costs of singing foreign song (aggression and/or low mating success), thus preventing mingling of the two song cultures and maintaining the distinction. Or, in species exhibiting dispersal across dialect boundaries, dialect stability may be maintained by post-dispersal imitation of the local songs (Podos and Warren, 2007).

But current evidence does not conclusively support either scenario as generalisable explanations (Salinas-Melgoza and Wright, 2012). Rather, simulation modelling (Planqué et al., 2013) suggests that song learning characteristics, individual dispersal behaviour, assortative mating, and environmental selection have complex interacting effects on dialect stability that defy simple explanation. For instance, whether the frequency-dependence of song learning is linear or non-linear can have a dramatic effect on dialect stability, as can the fitness costs/benefits of rare memes (Laland, 2004; Planqué et al., 2013). There is apparently no relationship between migratory behaviour and stability of dialects (Podos and Warren, 2007), though territorial species are thought to maintain dialects for longer than non-territorial colonial species (Trainer, 1989). Acoustic competition from other species or anthropogenic noise (Slabbekoorn and Ripmeester, 2008; Azar et al., 2014; Potvin and Clegg, 2015) can force dialect parameter shifts over time. So the factors involved in dialect stability are complex and currently difficult to separate.

Female song and other elaborate female traits: a major knowledge gap

As envisaged by Darwin (1871), reproduction is usually more costly for females, and so it pays females to be choosy, favouring high-quality males as mates.

Consequently, it is mostly males who must compete for access to females—the sex with limited availability. Sexual selection has driven the evolution of myriad elaborate male traits for use in male–male competition and as indicators of quality for female choice. Amongst birds, such traits can include elaborate courtship behaviours (Ligon et al., 2018), colourful plumage (Dale et al., 2015), and song (Oberweger and Goller, 2001).

The sexual selection theorem has strong support and is a cornerstone of modern theory. However, the emphasis on male competition and female choosiness has resulted in a longstanding assumption that elaborate signals such as song and colourful plumage must be predominantly male traits. Cases of female trait elaboration have traditionally been considered non-adaptive aberrations that might result 'accidentally' from shared genetic architecture with males (Darwin, 1871; Lande, 1980; see Tobias et al., 2012 for review). This view was maintained in part because most ornithological studies have been conducted in the northern hemisphere, where elaborate female traits tend to be rare. Thus, studies of elaborate bird traits have, until recently, focused almost entirely on males.

But recent work in birds has shown that elaborate female traits are not mere genetic carryovers from males. Far from being the exception, female song is widespread, occurring in 64% of surveyed oscine passerine species (Figure 1.2), especially concentrated in the tropics and southern hemisphere, and has been recovered as the ancestral condition of the songbirds (Langmore, 1998; Slater and Mann, 2004; Odom et al., 2014; Webb et al., 2016). Female song is also evolutionarily labile with respect to males (Hall et al., 2015; Price, 2015) as losses of female song have occurred independently of males (Price, 2019). Independent evolution of female song strongly suggests this trait is not functionless but is under selection. This is confirmed by direct evidence of female-specific functions of song in a growing number of species. For instance, female song functions in territory and resource defence, identity signalling, mate attraction, mate defence, pair-bonding, and links to female reproductive success (Searcy and Yasukawa, 1995; Langmore et al., 1996; Rogers et al., 2007; Brunton et al., 2008b, 2016; Templeton et al., 2011; Hall et al., 2015; Cain and Langmore, 2016; Keen et al., 2016; Demko and Mennill, 2018; Reichard et al., 2018; Magoolagan et al., 2019; Rose et al., 2019).

So there is compelling evidence that female song is indeed adaptive, and growing recognition of the female song research gap has elevated the subject to a 'hot topic' in behavioural biology (Hall et al., 2015; Odom and Benedict, 2018).

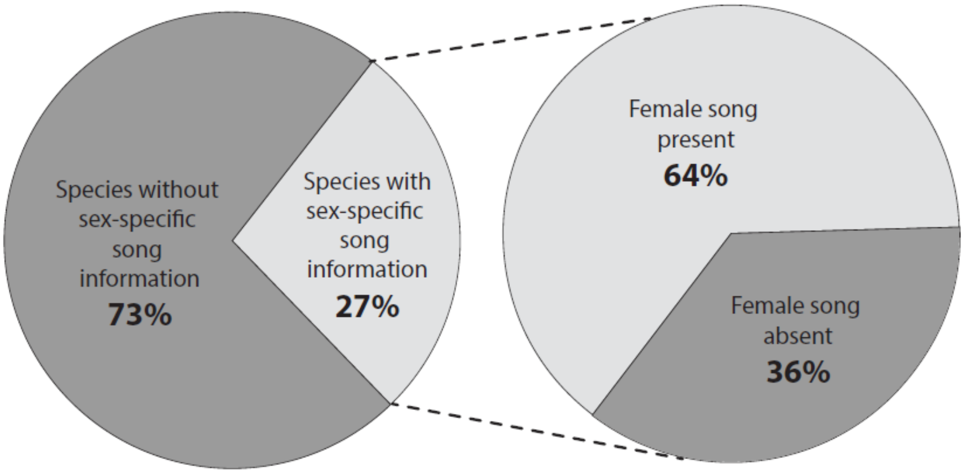


Figure 1.2—The current coverage of female song information (left) and known prevalence of female song (right).

Sex-specific information about song presence exists for only 27% of songbird species (Passeri). Of these, 64% have female song. Values from Webb et al. (2016) and figure concept adapted from Odom and Benedict (2018).

Female plumage elaboration is likewise common, apparently functional and evolutionarily labile with respect to the male phenotype (Price, 2019). While drab female plumage may appear the norm in northern temperate latitudes, elaborate and sexually monomorphic plumage elaboration is common in the tropics, where most songbird species occur (Dale et al., 2015; Odom and Benedict, 2018). Female plumage changes have played a greater role than male changes in the evolution of dichromatism in New World blackbirds (Icteridae; Irwin, 1994; Hofmann et al., 2008; Price and Eaton, 2014), tanagers (Thraupidae; Burns, 1998) and fairy-wrens (Maluridae; Johnson et al., 2013) and this pattern holds true for the order Passeriformes as a whole (Dale et al., 2015)—demonstrating evolutionary independence from the male phenotype (Shultz and Burns, 2017; Price, 2019). As for functional significance, female colouration is attractive to mates in some species (Amundsen et al., 1997; Smiseth and Amundsen, 2000; Murphy et al., 2009a, 2009b; Hasegawa et al., 2017) and frequently has roles in territory defence and female–female competition for non-sexual breeding resources (Doutrelant et al., 2008; Tobias et al., 2012; Morales et al., 2014; Berzins and Dawson, 2018; Enbody et al., 2018).

Thus, after being long overlooked as non-functional and aberrant, in fact female song and female plumage elaboration appear both adaptive and normative, highlighting elaborate female traits as an important knowledge gap.

Filling in the gaps: evolution of elaborate female traits, and comparing female song cultures to males

Understanding the evolutionary drivers of elaborate female traits is one key component of filling in the female song knowledge gap. This includes examining the evolutionary relationships *between* elaborate traits—for example, the evolutionary relationship between female song and female plumage elaboration. Do these traits evolve in a concerted fashion, in an antagonistic fashion, or in ways unrelated to each other? The answer could yield insights into whether the proposed high cost of song and elaborate plumage might impose an evolutionary trade-off across the songbirds (Darwin, 1871), and whether the traits have overlapping or separate functions.

Another important knowledge gap is female song culture. While female song dialects have been documented in bats (Esser and Schubert, 1998) and cetaceans (Whitehead and Rendell, 2014), little is known about cultural evolution of female song in songbirds (Mennill and Rogers, 2006; Graham et al., 2017, 2018a, 2018b). Do female song dialects exist, and if so, how do they arise? Are the processes underlying male and female cultural evolution different—perhaps dependent on sex-biased dispersal or song-learning modes?

New bioacoustics software is needed for large-scale analysis

Broad-scale evolutionary studies, and studies of female culture, are limited by the paucity of data on female song. Only 27% of songbird species have sex-specific song information available (Figure 1.2), and in the vast majority of these, female song has only been described anecdotally, not in a quantitative way. To advance understanding of female song, species-level studies of female song structure and complexity are vital to enable detailed comparisons between sexes and across species (Odom and Benedict, 2018).

Like many animal vocalisations, bird songs tend to be structured as a sequence of distinct acoustic units, where information is encoded in the types of units and sometimes their sequence structure (i.e. temporal arrangement) (Kershenbaum et al., 2016). To be able to analyse structure, complexity, and other metrics of song elaboration, a researcher must segment song recordings into constituent acoustic units (e.g. 'syllables', 'motifs' or suchlike), and classify acoustic units according to acoustic similarity (Figure 1.3).

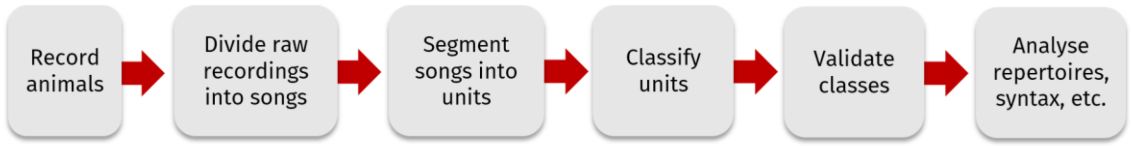


Figure 1.3—A typical process for studies in bioacoustics.

Classifying units involves identifying unit classes (i.e. producing a ‘catalogue' of class types) and then assigning units to those classes—only once units have been classified can repertoires and sequence structures be analysed and compared. Classification is a challenge, however; thus far, attempts at automation have not proven generalisable (Priyadarshani et al., 2018) and the primary approach for most species remains manual classification based on human visual and auditory perception (Kershenbaum et al., 2016).

To discern and assign classes manually, a researcher needs to view spectrograms and hear playback of units, comparing similarity of many units at a time. I know of no tools designed for this purpose; existing acoustics software offers visualisation and labelling functionality only within one or few files at a time, not on the large comparative scale required. As a result, researchers currently resort to cumbersome and non-integrated workflows that severely limit the scope of bioacoustics research. In an age where digital recording, fast computers, computer-based spectrograms and passive recorders are gathering large amounts of song data, new tools to expedite the large-scale classification and analysis of song could be of immense benefit. And if such tools were intuitive, accessible and free, they could democratise the wide-scale gathering and analysis of female song data urged by Odom and Benedict (2018).

New Zealand bellbirds: a model for comparative study of male and female song culture

I have identified a northern metapopulation of New Zealand bellbirds *Anthornis melanura* as an ideal system for studying cultural evolution of female song in comparison to males. Bellbirds are sexually dimorphic in size, plumage, and song, and juveniles are easily distinguished from adults (Roper, 2018). Both sexes of bellbird sing frequent and complex songs throughout the year in the defence of territories and food (Brunton and Li, 2006). Females disperse more frequently than males, and genetic analyses support an isolation-by-distance model of gene flow, with geographic distance strongly predicting genetic distance between populations (i.e. gene flow is limited by dispersal; Baillie et al., 2014).

Although extirpated from the mainland of the northern North Island (north of Hamilton; Craig and Douglas, 1984), bellbirds persist in a network of offshore islands spanning the Hauraki Gulf, including some peninsulas and mainland coastal forests recently colonised from island populations (Brunton et al., 2008a; Baillie et al., 2014). The sites in this archipelago differ in size, bellbird population histories, and connectivity to each other, providing an excellent opportunity to explore the influence of various spatial and social parameters governing song culture. While studies on male songbird culture have been conducted at a continental scale (e.g. yellowhammers *Emberiza citronella* across Europe; Diblíková et al., 2019) and localised scale (e.g. North Island saddleback *Philesturnus rufusater* within islands; Jenkins, 1978), an archipelago of this intermediate scale provides a tractable system but with the interesting complexities of dispersal processes. Furthermore, because males and females share genetic and environmental backgrounds yet differ in dispersal behaviour, the system makes it theoretically possible to isolate the effect of dispersal frequency on meme dynamics.

Thesis aims and outline

This thesis aims to address the female song knowledge gap in three ways:

(1) by elucidating the evolutionary relationship between two elaborate female traits: song and colourful plumage,

(2) by introducing new open-source software tools for analysis and comparison of female song structure/complexity,

(3) by applying those tools in the first large-scale study of female song culture in a natural system.

In **Chapter 2**, I study the evolutionary relationship between female song presence and female plumage colour elaboration across the songbirds (aim 1). I test whether these two elaborate traits have evolved in a concerted fashion, in an antagonistic fashion, or in ways unrelated to each other, to shed light on their respective functions and evolutionary drivers. This chapter is published in the peer-reviewed journal *Frontiers in Ecology and Evolution* and has been adapted for presentation in this thesis.

In **Chapter 3**, I describe the rationale and methods by which I collected and processed song recordings from a metapopulation of New Zealand bellbirds in the Hauraki Gulf. I explain how I produced a database of recordings—partitioned into songs, which in turn were segmented into their constituent syllables, which were finally classified according to acoustic similarity. The resultant database forms the basis of the analyses in Chapter 4, Chapter 5, and Appendices.

In **Chapter 4**, I introduce *Koe*, a tool designed to meet the need for detailed classification and analysis of song complexity. The program enables visualisation, segmentation, rapid classification and analysis of song structure. I demonstrate *Koe* with a case study of New Zealand bellbird song, showcasing the capabilities for large-scale bioacoustics research and application to female song (aim 2). This chapter is published as an application paper in the peer-reviewed journal *Methods in Ecology and Evolution* and has been adapted for presentation in this thesis.

In **Chapter 5**, I conduct one of the first large-scale comparative analyses of male and female song cultures (aim 3), using a metapopulation of New Zealand bellbirds spanning an archipelago. I compare male and female meme repertoires, meme prevalence, and patterns of meme sharing within and between sites in the island network to elucidate sexual differences in culture.

Finally, in **Chapter 6**, I draw out the key findings of my thesis and discuss their significance in the wider context. I describe how future work can build on this knowledge to expand the horizons of female song research.

Author contributions are indicated in each chapter.

References

Amundsen, T., Forsgren, E., and Hansen, L. T. T. (1997). On the function of female ornaments: male bluethroats prefer colourful females. *Proc. Biol. Sci.* 264, 1579–1586.

Aplin, L. M. (2019). Culture and cultural evolution in birds: a review of the evidence. *Anim. Behav.* 147, 179–187. doi:10.1016/j.anbehav.2018.05.001.

Araya-Salas, M., and Wright, T. (2013). Open-ended song learning in a hummingbird. *Biol. Lett.* 9, 20130625. doi:10.1098/rsbl.2013.0625.

Azar, J. F., Bell, B. D., and Borowiec, M. (2014). Temporal change of the song of a local population of the grey warbler (*Gerygone igata*): has its song changed over time? *Emu* 114, 80–85.

Baillie, S. M., Ritchie, P. A., and Brunton, D. H. (2014). Population genetic connectivity of an endemic New Zealand passerine after large-scale local extirpations: a model of re-colonization potential. *Ibis* 156, 826–839. doi:10.1111/ibi.12182.

Baker, M. C., and Gammon, D. E. (2008). Vocal memes in natural populations of chickadees: why do some memes persist and others go extinct? *Anim. Behav.* 75, 279–289. doi:10.1016/j.anbehav.2007.05.022.

Baptista, L. E., and Kroodsma, D. E. (2001). "Avian bioacoustics, a tribute to Luis Eelipe Baptista.," in *Handbook of birds of the world*, eds. J. Del Hoyo and A. Elliott (Barcelona, Spain: Lynx Edicions), 11–52.

Baptista, L. F., Boarman, W. I., and Kandianidis, P. (1983). Behavior and taxonomic status of Grayson's Dove. *The Auk*, 907–919.

Baptista, L. F., and Schuchmann, K.-L. (1990). Song learning in the Anna hummingbird (*Calypte anna*). *Ethology* 84, 15–26.

Berzins, L. L., and Dawson, R. D. (2018). Experimentally altered plumage brightness of female Tree Swallows (Tachycineta bicolor) influences nest site retention and reproductive success. *Can. J. Zool.* 96, 600–607. doi:10.1139/cjz-2017-0156.

Birkhead, T. (2008). *The Wisdom of Birds: An Illustrated History of Ornithology*. First Edition edition. New York: Walker Books.

Brunton, D. H., Evans, B. A., and Ji, W. (2008a). Assessing natural dispersal of New Zealand bellbirds using song type and song playbacks. *N. Z. J. Ecol.* 32, 147–154.

Brunton, D. H., Evans, B., Cope, T., and Ji, W. (2008b). A test of the dear enemy hypothesis in female New Zealand bellbirds (*Anthornis melanura*): female neighbors as threats. *Behav. Ecol.* 19, 791–798. doi:10.1093/beheco/arn027.

Brunton, D. H., and Li, X. (2006). The song structure and seasonal patterns of vocal behavior of male and female bellbirds (*Anthornis melanura*). *J. Ethol.* 24, 17–25. doi:10.1007/s10164-005-0155-5.

Brunton, D. H., Roper, M. M., and Harmer, A. M. (2016). Female song rate and structure predict reproductive success in a socially monogamous bird. *Front. Ecol. Evol.* 4, 13.

Burns, K. J. (1998). A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. *Evolution* 52, 1219–1224. doi:10.2307/2411252.

Cain, K. E., and Langmore, N. E. (2016). Female song and aggression show contrasting relationships to reproductive success when habitat quality differs. *Behav. Ecol. Sociobiol.* 70, 1867–1877.

Catchpole, C. K., and Slater, P. J. B. (2008). *Bird song: Biological themes and variations*. 2nd ed. New York, NY, US: Cambridge University Press doi:10.1017/CBO9780511754791.

Claidière, N., and André, J.-B. (2011). The transmission of genes and culture: a questionable analogy. *Evol. Biol.* 39, 12–24. doi:10.1007/s11692-011-9141-8.

Craig, J. L., and Douglas, M. E. (1984). Bellbirds in Auckland and Northland. *Notornis* 31, 82–86.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015). The effects of life history and sexual selection on male and female plumage colouration. *Nature* 527, 367–370.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Demko, A. D., and Mennill, D. J. (2018). Male and female signaling behavior varies seasonally during territorial interactions in a tropical songbird. *Behav. Ecol. Sociobiol.* 72, 84. doi:10.1007/s00265-018-2495-5.

Diblíková, L., Pipek, P., Petrusek, A., Svoboda, J., Bílková, J., Vermouzek, Z., et al. (2019). Detailed large-scale mapping of geographical variation of yellowhammer *Emberiza citrinella* song dialects in a citizen science project. *Ibis* 161, 401–414. doi:10.1111/ibi.12621.

Doutrelant, C., Grégoire, A., Grnac, N., Gomez, D., Lambrechts, M. M., and Perret, P. (2008). Female coloration indicates female reproductive capacity in blue tits. *J. Evol. Biol.* 21, 226–233. doi:10.1111/j.1420-9101.2007.01451.x.

Enbody, E. D., Boersma, J., Schwabl, H., and Karubian, J. (2018). Female ornamentation is associated with elevated aggression and testosterone in a tropical songbird. *Behav. Ecol.* 29, 1056–1066. doi:10.1093/beheco/ary079.

Esser, K.-H., and Schubert, J. (1998). Vocal dialects in the lesser spear-nosed bat *Phyllostomus discolor*. *Naturwissenschaften* 85, 347–349. doi:10.1007/s001140050513.

Feekes, F. (1977). Colony-specific song in *Cacicus cela* (Icteridae, Aves) - password hypothesis. *Ardea* 65, 197–202.

Filatova, O. A., Samarra, F. I. P., Deecke, V. B., Ford, J. K. B., Miller, P. J. O., and Yurk, H. (2015). Cultural evolution of killer whale calls: background, mechanisms and consequences. *Behaviour* 152, 2001–2038. doi:10.1163/1568539X-00003317.

Goller, F., and Suthers, R. A. (1996). Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.* 75, 867–876.

Graham, B. A., Heath, D. D., and Mennill, D. J. (2017). Dispersal influences genetic and acoustic spatial structure for both males and females in a tropical songbird. *Ecol. Evol.* 7, 10089–10102. doi:10.1002/ece3.3456.

Graham, B. A., Heath, D. D., Walter, R. P., Mark, M. M., and Mennill, D. J. (2018a). Parallel evolutionary forces influence the evolution of male and female songs in a tropical songbird. *J. Evol. Biol.* 31, 979–994. doi:10.1111/jeb.13279.

Graham, B. A., Heath, D. D., Walter, R. P., and Mennill, D. J. (2018b). Immigrant song: males and females learn songs after dispersal in a tropical bird. *Behav. Ecol.* 29, 711–723. doi:10.1093/beheco/ary029.

Hall, M. L., Rittenbach, M. R. D., and Vehrencamp, S. L. (2015). Female song and vocal interactions with males in a neotropical wren. *Behav. Evol. Ecol.* 3, 12. doi:10.3389/fevo.2015.00012.

Hardouin, L. A., Tabel, P., and Bretagnolle, V. (2006). Neighbour–stranger discrimination in the little owl, *Athene noctua*. *Anim. Behav.* 72, 105–112. doi:10.1016/j.anbehav.2005.09.020.

Hasegawa, M., Arai, E., Watanabe, M., and Nakamura, M. (2017). Reproductive advantages of multiple female ornaments in the Asian Barn Swallow Hirundo rustica gutturalis. *J. Ornithol.* 158, 517–532. doi:10.1007/s10336-016-1401-z.

Hofmann, C. M., Cronin, T. W., and Omland, K. E. (2008). Evolution of sexual dichromatism. 1. Convergent losses of elaborate female coloration in New World orioles (*Icterus* spp.). *The Auk* 125, 778–789.

Holveck, M.-J., and Riebel, K. (2014). Female zebra finches learn to prefer more than one song and from more than one tutor. *Anim. Behav.* 88, 125–135. doi:10.1016/j.anbehav.2013.11.023.

Irwin, R. E. (1994). The evolution of plumage dichromatism in the New World blackbirds: social selection on female brightness. *Am. Nat.* 144, 890–907.

Isler, M. L., Isler, P. R., Brumfield, R. T., and Zink, R. M. (2005). Clinal variation in vocalizations of an antbird (Thamnophilidae) and implications for defining species limits. *The Auk* 122, 433–444.

Janik, V. M., and Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A* 199, 479–489. doi:10.1007/s00359-013-0817-7.

Janik, V. M., and Slater, P. J. (1997). Vocal learning in mammals. *Adv. Study Behav.* 26, 59–100.

Jenkins, P. F. (1978). Cultural transmission of song patterns and dialect development in a free-living bird population. *Anim. Behav.* 26, 50–78. doi:10.1016/0003-3472(78)90007-6.

Johnson, A. E., Price, J. J., and Pruett-Jones, S. (2013). Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecol. Evol.* 3, 3030–3046. doi:10.1002/ece3.686.

Kapoor, V. (2016). The functional significance of microgeographic dialects in a hermit hummingbird. Available at: https://ecommons.cornell.edu/handle/1813/43723 [Accessed June 26, 2019].

Keen, S., Meliza, C. D., Pilowsky, J., and Rubenstein, D. R. (2016). Song in a social and sexual context: vocalizations signal identity and rank in both sexes of a cooperative breeder. *Front. Ecol. Evol.* 4, 46.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev.* 91, 13–52. doi:10.1111/brv.12160.

Kroodsma, D. E. (1982). Learning and the ontogeny of sound signals in birds. *Acoust. Commun. Birds* 2, 1–23.

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., et al. (2013). Behavioral evidence for song learning in the suboscine bellbirds (*Procnias* spp.; Cotingidae). *Wilson J. Ornithol.* 125, 1–14.

Laland, K. N. (2004). Social learning strategies. *Anim. Learn. Behav.* 32, 4–14. doi:10.3758/BF03196002.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292–305. doi:10.2307/2407393.

Langmore, N. E. (1998). Functions of duet and solo songs of female birds. *Trends Ecol. Evol.* 13, 136–140. doi:10.1016/S0169-5347(97)01241-X.

Langmore, N. E., Davies, N. B., Hatchwell, B. J., and Hartley, I. R. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 141–146.

Larsen, O. N., and Goller, F. (1999). Role of syringeal vibrations in bird vocalizations. *Proc. R. Soc. Lond. B Biol. Sci.* 266, 1609–1615. doi:10.1098/rspb.1999.0822.

Ligon, R. A., Diaz, C. D., Morano, J. L., Troscianko, J., Stevens, M., Moskeland, A., et al. (2018). Evolution of correlated complexity in the radically different courtship signals of birds-of-paradise. *PLoS Biol.* 16. doi:10.1371/journal.pbio.2006962.

Liu, W., Wada, K., Jarvis, E. D., and Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nat. Commun.* 4, 2082. doi:10.1038/ncomms3082.

Lynch, A., Plunkett, G. M., Baker, A. J., and Jenkins, P. F. (1989). A model of cultural evolution of chaffinch song derived with the meme concept. *Am. Nat.* 133, 634–653. doi:10.1086/284942.

Magoolagan, L., Mawby, P. J., Whitehead, F. A., and Sharp, S. P. (2019). The structure and context of male and female song in White-throated Dippers. *J. Ornithol.* 160, 195–205. doi:10.1007/s10336-018-1599-z.

Marler, P. (2004). "Chapter 1 - Science and birdsong: the good old days," in *Nature's Music*, eds. P. Marler and H. Slabbekoorn (Elsevier).

Marler, P. R., and Slabbekoorn, H. (2004). *Nature's music: the science of birdsong*. HAR/COM edition. Amsterdam ; Boston: Academic Press.

Marler, P., and Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science* 146, 1483–1486.

McNamara, A. (2011). Can we measure memes? *Front. Evol. Neurosci.* 3, 1. doi:10.3389/fnevo.2011.00001.

Mennill, D. J., and Rogers, A. C. (2006). Whip it good! Geographic consistency in male songs and variability in female songs of the duetting eastern whipbird *Psophodes olivaceus*. *J. Avian Biol.* 37, 93–100. doi:10.1111/j.0908-8857.2006.03548.x.

Morales, J., Gordo, O., Lobato, E., Ippi, S., Martínez-de la Puente, J., Tomás, G., et al. (2014). Female-female competition is influenced by forehead patch expression in pied flycatcher females. *Behav. Ecol. Sociobiol.* 68, 1195–1204. doi:10.1007/s00265-014-1730-y.

Murphy, T. G., Hernández-Muciño, D., Osorio-Beristain, M., Montgomerie, R., and Omland, K. E. (2009a). Carotenoid-based status signaling by females in the tropical streak-backed oriole. *Behav. Ecol.* 20, 1000–1006. doi:10.1093/beheco/arp089.

Murphy, T. G., Rosenthal, M. F., Montgomerie, R., and Tarvin, K. A. (2009b). Female American goldfinches use carotenoid-based bill coloration to signal status. *Behav. Ecol.* 20, 1348–1355. doi:10.1093/beheco/arp140.

Nelson, D. A., Hallberg, K. I., and Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology* 110, 879–908.

Nottebohm, F. (1972). The origins of vocal learning. *Am. Nat.* 106, 116–140.

Nottebohm, F. (1977). "Asymmetries in neural control of vocalization in the canary," in *Lateralization in the nervous system*, eds.S. Harnad, R.W. Doty, L. Goldstein, J. Jaynes, and G. Krauthamer (New York: Academic Press), 23–44.

Nottebohm, F. (1980). Testosterone triggers growth of brain vocal control nuclei in adult female canaries. *Brain Res.* 189, 429–436. doi:10.1016/0006-8993(80)90102-X.

Oberweger, K., and Goller, F. (2001). The metabolic cost of birdsong production. *J. Exp. Biol.* 204, 3379–3388.

Odom, K. J., and Benedict, L. (2018). A call to document female bird songs: Applications for diverse fields. *The Auk* 135, 314–325. doi:10.1642/AUK-17-183.1.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., and Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5. doi:10.1038/ncomms4379.

Oren, K., W., F. M., and Nicole, C. (2018). Integrative studies of cultural evolution: crossing disciplinary boundaries to produce new insights. *Philos. Trans. R. Soc. B Biol. Sci.* 373, 20170048. doi:10.1098/rstb.2017.0048.

Parker, K. A., Anderson, M. J., Jenkins, P. F., and Brunton, D. H. (2012). The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecol. Lett.* 15, 778–785. doi:10.1111/j.1461-0248.2012.01797.x.

Paton, J. A., Manogue, K. R., and Nottebohm, F. (1981). Bilateral organization of the vocal control pathway in the budgerigar, *Melopsittacus undulatus*. *J. Neurosci.* 1, 1279–1288.

Payne, R. B. (1981). "Population structure and social behavior: models for testing the ecological significance of song dialects in birds," in *Natural Selection and Social Behavior: Recent Research and New Theory*, eds. R. D. Alexander and D. W. Tinkle (New York: Chiron Press), 108–120.

Payne, R. B., Payne, L. L., and Doehlert, S. M. (1988). Biological and cultural success of song memes in indigo buntings. *Ecology* 69, 104–117. doi:10.2307/1943165.

Perreault, C. (2012). The pace of cultural evolution. *PLoS ONE* 7, e45150. doi:10.1371/journal.pone.0045150.

Peters, S. S., Searcy, W. A., and Marler, P. (1980). Species song discrimination in choice experiments with territorial male swamp and song sparrows. *Anim. Behav.* 28, 393–404. doi:10.1016/S0003-3472(80)80048-0.

Planqué, R., Britton, N. F., and Slabbekoorn, H. (2013). On the maintenance of bird song dialects. *J. Math. Biol.*, 1–27. doi:10.1007/s00285-012-0632-8.

Podos, J., and Warren, P. S. (2007). "The evolution of geographic variation in birdsong," in *Advances in the Study of Behavior*, eds. H. J. Brockmann, T. J. Roper, M. Naguib, K. E. Wynne-Edwards, C. Barnard, and J. Mitani (Academic Press), 403–458. Available at: http://www.sciencedirect.com/science/article/pii/S0065345407370095.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., and Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature* 434, 455.

Potvin, D. A., and Clegg, S. M. (2015). The relative roles of cultural drift and acoustic adaptation in shaping syllable repertoires of island bird populations change with time since colonization. *Evolution* 69, 368–380. doi:10.1111/evo.12573.

Prat, Y., Azoulay, L., Dor, R., and Yovel, Y. (2017). Crowd vocal learning induces vocal dialects in bats: Playback of conspecifics shapes fundamental frequency usage by pups. *PLOS Biol.* 15, e2002556. doi:10.1371/journal.pbio.2002556.

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Behav. Evol. Ecol.* 3, 40. doi:10.3389/fevo.2015.00040.

Price, J. J. (2019). Sex differences in song and plumage color do not evolve through sexual selection alone: new insights from recent research. *J. Ornithol.* 160, 1213–1219. doi:10.1007/s10336-019-01681-8.

Price, J. J., and Eaton, M. D. (2014). Reconstructing the evolution of sexual dichromatism: current color diversity does not reflect past rates of male and female change. *Evolution* 68, 2026–2037. doi:10.1111/evo.12417.

Priyadarshani, N., Marsland, S., and Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. *J. Avian Biol.* 49, jav-01447. doi:10.1111/jav.01447.

Reichard, D. G., Brothers, D. E., George, S. E., Atwell, J. W., and Ketterson, E. D. (2018). Female dark-eyed juncos *Junco hyemalis thurberi* produce male-like song in a territorial context during the early breeding season. *J. Avian Biol.* 49, jav-01566. doi:10.1111/jav.01566.

Richerson, P. J., Boyd, R., and Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *Proc. Natl. Acad. Sci.* 107, 8985–8992. doi:10.1073/pnas.0914631107.

Riebel, K., Lachlan, R. F., and Slater, P. J. B. (2015). "Chapter Six - Learning and cultural transmission in chaffinch song," in *Advances in the Study of Behavior*, eds. M. Naguib, J. C. Mitani, L. W. Simmons, L. Barrett, S. Healy, and P. J. B. Slater (Academic Press), 181–227. Available at: http://www.sciencedirect.com/science/article/pii/S0065345415000029.

Rogers, A. C., Langmore, N. E., and Mulder, R. A. (2007). Function of pair duets in the eastern whipbird: cooperative defense or sexual conflict? *Behav. Ecol.* 18, 182–188. doi:10.1093/beheco/arl070.

Roper, M. M. (2018). Sexual dimorphism of song and life history trade-offs in the New Zealand bellbird: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph. D.) in Ecology at Massey University, Albany, New Zealand.

Rose, E. M., Coss, D. A., Haines, C. D., Danquah, S. A., Studds, C. E., and Omland, K. E. (2019). Why do females sing?—pair communication and other song functions in eastern bluebirds. *Behav. Ecol.* 30, 1653–1661. doi:10.1093/beheco/arz130.

Rothenberg, D., Roeske, T. C., Voss, H. U., Naguib, M., and Tchernichovski, O. (2014). Investigation of musicality in birdsong. *Hear. Res.* 308, 71–83. doi:10.1016/j.heares.2013.08.016.

Rothstein, S. I., and Fleischer, R. C. (1987). Vocal dialects and their possible relation to honest status signalling in the brown-headed cowbird. *Condor*, 1–23.

Salinas-Melgoza, A., and Wright, T. F. (2012). Evidence for vocal learning and limited dispersal as dual mechanisms for dialect maintenance in a parrot. *PLoS ONE* 7, e48667. doi:10.1371/journal.pone.0048667.

Sanvito, S., Galimberti, F., and Miller, E. H. (2007). Observational evidences of vocal learning in southern elephant seals: a longitudinal study. *Ethology* 113, 137–146. doi:10.1111/j.1439-0310.2006.01306.x.

Searcy, W. A., and Yasukawa, K. (1995). *Polygyny and sexual selection in red-winged blackbirds*. Princeton University Press.

Shennan, S. (2011). Descent with modification and the archaeological record. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 1070–1079. doi:10.1098/rstb.2010.0380.

Shultz, A. J., and Burns, K. J. (2017). The role of sexual and natural selection in shaping patterns of sexual dichromatism in the largest family of songbirds (Aves: Thraupidae). *Evolution* 71, 1061–1074. doi:10.1111/evo.13196.

Slabbekoorn, H., and Ripmeester, E. a. P. (2008). Birdsong and anthropogenic noise: implications and applications for conservation. *Mol. Ecol.* 17, 72–83. doi:10.1111/j.1365-294X.2007.03487.x.

Slagsvold, T., Hansen, B. T., Johannessen, L. E., and Lifjeld, J. T. (2002). Mate choice and imprinting in birds studied by cross-fostering in the wild. *Proc. R. Soc. B Biol. Sci.* 269, 1449–1455. doi:10.1098/rspb.2002.2045.

Slater, P. J. B. (1986). The cultural transmission of bird song. *Trends Ecol. Evol.* 1, 94–97.

Slater, P. J. B. (2003). Fifty years of bird song research: a case study in animal behaviour. *Anim. Behav.* 65, 633–639. doi:10.1006/anbe.2003.2051.

Slater, P. J. B., and Mann, N. I. (2004). Why do the females of many bird species sing in the tropics? *J. Avian Biol.* 35, 289–294. doi:10.1111/j.0908-8857.2004.03392.x.

Smiseth, P. T., and Amundsen, T. (2000). Does female plumage coloration signal parental quality? A male removal experiment with the bluethroat (*Luscinia s. svecical*). *Behav. Ecol. Sociobiol.* 47, 205–212.

Soltis, J. (2009). Vocal production, affect expression, and communicative function in African elephant rumble vocalizations. *J. Acoust. Soc. Am.* 126, 2210–2210. doi:10.1121/1.3248716.

Sperber, A. L., Werner, L. M., Kappeler, P. M., and Fichtel, C. (2017). Grunt to go—Vocal coordination of group movements in redfronted lemurs. *Ethology* 123, 894–905. doi:10.1111/eth.12663.

Starnberger, I., Preininger, D., and Hödl, W. (2014). From uni- to multimodality: towards an integrative view on anuran communication. *J. Comp. Physiol. A* 200, 777–787. doi:10.1007/s00359-014-0923-1.

Templeton, C. N., Rivera-Cáceres, K. D., Mann, N. I., and Slater, P. J. B. (2011). Song duets function primarily as cooperative displays in pairs of happy wrens. *Anim. Behav.* 82, 1399–1407. doi:10.1016/j.anbehav.2011.09.024.

Tobias, J. A., Montgomerie, R., and Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2274–2293. doi:10.1098/rstb.2011.0280.

Trainer, J. M. (1989). Cultural evolution in song dialects of yellow-rumped caciques in Panama. *Ethology* 80, 190–204. doi:10.1111/j.1439-0310.1989.tb00739.x.

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., and Dale, J. (2016). Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. *Front. Ecol. Evol.* 4, 22.

Whitehead, H., and Rendell, L. (2014). *The cultural lives of whales and dolphins*. University of Chicago Press.

Whiten, A., Hinde, R. A., Laland, K. N., and Stringer, C. B. (2011). Culture evolves. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 938–948. doi:10.1098/rstb.2010.0372.

Wilkins, M. R., Seddon, N., and Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. *Trends Ecol. Evol.* 28, 156–166. doi:10.1016/j.tree.2012.10.002.

Winn, H. E., Thompson, T. J., Cummings, W. C., Hain, J., Hudnall, J., Hays, H., et al. (1981). Song of the humpback whale — Population comparisons. *Behav. Ecol. Sociobiol.* 8, 41–46. doi:10.1007/BF00302842.

Wright, T. F., and Dahlin, C. R. (2018). Vocal dialects in parrots: patterns and processes of cultural evolution. *Emu - Austral Ornithol.* 118, 50–66. doi:10.1080/01584197.2017.1379356.

2 Female song occurs in songbirds with more elaborate female colouration and reduced sexual dichromatism

2.1 Abstract

Elaborate plumages and songs in male birds provide classic evidence for Darwinian sexual selection. However, trait elaboration in birds is not gender-restricted: female song has recently been revealed as a taxonomically-widespread trait within the songbirds (oscine Passerines), prompting increased research into likely functions and social/ecological correlates. Here we use phylogenetically-informed comparative analysis to test for an evolutionary association between female song and plumage colour elaboration in songbirds. If there is an evolutionary trade-off between signalling modes, we predict a negative correlation between acoustic and visual elaboration. This trade-off hypothesis has been commonly proposed in males but has mixed empirical support. Alternatively, if song and plumage have similar or overlapping functions and evolve under similar selection pressures, we predict a positive correlation between female song and female plumage elaboration. We use published data on female song for 1023 species of songbirds and a novel approach that allows for the reliable and objective comparison of colour elaboration between species and genders. Our results reveal a significant positive correlation between female colourfulness and female song presence. In species where females sing, females (but not males) are on average more colourful—with concomitantly reduced average sexual dichromatism. These results suggest that female plumage colour and female song likely evolved together under similar selection pressures and that their respective functions are reinforcing. We discuss the potential roles of sexual versus social selection in driving this relationship, and the implications for future research on female signals.

2.2 Introduction

Elaborate male traits often provide evidence for classic sexual selection (Darwin, 1871), but the possession of elaborate traits by females is less well understood. A traditional perspective holds that female trait elaboration is non-adaptive, perhaps detrimental, and results from "shared genetic architecture" with males (Darwin, 1871; Lande, 1980; see Tobias et al., 2012 for review). While current evidence does suggest that female ornamentation is correlated with conspecific male ornamentation to some extent (Bonduriansky and Chenoweth, 2009; Dale et al., 2015b) the view that elaborate female traits are purely non-adaptive pleiotropic effects has been strongly refuted by recent research.

Bird song and elaborate plumage are often considered costly traits (Song: Oberweger and Goller, 2001; Nowicki et al., 2002; Berg et al., 2005; Schmidt et al., 2013; but see Gil et al., 2006. Plumage: McGraw et al., 2002; Walther and Clayton, 2005; Simpson et al., 2015). In females, both traits can be evolutionarily labile with respect to the conspecific male phenotype. First, female plumage changes have played a greater role than male changes in the evolution of dichromatism in New World blackbirds (Icteridae; Irwin, 1994; Hofmann et al., 2008; Price and Eaton, 2014), tanagers (Thraupidae; Burns, 1998) and fairy-wrens (Maluridae; Johnson et al., 2013) and this pattern was shown to hold true for the order Passeriformes as a whole (Dale et al., 2015b).

Second, female song has been recovered as the ancestral condition of songbirds (Odom et al., 2014), indicating that multiple losses of female song have occurred over evolutionary time despite conspecific males retaining the trait (Price, 2015). This shows that female traits are not necessarily tightly constrained by the male phenotype, but are able to evolve rapidly and independently.

Furthermore, female-specific functions of song and plumage ornamentation have been resolved for a growing number of species, revealing these traits to be potentially adaptive for females. For instance, female song has been shown to function in territory and resource defence, mate attraction, mate defence, and pair-bonding (Searcy and Yasukawa, 1995; Langmore et al., 1996; Rogers et al., 2007; Brunton et al., 2008; Templeton et al., 2011; Hall et al., 2015). Female colouration is attractive to mates in some species (Amundsen et al., 1997; Smiseth and Amundsen, 2000; Murphy et al., 2009a, 2009b) and frequently has roles in female–female competition for non-sexual breeding resources (Tobias et al., 2012; Morales et al., 2014).

If elaborated female traits are adaptive, this raises the question: what is the evolutionary relationship between female song and plumage colour elaboration?

We outline three hypotheses regarding the evolution of multiple elaborate traits: the traits may evolve "in a concerted fashion, in an antagonistic fashion, or in ways unrelated to each other" (Shutler and Weatherhead, 1990). In our context, these options are detailed as follows:

(i) The "trade-off" hypothesis (Darwin, 1871) argues that if two modes of signalling are both costly to produce or maintain, then selection might favour doing one thing well rather than two things poorly. This predicts an evolutionary trade-off (an inverse correlation) between signalling modes (Shutler and Weatherhead, 1990). Thus, species with female song are predicted to have less colourful females on average than species lacking female song.

(ii) The "reinforcing signals" hypothesis (inspired by the "redundant signal" hypothesis of Møller and Pomiankowski, 1993) proposes that the two ornamental traits have overlapping, reinforcing functions, acting in concert to convey the condition or status of the signaller. Whereas either trait on its own provides a partial indication of signaller condition or status, in combination the multiple ornaments enable a more accurate assessment by rivals or mates. This predicts a co-evolution of the two traits; that is, a positive correlation between song and plumage colour elaboration in females. Thus, species with female song are predicted to have more colourful females on average than species lacking female song.

(iii) The "multiple messages" hypothesis (adapted from Møller and Pomiankowski, 1993) posits that the two traits reveal different information about their bearers. This implies that the characters are driven by different selective pressures, and therefore elaborate vocal and visual sexual signals should evolve independently.

Thus, there should be no difference in female colour elaboration between species with female song, and those without female song.

Current evidence is inconclusive regarding the three hypotheses. In line with the trade-off hypothesis, an inverse relationship has been found between male plumage brightness and song complexity in cardueline finches (Badyaev et al., 2002). But in support of the reinforcing signals hypothesis, a positive correlation has been found between song length and number of coloured patches among Asian barbets (Gonzalez-Voyer et al., 2013), and a positive correlation between the degree of dichromatism and time spent singing among wood warblers (Shutler and Weatherhead, 1990). Finally, Ornelas et al. (2009) found no relationship between dichromatism and song complexity among trogons, and Mason et al. (2014) found no correlation between song and plumage complexity among the tanagers. These conflicting results likely reflect biological and evolutionary differences among focal taxa, as well as methodological differences in how song and plumage were quantified (Mason et al., 2014). Furthermore, all these studies focus on males. The possibility of a generalised macroevolutionary association between multiple ornamental traits in females has not yet been investigated.

In this study we test for an evolutionary correlation between female song and plumage elaboration (male, female and dichromatism) across the songbirds (i.e., Oscines; order Passeriformes, suborder Passeri). We perform phylogenetically-informed comparative analysis using song and plumage data from repositories supplemented with additional data gleaned from the literature.

2.3 Methods

2.3.1 Scoring female song

We compiled data describing the presence or absence of song in male and female songbirds (Oscines) from Odom et al. (2014) and del Hoyo et al. (2004–2011). We gave each species one of four scores according to the criteria of Odom et al. (2014, full details therein). Scores included: “present,” both males and females of the species sing; “absent,” only the male sings; “songless,” neither sex sings; or “not enough information” if we could not reliably make a designation. (Note that no species where only females sing has been described.) Out of all 4814 songbird species, 1314 had sufficient information on song to reliably score the species. Because the lack of female song in songless species might be the product of different selection pressures than in species with male song, we omitted songless species (291 species) from our analysis. Our final species pool included 1023 singing species comprised of 656 species where both sexes sing (64%), and 367 species where only males sing (36%).

2.3.2 Scoring plumage elaboration

Plumage colour scores for the 1023 songbird species were obtained from Dale et al. (2015a). Briefly, for both sexes of each species of passerine (Order: Passeriformes), the mean red, green and blue (RGB) values on 3 dorsal and 3 ventral patches were measured using digital image processing software (Valcu and Dale, 2013) on scanned images from handbook plates. For each patch of each sex of each species, it was determined how “male-like” that patch is by scoring the proportion of males in the nearest 1% of similarly coloured patches in other species. The method results in scores where low values correspond to males or females with drab, classically-“female-like” plumage, and high scores correspond to males or females with elaborate, classically-“male-like” plumage. This approach is transferable to other colour quantification methods. There is a high correlation between scores determined with handbook plates vs. analogous scores determined with UV-VIS (ultraviolet to visible) reflectance spectra from museum specimens (Dale et al., 2015b). This result provides critical validation of the method because although human and avian vision have considerable overlap (Badyaev and Hill, 2003; Seddon et al., 2010), birds can also see UV light not visible to humans (Cuthill, 2006). See Dale et al. (2015b) for detailed methodology. Sexual dichromatism was calculated for each species as the male plumage colour score minus the female plumage score.

2.3.3 Trait correlation

To test for an evolutionary correlation between female song and female plumage elaboration, we first performed Pagel's correlation test (Pagel, 1994), in R 3.1.2 (R Development Core Team, 2014) using the “geiger” and “phytools” packages (Harmon et al., 2008; Revell, 2012). The Pagel test controls for phylogenetic relatedness and requires no designation of independent and response variables. We assigned a song character state and a plumage character state to each tip of a phylogeny, and tested the null hypothesis that the two traits had evolved independently. As the Pagel test requires both traits to be binary, plumage scores were binned into binary characters according to an arbitrary cut-off, which was moved in integer increments from 35 (1022 of 1023 species with female plumage elaboration present) to 71 (1 of 1023 species), to study the sensitivity of the correlation test to changes in plumage cutoff value. The phylogenies used for this analysis were obtained from the Hackett backbone (Hackett et al., 2008) supertrees at http://birdtree.org (Jetz et al., 2012).

2.3.4 MCMCglmm analysis

To estimate the strength of the correlation between female song presence and female plumage elaboration, we fit a multivariate generalised linear mixed model using the "MCMCglmm" package (Monte Carlo Markov Chain generalised linear mixed model; Hadfield, 2010) in R (version 3.1.2). MCMCglmm allowed us to fit a model which had a response vector that contained a mixture of Gaussian and non-Gaussian distributed variables. Female plumage elaboration is a continuous measure (see above); hence we assumed a Gaussian error distribution. For female song, the response vector contained binary presence scores (0 = female song absent, and 1 = female song present) and accordingly we assumed a Bernouli error distribution and used a logit link function. Fixed effects in our model included female plumage elaboration, the presence of female song, and male plumage elaboration as a covariate. Phylogeny was fit as a random effect using the methods described in Hadfield and Nakagawa (2010) to calculate the inverse numerator relationship for phylogenetic effects. For the phylogenetic effects we also allowed separate random intercepts for female song and female plumage elaboration and a non-zero covariance between these two traits by assuming an unstructured variance-covariance structure.

Priors for the location effects were diffuse about zero and had a large variance (10^8). For the variance components we used priors conforming to a scaled non-central F-distribution (Gelman, 2006) with the location parameter equal to zero. The scale parameter for female colour elaboration was equal to half of the phenotypic variation in female colour elaboration, and for female song the scale parameter was equal to p(1−p), where p is the mean probability of female song across the dataset. For the residual covariance matrix we assumed an inverse-Wishart distributed prior for female colouration. For female song (which is a binary trait), it is not possible to estimate a residual variance, so we fixed the prior at a value of 1 (Hadfield, 2014). The MCMC chain had 20,600,000 iterations, with a burn-in of 600,000 and a thinning interval of 20,000, resulting in ~1000 samples of the posterior distribution of the parameters. Model fit was confirmed by ensuring that autocorrelation was low and the trait means lay within the 95% highest posterior density (HPD) intervals of the posterior predictive distribution of each trait. To incorporate some of the uncertainty in the phylogenetic relationships among bird species, we applied the statistical model described above to 10 different phylogenetic trees randomly selected from http://birdtree.org (Jetz et al., 2012). Finally, we examined the convergence of the phylogenetic variances and covariances estimated from the 10 models (each using different trees, and therefore with different numerator relationship matrices for the phylogenetic effects), with the Gelman and Rubin (1992) diagnostic, R. For these 10 trees, the point estimate was $R = 1.2$ indicating moderate convergence. This phylogenetic uncertainly is incorporated in all the estimates of the posterior means and the HPD intervals we present.

To assess the significance of the phylogenetic correlation between female song (FS) and female plumage elaboration (FP), we first calculated the posterior distribution of the correlation using

$$cor_{(FS,FP)} = cov_{(FS,FP)} \div (var_{FS} \times var_{FP})^{0.5}$$

If the 95% HPD intervals of the posterior distribution of the correlation did not overlap zero, we interpreted this as evidence for a significant phylogenetic correlation between female plumage elaboration and female song. Similarly, to assess the significance of the association between male colour elaboration and our two female traits, we tested whether the 95% HPD intervals of the fixed interactions between male plumage and female plumage elaboration, as well as between male plumage and female song, overlapped zero.

2.4 Results

2.4.1 Female song in relation to plumage colour elaboration

For many species in our sample, males and females have similar plumage elaboration scores (Figure 2.1A, points lying along the diagonal). However, there are also many sexually dichromatic species where the male is more colourful and the female is more drab (Figure 2.1A, points in the upper left of the distribution). Overlaying the density of female song presence on this plumage elaboration scatter (Figure 2.1B) reveals that female song presence is most concentrated amongst species with high female plumage scores and reduced average sexual dichromatism.

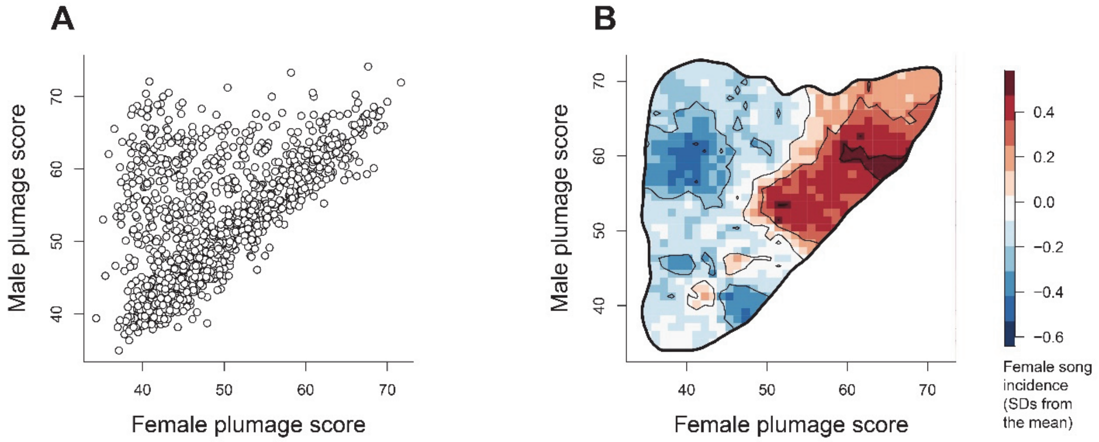


Figure 2.1—Female song incidence in relation to male and female plumage elaboration

(A) Male plumage score vs. female plumage score for each species (N = 1023). **(B)** Contour map depicting the average of the (scaled) female song values overlaid within the 99.8% volume contour of the male vs. female plumage score distribution. Values within the plot were calculated by superimposing a 50 × 50 grid over the scatter occurring between 30 and 80 and then calculating at each grid point the mean female song scores of the closest 3% of species. The units represent standard deviations away from the mean incidence of song, with redder values indicating increasingly higher occurrence of female song.

2.4.2 Phylogenetic distribution

The phylogenetic distribution of the co-occurrence of female elaboration traits is visualised in Figure 2.2, where colour-coded species tips correspond to the presence/absence of female song and female plumage elaboration. Instances where both traits co-occur (257 species, 25% of the total sample) are not concentrated within one or a few clades but are dispersed throughout the tree, suggesting many separate origins of the co-evolution of these traits (Figure 2.2). Wide phylogenetic dispersion persists when higher plumage elaboration cutoff values or different trees from Jetz et al. (2012) are used. Also numerous and widely dispersed around the tree are species with only one trait or the other: female song but not female plumage elaboration (399 species, 39% of the total sample); female plumage elaboration but not female song (72 species, 7% of the total sample), and neither trait present (295 species, 29% of the total sample).

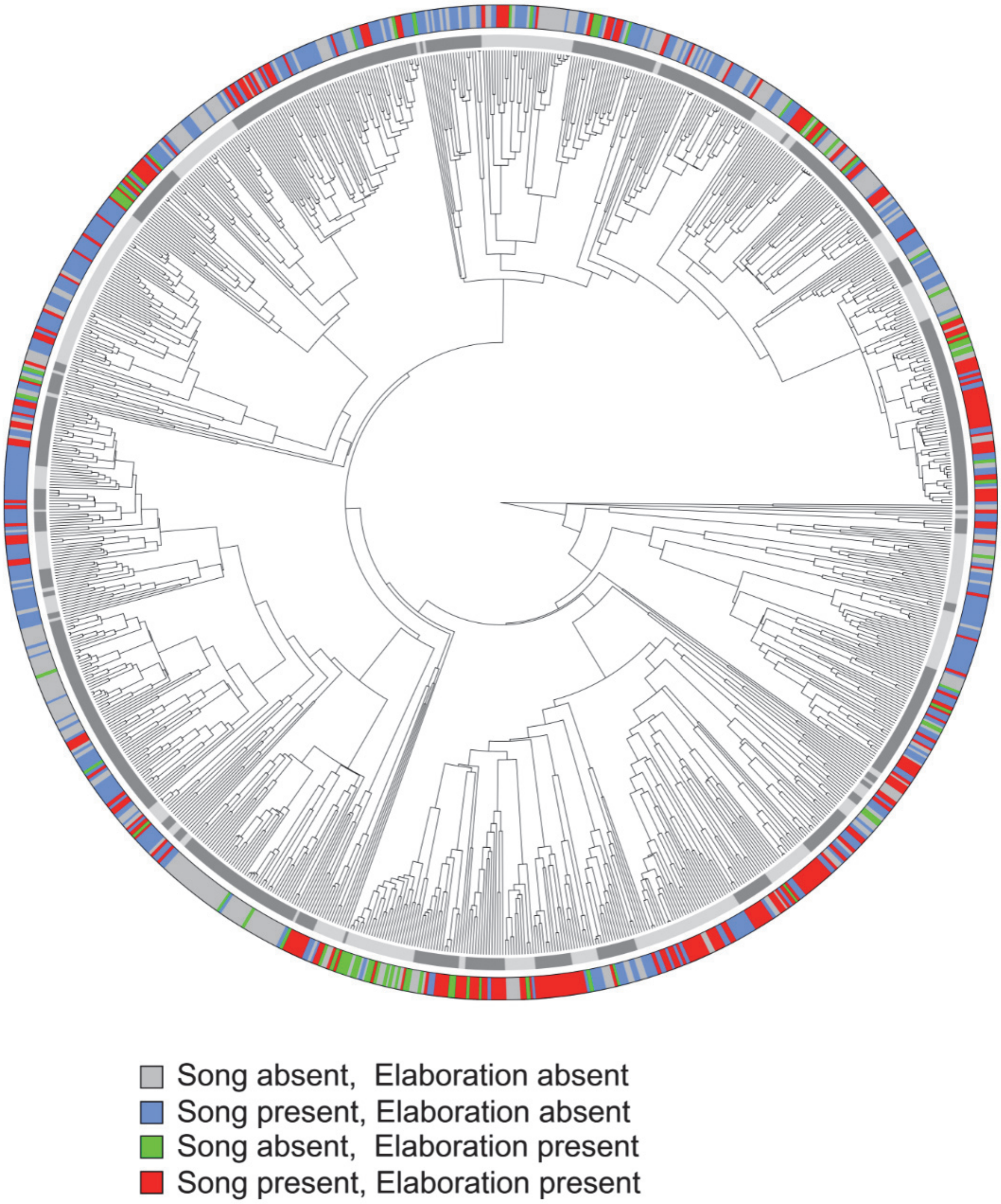


Figure 2.2—Phylogenetic distribution of female song and female plumage elaboration.

Plumage elaboration is here treated as a binary character with cutoff = plumage score of 50. Red labels represent species where female song and female plumage elaboration (plumage score >50) co-occur. The inner ring of alternating grey tones depicts boundaries of the 86 oscine families in our sample.

2.4.3 Pagel correlation test

The Pagel test provides strong support for correlated evolution of female song presence and female plumage elaboration for all plumage cut-off values between 41 and 54 (Figure 3, $P < 0.0001$). This is a large range by comparison to the entire range of female plumage scores (35–71), and even at a cut-off as high as 60, the correlation is still statistically significant at the 5% level. The presence of an evolutionary association between female song and female colour elaboration is therefore highly robust to alternative cut-off values used to categorize female colouration as elaborated.

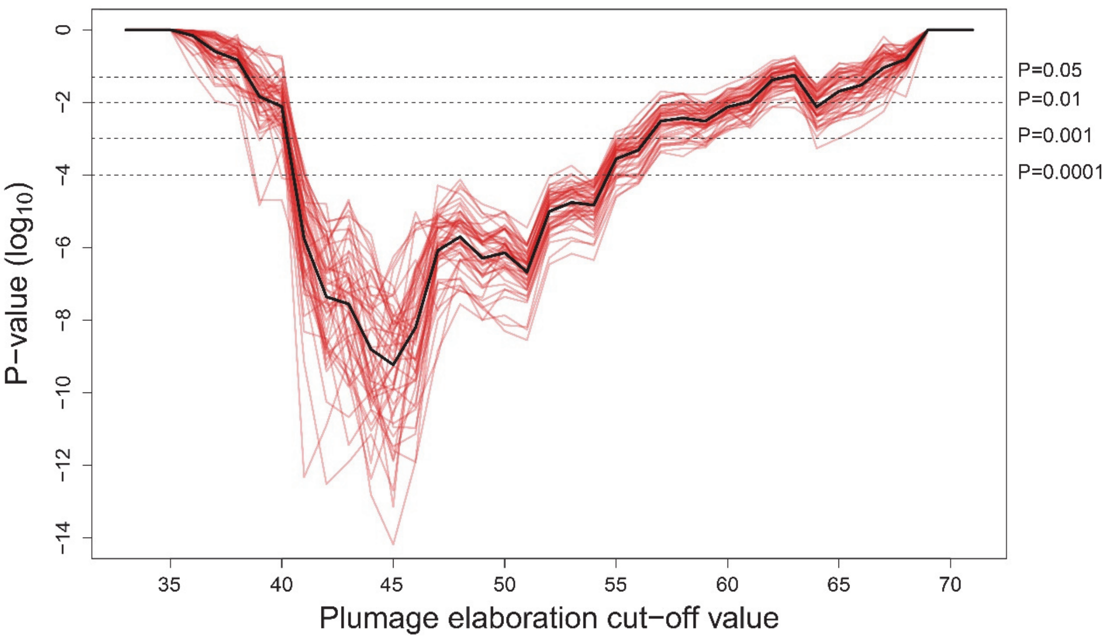


Figure 2.3—Pagel correlation test *P*-values (Pagel, 1994) at all possible plumage elaboration cutoff values, for each of 50 phylogenetic trees from http://birdtree.org (Jetz et al., 2012).

The black line indicates the mean of the 50 runs (red lines). Evidence of correlated evolution is strong (P<0.0001) for all plumage cut-off values between 41 and 54.

2.4.4 MCMCglmm

Despite the phylogenetic uncertainty introduced by using 10 randomly selected phylogenetic trees from Jetz et al. (2012), we found a strongly significant evolutionary correlation between female plumage elaboration and the presence of female song (Figure 2.4, mean $cor_{(FS,FP)}$ = 0.402, 95% HPD = 0.220–0.583). Accounting for evolutionary relationships among bird species, we found that male and female colour elaboration were positively associated (posterior mean = 0.472, 95% HPD = 0.422–0.527), confirming the apparent pattern seen in Figure 2.1. In addition there was a positive, but not statistically significant, relationship between the degree of male colour elaboration and the presence of female song (posterior mean = 0.044, 95% HPD = −0.014–0.102).

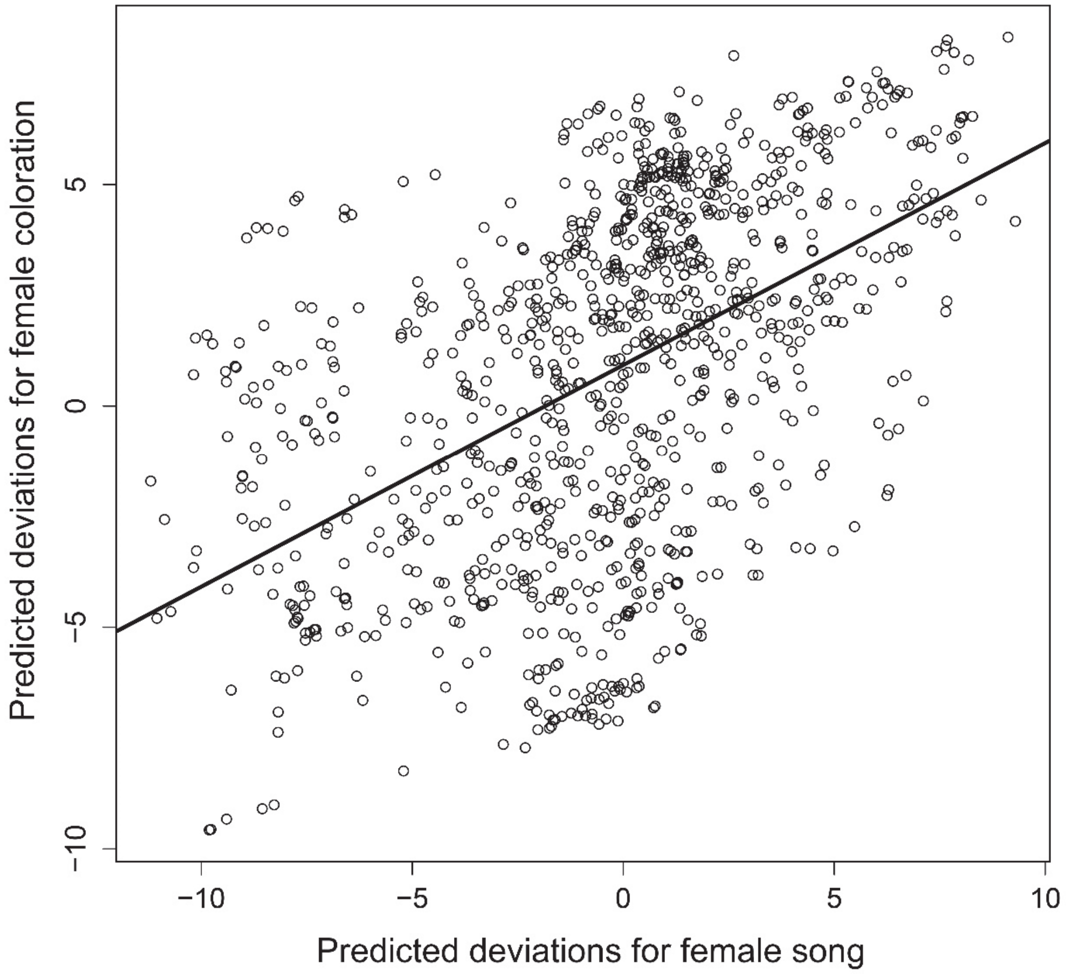


Figure 2.4—Plot of best linear unbiased predictions of the phylogenetic association between female song and female plumage colour elaboration.

Points indicate the predicted phylogenetic deviations of each species (i.e., the 1023 species tips within the pruned phylogenetic tree) for the two female display traits, and solid line denotes the posterior mean phylogenetic correlation between female song with female plumage.

2.5 Discussion

We used phylogenetically-informed analysis to reveal the relationship between female song presence and plumage elaboration across the songbirds. Our study shows that (1) female song is more concentrated amongst species with elaborated (classically “male-like”) female plumage, (2) the co-occurrence of female song and female plumage elaboration is widely dispersed across the songbird phylogeny, suggesting many independent origins of this evolutionary association, (3) Pagel correlation tests demonstrate clear support for an evolutionary correlation between female song and female plumage colour elaboration, and (4) estimates from MCMCglmm models suggest that the presence of female song accounts for 16.2% ($R = 0.402$) of the variation in female plumage elaboration (and vice versa).

Mason et al. (2014) provided the largest comparative study of multimodal sexual signalling to date, finding no relationship between plumage elaboration and song elaboration across the males of 301 tanager species (Thraupidae). In contrast, our results show a clear positive relationship between female song and female plumage colour elaboration across the songbirds. The difference between our results and those of Mason et al. could stem from a number of factors. First, Mason et al. consider only male elaboration. The selective pressures acting on males and females are not necessarily equivalent, and so female traits may exhibit a different relationship to each other than do male traits (Tobias et al., 2012). Second, we use different methods of scoring song and plumage elaboration; notably, the song data of Mason et al. are quantitative, whereas our data are presence/absence. Finally, there are differences in scale between our studies. The presence of negative or non-significant relationships within some families such as the Thraupidae does not preclude a positive correlation at the broader taxonomic level of the songbirds as a whole. These considerations apply equally to other studies of multimodal signalling which find a negative relationship or no relationship between song and plumage (Badyaev et al., 2002; Ornelas et al., 2009).

Our results are most consistent with the "reinforcing signals" hypothesis; a positive correlation between female song and female plumage elaboration is expected if the two signalling modes tend to have reinforcing, overlapping functions, and thus have co-evolved together in response to similar selective pressures. Females may use song and plumage as a multimodal signal to reinforce the reliability and/or potency of the message to receivers.

A central tenet of sexual selection theory is that males and females often differ in their routes to reproductive success; for males, competition for mates is paramount, whereas for females, access to resources that affect fecundity is thought to be of greater importance (Rosvall, 2011). Therefore, the primary selection operating on females is likely to be non-sexual "social selection" for ecological or social resources, such as foraging territories, nest sites and paternal investment (West-Eberhard, 1979, 1983; Tobias et al., 2012). Such a view would suggest that a key function of both female song and plumage colour elaboration may be signalling the female status (or resource-holding potential: Searcy and Nowicki, 2005) to competitive rivals, a prediction that is borne out in a number of studies (Tobias et al., 2012). This is not to deny that sexual selection for elaborate female traits may also be important (Clutton-Brock, 2007; Rosvall, 2011; Clutton-Brock and Huchard, 2013); indeed, the same ornaments can perform both sexual and (non-sexual) social functions (Kraaijeveld et al., 2004).

The co-occurrence of both visual and acoustic indicators of status (or resource-holding potential) likely reinforces the overall message and facilitates effective communication under different signalling scenarios. Song can be communicated over longer distances, without a clear line of sight; and as sound is propagated radially (Fahy and Gardonio, 2007), vocal communication does not depend on precise directionality of the signaller in relation to the receiver. Song is amenable to rapid temporal changes, conveying the short-term intentions of the individual and encoding complex information about signaller identity. Plumage colouration, by contrast, is a more permanent feature (though birds may be able to mediate plumage display behaviourally). In general, plumage colouration is a more direction-sensitive close-range signal than song. Thus, to signal quality and ward off rivals, selection could favour song for long-range broadcasting and plumage elaboration for close encounters, even if the message and intended receiver are the same.

If differing signalling scenarios promote female song or plumage elaboration differentially, this might explain the many species in our sample with only one trait or the other. Given that 16.2% of variation in female plumage elaboration is explained by female song presence (and vice versa), a remaining 83.8% of variation in each trait is thus attributable to other factors. That is, the relationship between female song and female plumage elaboration is complex and likely mediated by additional effects of, for example, habitat type, predation risk, territory size, and social structure. Closed habitat might strongly favour female song for effective communication if the range and efficacy of visual signalling is impaired. Or, for species experiencing high predation, elaborate female plumage may impose too great a risk around the nest (Martin and Badyaev, 1996) and be selected against. Or, in colony-living species where individuals remain in close proximity, close-range visual signalling may be sufficient for female signalling needs, making female song redundant. In short, there is much work left to do in identifying and quantifying the factors that contribute to visual and acoustic ornamentation in female songbirds.

If song and elaborate plumage are both costly (e.g. Oberweger and Goller, 2001; McGraw et al., 2002), as expected if the signals are “honest” indicators of aspects of female quality, then why have we not observed a trade-off between signalling modes? There are a number of potential explanations. First, while our results demonstrate that *presence* of female song does not trade off with female plumage elaboration, we lack information on the degree of female song *elaboration* (e.g., in terms of repertoire size, vocal agility, and time spent singing) and thus we cannot evaluate whether there is some degree of trade-off of resource allocation between modes. As more studies focus on quantifying the vocal performance of female birds, addressing whether singing females with bright plumage have “cheaper” songs than those with drab plumage will be practicable. It is also conceivable that when comparing across species, trade-offs might not manifest if different species are selected to invest different levels of resources into the overall message.

Another possibility is that resources involved in song and plumage development are not limiting for species where status signalling is strongly favoured by selection. Rather than functioning as condition-dependent indices or handicaps, the honesty of these signals may instead be socially enforced, with cheaters (i.e., individuals who signal having greater status than they actually have) being punished through increased aggression by conspecifics (Tibbetts and Dale, 2004; Tibbetts and Izzo, 2010). The social cost in such cases will depend on how accurately the signal reflects true quality of the individual, rather than the number of ornaments involved in the signal *per se*. That is, under social costs, female song and plumage elaboration may be no more costly to produce than either trait alone, in which case we would not expect a trade-off between traits.

2.6 Conclusion and future directions

We have demonstrated strong evidence for a positive co-evolutionary relationship between plumage elaboration and song in female songbirds, a result which supports an overlapping function of the two traits (i.e., the reinforcing signals hypothesis). We have suggested, in light of current selection theory (Tobias et al., 2012), that the primary context for this multimodal signalling is non-sexual social competition for ecological or social resources, and that the different signalling ranges of plumage and song may have favoured the evolutionary maintenance of both traits. Our finding raises several questions for future research, including: (1) Does this pattern hold for the sub-oscines? Though generally poorer singers than the songbirds, and lacking vocal learning (but see Kroodsma et al., 2013), many sub-oscines vocalize for mate attraction and territorial defence (Chelén et al., 2005) and thus may be subject to similar evolutionary pressures. (2) Are female ornaments gained and lost more frequently than male ornaments (Kraaijeveld, 2014) and is there a consistent order of female trait evolution? That is, do gains (or losses) of song follow gains (or losses) in plumage elaboration, or vice versa? And (3), what are the social and ecological drivers of the evolutionary association of female song and female plumage elaboration? It is our hope that the patterns reported in this study contribute to future research on the functions of both visual and acoustic ornamentation in females.

2.7 Author contributions

Wesley Webb, James Dale, and Dianne Brunton conceived of the study. Wesley Webb collected the song data. James Dale and Mihai Valcu collected the plumage data. Wesley Webb, James Dale, David Aguirre and Daniel Thomas analysed the data. Wesley Webb, James Dale, and David Aguirre wrote the paper with input from the other authors.

2.8 Supporting information

This chapter has been adapted from the following published paper:

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., & Dale, J. (2016). Female Song Occurs in Songbirds with More Elaborate Female Coloration and Reduced Sexual Dichromatism. *Frontiers in Ecology and Evolution*, 22. https://doi.org/10.3389/fevo.2016.00022

See Appendix 7.1 for a reprint of the published paper and associated research outputs.

2.9 References

Amundsen, T., Forsgren, E., and Hansen, L. T. T. (1997). On the function of female ornaments: Male bluethroats prefer colourful females. *Proc. Biol. Sci.* 264, 1579–1586.

Badyaev, A. V., and Hill, G. E. (2003). Avian sexual dichromatism in relation to phylogeny and ecology. *Annu. Rev. Ecol. Evol. Syst.* 34, 27–49.

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

Berg, M. L., Beintema, N. H., Welbergen, J. A., and Komdeur, J. (2005). Singing as a handicap: the effects of food availability and weather on song output in the Australian reed warbler *Acrocephalus australis*. *J. Avian Biol.* 36, 102–109.

Bonduriansky, R., and Chenoweth, S. F. (2009). Intralocus sexual conflict. *Trends Ecol. Evol.* 24, 280–288.

Brunton, D. H., Evans, B., Cope, T., and Ji, W. (2008). A test of the dear enemy hypothesis in female New Zealand bellbirds (*Anthornis melanura*): female neighbors as threats. *Behav. Ecol.* 19, 791–798. doi:10.1093/beheco/arn027.

Burns, K. J. (1998). A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. *Evolution* 52, 1219–1224. doi:10.2307/2411252.

Chelén, A. A. R., Garcia, C. M., and Riebel, K. (2005). Variation in the song of a sub-oscine, the vermilion flycatcher. *Behaviour* 142, 1115–1132. doi:10.1163/156853905774405326.

Clutton-Brock, T. (2007). Sexual selection in males and females. *Science* 318, 1882–1885. doi:10.1126/science.1133311.

Clutton-Brock, T. H., and Huchard, E. (2013). Social competition and selection in males and females. *Philos. Trans. R. Soc. B Biol. Sci.* 368, 20130074.

Cuthill, I. C. (2006). "Volume 1: Mechanisms and Measurements," in *Bird Coloration*, eds. G. E. Hill and K. J. McGraw (Harvard University Press), 3–40.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015a). Data from: The effects of life history and sexual selection on male and female plumage colouration. doi:10.5061/dryad.1rp0s.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015b). The effects of life history and sexual selection on male and female plumage colouration. *Nature* 527, 367.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Fahy, F. J., and Gardonio, P. (2007). *Sound and structural vibration: radiation, transmission and response*. Academic press.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1, 515–534.

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.*, 457–472.

Gil, D., Naguib, M., Riebel, K., Rutstein, A., and Gahr, M. (2006). Early condition, song learning, and the volume of song brain nuclei in the zebra finch (*Taeniopygia guttata*). *J. Neurobiol.* 66, 1602–1612. doi:10.1002/neu.20312.

Gonzalez-Voyer, A., den Tex, R.-J., Castelló, A., and Leonard, J. A. (2013). Evolution of acoustic and visual signals in Asian barbets. *J. Evol. Biol.* 26, 647–659. doi:10.1111/jeb.12084.

Hackett, S. J., Kimball, R. T., Reddy, S., Bowie, R. C., Braun, E. L., Braun, M. J., et al. (2008). A phylogenomic study of birds reveals their evolutionary history. *Science* 320, 1763–1768.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* 33, 1–22.

Hadfield, J. D. (2014). MCMCglmm Course Notes. Available at: https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf [Accessed February 13, 2019].

Hadfield, J. D., and Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* 23, 494–508.

Hall, M. L., Rittenbach, M. R. D., and Vehrencamp, S. L. (2015). Female song and vocal interactions with males in a neotropical wren. *Behav. Evol. Ecol.* 3, 12. doi:10.3389/fevo.2015.00012.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). GEIGER: investigating evolutionary radiations. *Bioinformatics* 24, 129–131.

Hofmann, C. M., Cronin, T. W., and Omland, K. E. (2008). Evolution of sexual dichromatism. 1. Convergent losses of elaborate female coloration in New World orioles (*Icterus* spp.). *The Auk* 125, 778–789.

del Hoyo, J., Elliott, A., and Christie, D. (2004–2011). *Handbook of the birds of the world Vol. 9–16*. Lynx Edicions.

Irwin, R. E. (1994). The evolution of plumage dichromatism in the New World blackbirds: social selection on female brightness. *Am. Nat.* 144, 890–907.

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., and Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature* 491, 444–448. doi:10.1038/nature11631.

Johnson, A. E., Price, J. J., and Pruett-Jones, S. (2013). Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecol. Evol.* 3, 3030–3046. doi:10.1002/ece3.686.

Kraaijeveld, K. (2014). Reversible trait loss: the genetic architecture of female ornaments. *Annu. Rev. Ecol. Evol. Syst.* 45, 159–177.

Kraaijeveld, K., Gregurke, J., Hall, C., Komdeur, J., Mulder, R. A., and others (2004). Mutual ornamentation, sexual selection, and social dominance in the black swan. *Behav. Ecol.* 15, 380–389.

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., et al. (2013). Behavioral evidence for song learning in the suboscine bellbirds (*Procnias* spp.; Cotingidae). *Wilson J. Ornithol.* 125, 1–14.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292–305. doi:10.2307/2407393.

Langmore, N. E., Davies, N. B., Hatchwell, B. J., and Hartley, I. R. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 141–146.

Martin, T. E., and Badyaev, A. V. (1996). Sexual dichromatism in birds: importance of nest predation and nest location for females versus males. *Evolution* 50, 2454–2460. doi:10.2307/2410712.

Mason, N. A., Shultz, A. J., and Burns, K. J. (2014). Elaborate visual and acoustic signals evolve independently in a large, phenotypically diverse radiation of songbirds. *Proc. R. Soc. B Biol. Sci.* 281, 20140967. doi:10.1098/rspb.2014.0967.

McGraw, K. J., Mackillop, E. A., Dale, J., and Hauber, M. E. (2002). Different colors reveal different information: how nutritional stress affects the expression of melanin- and structurally based ornamental plumage. *J. Exp. Biol.* 205, 3747–3755.

Møller, A. P., and Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behav. Ecol. Sociobiol.* 32, 167–176.

Morales, J., Gordo, O., Lobato, E., Ippi, S., Martínez-de la Puente, J., Tomás, G., et al. (2014). Female-female competition is influenced by forehead patch expression in pied flycatcher females. *Behav. Ecol. Sociobiol.* 68, 1195–1204. doi:10.1007/s00265-014-1730-y.

Murphy, T. G., Hernández-Muciño, D., Osorio-Beristain, M., Montgomerie, R., and Omland, K. E. (2009a). Carotenoid-based status signaling by females in the tropical streak-backed oriole. *Behav. Ecol.* 20, 1000–1006. doi:10.1093/beheco/arp089.

Murphy, T. G., Rosenthal, M. F., Montgomerie, R., and Tarvin, K. A. (2009b). Female American goldfinches use carotenoid-based bill coloration to signal status. *Behav. Ecol.* 20, 1348–1355. doi:10.1093/beheco/arp140.

Nowicki, S., Searcy, W., and Peters, S. (2002). Brain development, song learning and mate choice in birds: a review and experimental test of the" nutritional stress hypothesis". *J. Comp. Physiol. A* 188, 1003–1014.

Oberweger, K., and Goller, F. (2001). The metabolic cost of birdsong production. *J. Exp. Biol.* 204, 3379–3388.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., and Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5. doi:10.1038/ncomms4379.

Ornelas, J. F., González, C., and Espinosa De Los Monteros, A. (2009). Uncorrelated evolution between vocal and plumage coloration traits in the trogons: a comparative study. *J. Evol. Biol.* 22, 471–484. doi:10.1111/j.1420-9101.2008.01679.x.

Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. Biol. Sci.* 255, 37–45.

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Behav. Evol. Ecol.* 3, 40. doi:10.3389/fevo.2015.00040.

Price, J. J., and Eaton, M. D. (2014). Reconstructing the evolution of sexual dichromatism: current color diversity does not reflect past rates of male and female change. *Evolution* 68, 2026–2037. doi:10.1111/evo.12417.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things): phytools: R package. *Methods Ecol. Evol.* 3, 217–223. doi:10.1111/j.2041-210X.2011.00169.x.

Rogers, A. C., Langmore, N. E., and Mulder, R. A. (2007). Function of pair duets in the eastern whipbird: cooperative defense or sexual conflict? *Behav. Ecol.* 18, 182–188. doi:10.1093/beheco/arl070.

Rosvall, K. A. (2011). Intrasexual competition in females: evidence for sexual selection? *Behav. Ecol.*, arr106.

Schmidt, K. L., Moore, S. D., MacDougall-Shackleton, E. A., and MacDougall-Shackleton, S. A. (2013). Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Anim. Behav.* 86, 25–35.

Searcy, W. A., and Nowicki, S. (2005). *The evolution of animal communication: reliability and deception in signaling systems*. Princeton University Press Available at: https://books.google.co.nz/books?hl=en&lr=&id=fUFLKFbvt0gC&oi=fnd&pg=PR13&dq=Searcy,+W.+A.+%26+Nowicki,+S.+2005+The+evolution+of+animal+communication.+Princeton,+NJ:+Princeton+University+Press+&ots=qhgO0Mpd9Q&sig=agE2AzwKaQ_GYxj_J_vyjdDprkM [Accessed November 2, 2015].

Searcy, W. A., and Yasukawa, K. (1995). *Polygyny and sexual selection in red-winged blackbirds*. Princeton University Press.

Seddon, N., Tobias, J. A., Eaton, M., and Ödeen, A. (2010). Human vision can provide a valid proxy for avian perception of sexual dichromatism. *The Auk* 127, 283–292. doi:10.1525/auk.2009.09070.

Shutler, D., and Weatherhead, P. J. (1990). Targets of sexual selection: song and plumage of wood warblers. *Evolution* 44, 1967–1977. doi:10.2307/2409607.

Simpson, R. K., Johnson, M. A., and Murphy, T. G. (2015). Migration and the evolution of sexual dichromatism: evolutionary loss of female coloration with migration among wood-warblers. *Proc. R. Soc. Lond. B Biol. Sci.* 282, 20150375. doi:10.1098/rspb.2015.0375.

Smiseth, P. T., and Amundsen, T. (2000). Does female plumage coloration signal parental quality? A male removal experiment with the bluethroat (*Luscinia s. svecical*). *Behav. Ecol. Sociobiol.* 47, 205–212.

Templeton, C. N., Rivera-Cáceres, K. D., Mann, N. I., and Slater, P. J. B. (2011). Song duets function primarily as cooperative displays in pairs of happy wrens. *Anim. Behav.* 82, 1399–1407. doi:10.1016/j.anbehav.2011.09.024.

Tibbetts, E. A., and Dale, J. (2004). A socially enforced signal of quality in a paper wasp. *Nature* 432, 218–222.

Tibbetts, E. A., and Izzo, A. (2010). Social punishment of dishonest signalers caused by mismatch between signal and behavior. *Curr. Biol.* 20, 1637–1640.

Tobias, J. A., Montgomerie, R., and Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2274–2293. doi:10.1098/rstb.2011.0280.

Valcu, M., and Dale, J. (2013). *colorZapper: a platform for scoring biological coloration*. Available at: http://github.com/valcu/colorZapper.

Walther, B. A., and Clayton, D. H. (2005). Elaborate ornaments are costly to maintain: evidence for high maintenance handicaps. *Behav. Ecol.* 16, 89–95.

West-Eberhard, M. J. (1979). Sexual selection, social competition, and evolution. *Proc. Am. Philos. Soc.*, 222–234.

West-Eberhard, M. J. (1983). Sexual selection, social competition, and speciation. *Q. Rev. Biol.* 58, 155–183.

3 General Methods

The chapters that follow (Chapter 4 and Chapter 5) are based on an extensive database of New Zealand bellbird *Anthornis melanura* (hereafter 'bellbird') song collected from an archipelago metapopulation in northeastern New Zealand. The purpose of this chapter is to describe the rationale and methods by which I collected the data and created the database.

3.1 Syllables as memes

Bellbird song structure is highly flexible; the ways in which syllables are combined into songs varies both within and between individuals (see Appendix 7.2). Hence, I have used the smallest units of song—syllables—as the basic cultural unit (meme) for analysis. I defined syllables as a vocal sound from an individual (consisting of one or more 'elements') separated from other sounds by a gap of at least 15 ms (see Section 3.5). Analysis at the syllable level provides a robust and flexible approach where syllable diversity and sharing can be quantified, syllables can be grouped and examined at broader classification granularities (e.g. 'families' of syllables), and syllable associations and song syntax can be analysed.

I followed a typical bioacoustics workflow (Figure 3.1) to produce a database of recordings that have been partitioned into songs, which have in turn been segmented into their constituent syllables, which were then classified according to acoustic similarity. Below I describe each step of the process, beginning with an explanation of the study system.

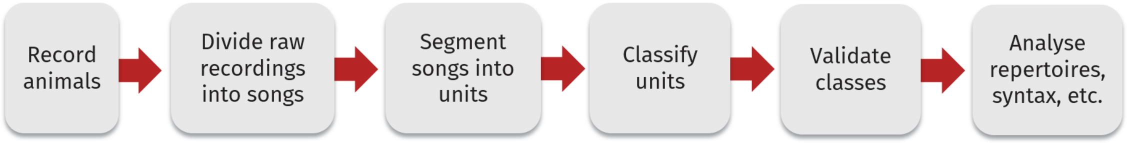


Figure 3.1—The bioacoustics workflow followed in this thesis.

I followed this workflow to create the bellbird song database, on which future chapters are based.

3.2 Basic experimental design

Bellbirds are present on most forested offshore islands (Heather and Robertson, 2000). For my study I chose six sites in the wider Hauraki Gulf region (i.e. from the Poor Knights Islands 35°28′06″S 174°44′13″E in the north to Tiritiri Matangi 36°36′01″S 174°53′24″E in the south, extending to Repanga 36°26′11″S 175°46′16″E in the east) based on their abundant bellbird populations (good return for sampling effort) and their variety in size, connectivity and bellbird population histories, allowing me to explore the effect/interplay of these factors on meme dynamics (Figure 3.2).

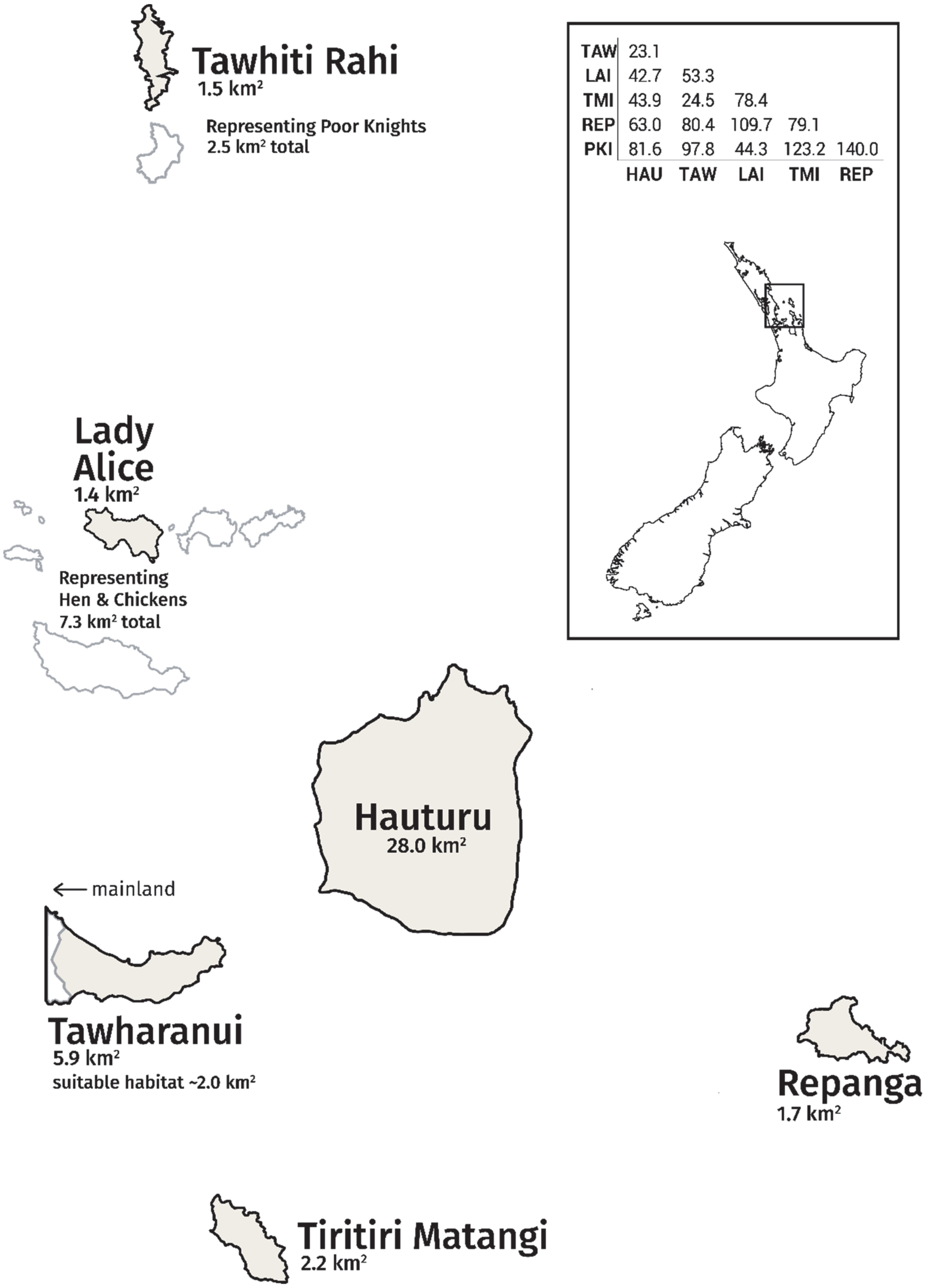
The Hauraki Gulf archipelago provides a unique opportunity. Compared to an open mainland situation, the islands afford more discrete populations with more controlled degrees of connectivity between them; the ocean forms a barrier that regulates the degree of migration. Previous work on bellbird gene-flow (Baillie, 2011; Baillie et al., 2014) has found ongoing dispersal on an ‘isolation-by-distance’ basis, with higher gene flow between closer sites.

If geographic distance regulates population connectivity within the Hauraki archipelago, it provides the opportunity to test whether more connected sites have more similar song (Chapter 5). Table 3.1 summarises basic geographic information about each of the 6 sites, including the dates of recording trips.

Table 3.1—Geographic and sampling information for the six study sites.

Tawhiti Rahi represents the larger Poor Knights Islands group; likewise, Lady Alice Island represents the larger Hen & Chickens Islands group. Elevation data from www.topomap.co.nz

Sites	GPS Coordinates	Area (km^2)	Elevation	Recording dates and number of recordists	Sample size
Tawhiti Rahi (representing Poor Knights Island group)	35°28’06”S 174°44’13”E	1.5 (2.5 including neighbouring island, Aorangi)	191 m (216 m for Aorangi)	Single trip 23/02/2017–27/02/2017 3 recordists	Birds: 128 M, 48 F Songs: 309 M, 61 F
Lady Alice (represents larger Hen and Chickens Islands group)	35°53’23”S 174°43’41”E	1.4 (7.3 including entire Hen and Chickens group	158 m (417 m for Hen)	Single trip 27/05/2015–30/05/2015 4 recordists	Birds: 149 M, 116 F Songs: 181 M, 126 F
Hauturu	36°11’57”S 175°04’53”E	28.0	722 m	Single trip 05/04/2016–09/04/2016 4 recordists	Birds: 57 M, 63 F Songs: 67 M, 98 F
Tawharanui	36°22’20”S 174°50’30”E	5.9	91 m	20/04/2016, 18/08/2016, 01/09/2016–07/09/2016, 22/09/2016 1–3 recordists	Birds: 66 M, 26 F Songs: 206 M, 37 F
Repanga	36°26’11”S 175°46’16”E	2.0	214 m	Single trip 03/11/2016 – 09/11/2016 5 recordists	Birds: 122 M, 141 F Songs: 328 M, 196 F
Tiritiri Matangi	36°36’01”S 174°53’24”E	2.2	80 m	Frequent recordings in non-winter months 2013–2015 1–2 recordists	Birds: 91 M, 58 F Songs: 389 M, 138 F



	HAU	TAW	LAI	TMI	REP
TAW	23.1				
LAI	42.7	53.3			
TMI	43.9	24.5	78.4		
REP	63.0	80.4	109.7	79.1	
PKI	81.6	97.8	44.3	123.2	140.0

Figure 3.2—Simplified map of the six study sites.

Tawhiti Rahi represents the larger Poor Knights Islands group; likewise, Lady Alice Island represents the larger Hen and Chickens Islands group. The other islands of these wider groups are shown in white. The distance matrix between sites (measured in km) is provided (HAU, Hauturu; TAW, Tawharanui; LAI, Lady Alice; TMI, Tiritiri Matangi; REP, Repanga, PKI, Tawhiti Rahi). Distances not to scale on simplified map.

Hauturu (Little Barrier Island)

Te Hauturu-o-Toi, or Little Barrier Island (hereafter Hauturu) is a large and rugged eroded volcano, central to the Hauraki Gulf. It was established as New Zealand's first nature reserve in 1896 (McLintock et al., 1966). Feral cats arrived and became established on the island after 1867, and together with Pacific rats *Rattus exulans* drove decline in several sea-bird, land-bird and reptile species, including the local extinction of North Island saddleback *Philesturnus carunculatus rufusater* (Girardet et al., 2001). By 1894, approximately a third of the island's forest had been burned or felled. Since becoming protected, forest regenerated, cats were eradicated by 1980 (Veitch, 2001), saddleback were re-introduced in 1986 (Lovegrove, 1996), and rats eliminated by 2004 (Bellingham et al., 2010). Today Hauturu can be considered one of New Zealand's most intact ecosystems (Girardet et al., 2001). Hauturu supports an ancient, stable population of bellbirds (Craig and Douglas, 1984), and given its centrality to the Hauraki Gulf archipelago, represents a likely source of historic and ongoing bellbird dispersal to surrounding sites. Genetic sampling of bellbirds on Hauturu did not detect first-generation migrants (i.e. birds born outside of Hauturu), but gene frequencies of migrants and natives may well have been too similar to discern between them, rather than reflecting a true lack of immigration (Baillie, 2011 p. 58).

Tawharanui Regional Park

Tawharanui Regional Park (hereafter Tawharanui) consists of a patchwork of native forest and pastureland on a mainland peninsula jutting into the Hauraki Gulf, 23 km west of Hauturu. Bellbirds disappeared from Tawharanui—as well as virtually all the North Island mainland north of Hamilton (37°45' S)—in the mid-1800s, likely due to deforestation, introduced mammalian predators, and disease (Craig and

Douglas, 1984; Bartle and Sagar, 1987; Duncan and Blackburn, 2004). Following pest eradication and the implementation of a predator-proof fence in 2004, Tawharanui was recolonised by a large flock of bellbirds (*c*. 100 individuals), first seen in February 2005 (Brunton et al., 2008). The flock consisted of both young and adult birds, both males and females (Baillie, 2011), and the source population was determined to be primarily Hauturu, based on proximity, matching song dialect (Brunton et al., 2008) and genetics (Baillie, 2011). The Tawharanui bellbird population shows no evidence of a genetic bottleneck (Baillie, 2011). Since colonization the bellbird population at this site has rapidly increased, and today appears to be at or approaching carrying capacity (Ornithological Society of New Zealand OSNZ, unpublished data).

Tiritiri Matangi

Tiritiri Matangi Island (hereafter Tiri) is a low-lying island 4 km east of the Whangaparoa peninsula. It was deforested over many centuries of Maori and European occupation and farming, leaving only 6% native vegetation cover by 1984 (Graham and Veitch, 2002; Anderson and Craig, 2003). Farming ceased in 1971 and intensive restoration 1984–1993 included the planting of *c*. 280,000 native trees, raising native vegetation cover to 60% (Galbraith and Hayson, 1995), with an additional 20,000 trees planted since (Graham and Veitch, 2002). A number of endangered bird and reptile species have been translocated to the island, and today Tiri is an ‘open sanctuary’ for native fauna, upheld as a model of community participation and conservation advocacy (Galbraith and Cooper, 2013).

Prior to restoration efforts, there were only *c*. 150 bellbirds on the island in 1977–1978 (Anderson and Craig, 2003) and possibly even fewer beforehand (pers. comm.

Dianne Brunton). Since the replanting/regeneration of forest, eradication of mammalian predators (Pacific rats), and introduction of supplementary sugar-water feeders (Roper, 2012; Galbraith and Cooper, 2013), the bellbird population expanded to approximately 1200 individuals as of 2010, as a conservative estimate (Roper, 2012), and now appears to be at carrying capacity (Roper, 2018). Genetically, the population is distinct and non-diverse, compared to other populations in our sample (Baillie, 2011).

Repanga (Cuvier Island)

Repanga is a rugged, cliff-dominated island 23 km off the eastern coast of the Coromandel Peninsula. It is mostly forested, with grassy clearings on the eastern slopes which are dotted with pohutukawa *Metrosideros excelsa*. Bellbirds are likely the most abundant passerine on Repanga (Bellingham et al., 1981), and have remained present on the island through ecological degradation in the early 1800s to 1960s, as described below.

In the early 1800s, Maori periodically inhabited Repanga, and Pacific rats were present (Merton, 1972). In the late 1880s, the island became a lighthouse reserve and a quarter was farmed. By 1957, wild goats and wandering stock had destroyed the forest understorey, and predation by cats had driven the local extinction of North Island saddleback, tui *Prosthemadera novaeseelandiae*, and red-crowned parakeet *Cyanoramphus novaezelandiae* (Merton, 1972; Bellingham et al., 1981). Restoration began in 1961 with the removal of wild goats and excess stock, eradication of cats in 1964, and rats in 1993 (Towns et al., 1995). The understorey rapidly regenerated, and saddleback and red-crowned parakeets were successfully re-introduced to Repanga in 1968 and 1974, respectively (Bellingham et al., 1981;

Veitch and Newton, 1989). Today the island is a nature reserve managed for scientific and species-protection purposes.

Repanga lies 63 km southeast of Hauturu. A direct path from Hauturu to Repanga passes through the Colville Channel between Great Barrier Island (Aotea) and the Coromandel Peninsula. It is possible that a bellbird flying from Hauturu might stop *en route* on the Coromandel, as a 'stepping stone' assisting the journey to Repanga. Alternatively, if bellbirds encounter suitable habitat in the Coromandel (which holds an established bellbird population) they may settle there, rather than continuing on to Repanga. It is therefore unclear to what extent the Coromandel promotes or obstructs bellbird movement to Repanga from Hauturu.

Lady Alice Island (Hen and Chickens Islands group)

Lady Alice Island (Motu Muka) forms part of the Hen and Chickens Islands group, a nature reserve 10 km from the Whangarei Heads at the closest point. The Hen and Chickens consists of seven islands; Hen (Taranga) Island is the largest at 4.7 km^2, and the Chickens (Marotere Islands) comprise a linear chain of six much smaller, closely-spaced islands about 6 km to the north: Wareware and Muriwhenua—small outcrops totalling 3 ha; Mauitaha—a rugged 20 ha scrub-covered rock; Lady Alice Island (140 ha); Whatupuke (100 ha); Coppermine Island (75 ha).

All the islands were occupied by Maori, probably continuously on Taranga and seasonally on the smaller islands (Hayward et al., 1978; Prickett, 1984). Regular use of the islands by Maori likely ended in 1821 (Cranwell and Moore, 1935).

Hen was likely mostly deforested for gardening, prior to Maori departure (Court, 1978). Its forest consists of a complex mosaic of "mixed broad-leaved, hardwood (including extensive *Beilschmeidia tarairi*), and coastal forest regenerating through kanuka *Kunzia ericoides* and pohutukawa" (Towns et al., 2007).

Lady Alice Island experienced periodic localised disturbance from burning until 1903 (Bellingham, 1984), and had free-ranging cattle from 1890 until removal in 1924 (Hayward and McCallum, 1984). Today the island is covered with coastal forest with kohekohe *Dysoxylum spectabile* regenerating through kanuka and pohutukawa (Towns et al., 2007).

Coppermine Island has a history of fires, prospecting and mining (Atkinson, 1968). It retains pockets of mature forest, but largely consists of kohekohe regenerating through extensive kanuka and pohutukawa (Towns et al., 2007).

Whatupuke Island is the least disturbed of the Chickens and contains mature coastal forest with kohekohe, some areas regenerating through kanuka and pohutukawa (Towns et al., 2007).

Pacific rats, present since the early 1800s (at least on Lady Alice Island; Brook, 1999), are the only mammalian predators to have existed on the islands, and were eradicated from Whatupuke in 1993, Lady Alice Island in 1994, Coppermine in 1997 (Towns et al., 2007), and Taranga in 2011 (Department of Conservation, 2013).

Bellbirds have apparently been present continuously on the Hen and Chickens.

Tawhiti Rahi (Poor Knights Islands group)

Tawhiti Rahi (1.5 km^2) and Aorangi (1.0 km^2) together comprise the Poor Knights Islands group, the northernmost site in my study. The islands are rugged and cliff-bound, covered with mostly pohutukawa forest, with some broadleaf forest remnants and a few grassy clearings (Stringer et al., 2003). Both islands support great densities of Buller's shearwaters *Puffinus bulleri* (Waugh et al., 2013).

The Poor Knights were intermittently inhabited by Ngatiwai Maori from *c.* 1500, who used the islands for vegetable gardens and as a base for harvesting seafood (Fraser, 1925; Hayward, 1993). After a massacre by a raiding party from the mainland in 1820, the islands were declared 'tapu' (off-limits) and have been uninhabited since (Fraser, 1925; Taylor et al., 2011). Pigs *Sus scrofa* remained on the island and proliferated, taking a heavy toll on burrowing seabirds and tuatara (Buddle, 1941). Pigs were eradicated from the islands in 1936 (Bellingham et al., 2010) and the islands left to recover. Reserve status was conferred in 1975 and full protection in 1998 (Taylor et al., 2011), with access now restricted to scientific studies.

Relative to other sites in my sample, the Poor Knights are geographically isolated; the nearest island is Lady Alice Island in the Hen and Chickens group, 44 km to the south. The isolation of this population is evidenced by high genetic distinctness and a lack of first-generation migrants detected (Baillie, 2011).

The Poor Knights bellbird population survived the mainland extirpation of the 1800s (Craig and Douglas, 1984), but shows strong evidence of a population bottleneck (Baillie, 2011). I surmise that the bottleneck likely resulted from drought, as fresh water on the islands is scarce and the climate harsh (Buddle, 1941; Bartle and Sagar, 1987).

3.3 Field methods: Audio recording

3.3.1 Trip logistics

The fieldwork undergirding this project began with Michelle Roper's doctorate work on Tiri (2013–2015; Roper, 2018). After commencement of my own doctorate this expanded to include the other sites of the archipelago (2015–2017). The work involved multi-day recording visits to 6 island and mainland sites in the Hauraki Gulf, northeastern New Zealand, and exploratory trips to several others. Many of the islands of the Hauraki Gulf are scientific reserves with strictly controlled access. I obtained landing and research permits through consultation with iwi (Ngati Manuhiri, Ngati Wai), the New Zealand Department of Conservation (47948-FAU, 34833–FAU, 41756-FAU, 48000-FAU), and Massey University Animal Ethics Committee (15/21) with permission to record song and undertake mist-netting and blood-sampling of bellbirds. My final sample consisted of 6 sites: Tawhiti Rahi, Lady Alice, Hauturu, Tawharanui, Repanga and Tiri. I applied for and conducted exploratory trips to other sites—Whangarei Heads, Motutapu, Shakespear Regional Park, Coromandel Forest Park, and Hunua Ranges Regional Park—but determined that bellbird density was too low for efficient sampling. Additionally, a boat was chartered and full day of recording with five recordists conducted on the small (50 ha) naturally-regenerating island of Pokohinu (situated in the Mokohinau Islands group, 35°54'18"S, 175°06'48"E). Although female song was prolific on this day, males were silent. I thus excluded this island from my final sample.

I was responsible for recruiting volunteers, organising gear, transport, food and shelter for recording trips. Given the protected status of many sites, pre-departure quarantine procedures were stringent and involved cleaning gear with TriGene disinfectant (Ethical Agents International Ltd), removing all traces of biotic matter from gear and clothing, abiding by food restrictions, and sealing gear in airtight containers until arrival.

Some sites contained sensitive *tapu* areas of significance to Maori which restricted where we could record. This was especially true of Hauturu; ascending the island peak beyond the designated tracks was not permitted, which constrained our sampling to the southwest quadrant of the island. To increase spatial coverage, the island ranger took me and colleague Michelle Roper by boat to the northeastern coast of the island, Pohutukawa Flats, to sample for a day. Distinctly different bellbird song-types were noted here, but vegetation was so tall (and bellbird activity so far above the ground) that recording opportunities were limited.

3.3.2 Recording bellbird song

The approximate total number of recordist-hours spent actively tracking and recording bellbirds at each site was as follows. Tawhiti Rahi, 108; Lady Alice, 110; Hauturu, 140; Tawharanui, 132; Repanga, 260; Tiri, many hundreds of hours (mostly by Michelle Roper), plus ~100 hours at other exploratory sites not included in the final sample. In total then, well over 1000 hours of active recording effort were spent. The final database consisted of 2137 good quality song recordings of adult birds. For each of the 6 sites in the final sample, an average of 2–4 useable songs were recorded for every hour of effort. See Table 3.1 for sample sizes of songs and individuals for each site.

My aim was to sample bellbird song at each site, as systematically and representatively as possible. Recordists coordinated their sampling to maximise spatial coverage of the sites and minimise pseudo-replication by searching and recording in separate, designated areas at each site. Population sizes were large, making pseudo-replication unlikely.

Recordings were made between dawn and dusk with Marantz PMD661 portable solid-state recorders paired with Sennheiser ME-66 shotgun (directional) microphones. Recordists used a sampling rate of 48 kHz and 24-bit depth, using ‘dual mono’ mode. This mode records to two channels simultaneously at different volume levels; the quieter channel provides a backup in case of clipping distortion on the primary channel.

Recordists moved quietly, searching for adult bellbirds. Adulthood and sex of the focal bird was identified based on size and plumage, and individual birds were typically followed for up to 10 minutes. Once a recording was collected the recordist focussed on a different bird or relocated to find new birds. Playback was not used to elicit songs in any of the songs in the final sample. The recordist maximised signal-to-noise ratio by moving as close to the focal bird as possible, angling the microphone to minimise background noise (from non-focal birds, wind, streams, ocean, vehicles, etc). Information about focal bird and environment were dictated softly into the microphone during lulls in singing activity. This information included age (juvenile or adult), sex and behaviour of the bird, specific location, social context, vegetation type, and weather conditions. Locations of each recording were logged with a Garmin GPS unit or approximated using *Google Earth Pro* with track maps digitally overlaid on the base map.

3.3.3 File naming and storage

After each trip, recordings were transferred from SD cards to a shared drive and renamed with a systematic naming structure, using the software *Bulk Rename Utility* (www.bulkrenameutility.co.uk) as shown in Figure 3.3.

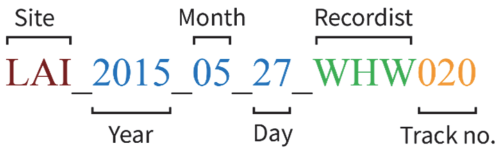


Figure 3.3—Naming scheme for recordings.

This example is the twentieth track recorded by recordist WHW on 27th May 2015, Lady Alice Island (LAI).

3.4 Transcribing recordings and partitioning into songs

Each recording was opened in *Raven Pro* version 1.4 or 1.5 beta (The Cornell Lab of Ornithology, 2014); I listened to the track while viewing the spectrogram, transcribing all relevant track metadata into a transcript spreadsheet in *Microsoft Excel* (Office 365). At the same time, I partitioned the track into 'songs', defined as vocalisation bouts by an individual, separated by a gap of >3 seconds. There are no universally applicable criteria for delineating units; criteria tend to be species-specific (See Kershenbaum et al., 2016). I chose the value of 3 seconds by trial and error; long male songs often have long gaps between syllables. Three seconds was long enough to avoid every syllable being considered a separate song, but short enough to avoid multiple bouts being included as one song. For each song selection, I transcribed song metadata into a selection table in *Raven Pro*; the song selection table was then exported as a text file. Track metadata and Song metadata are described below.

3.4.1 Track metadata

Track metadata relate to the entire recording and consist of the fields in Table 3.2.

Table 3.2—Track metadata fields and their descriptions.

Metadata field	Description
site	The site of the recording: Hauturu, Tawharanui, Lady Alice Island, Tiri, Repanga, Tawhiti Rahi
recordist	Initials of recordist, e.g. WHW
date	The date the recording was made, in format yyyy_mm_dd.
track	The name of the WAV file
time	The time the recording was made, e.g. 13:52
nest	ID of nest if recording birds on/in a nest
stimulus	Is song playback on the recording? Y [yes] / N [no] (Used to filter for playback experiments).
behaviours	Comments specifically about the behaviour of the bird
specific_locality	Specific locality of the recording, e.g. using track markers, landmarks… (free format)
GPS	GPS coordinates (degrees, minutes, seconds) , extracted from Garmin GPS unit or estimated from Google Earth
notes	Free format field for incorporating any information from the commentary or observations of the recording. For example, phrases like "backlighting made gender ID difficult"... "approximately 10 BB in mixed-gender flock feeding in tall kanuka"... "Long sequence (8 sels) of one unbanded M singing tight loop of 5 syllables" … etc.

3.4.2 Song selection metadata

Song selection metadata relate specifically to the song selection. A song selection table is created in *Raven* as you make selections of songs on the spectrogram. The purpose of the song selection table is to reference where each song occurs within the raw recording, and to note song quality, ID of singing individual, and social/behavioural/acoustic context of the songs. Song metadata consists of the fields in Table 3.3.

Table 3.3—Song selection metadata fields and their descriptions.

Annotation columns are marked with ☐, and with an additional * if the field is essential to fill out.

Metadata field	Description
Selection	The song selection number in *Raven* (selections are numbered chronologically)
View	Whether the row of data relates to the waveform view or the spectrogram view in *Raven*.
Channel	Which channel the audio was recorded on (1 or 2).
Begin File	The name of the track WAV file on which the song selections have been made.
Begin Time (s)	The time at which the selection begins, in seconds.
End Time (s)	The time at which the selection ends, in seconds.
Low Freq (Hz)	The floor of the selection box drawn in *Raven*.
High Freq (Hz)	The ceiling of the selection box drawn in *Raven*.
sex ☐*	The sex of the singing individual. **M** [male] **F** [female] **U** [unknown] **MP** [probably male] **FP** [probably female]
quality ☐*	Quality of the song selection. **B** [bad; you would only make such a selection if there was some interesting aspect to document] **OK** [moderate quality; not fit for detailed audio analysis but you can make out the song/syllable types] **G** [good; clear but low volume/high background noise] **VG** [very good; clear, close and suitable for detailed audio analysis] **EX** [excellent; close, studio-quality recording with excellent signal-to-noise ratio]
bird_ID ☐*	Unique identifier of the singing bird. The ID functions to group songs from the same bird together for analyses. If a bird produces multiple song selections, label the first-

	occurring song selection from that bird as the audio track name, plus a letter. For example, if you are working on audio track LAI_2015_05_30_DHB001, label the bird's first song selection as LAI_2015_05_30_DHB001A. For every subsequent selection by the same bird, whether within the same recording or in subsequent recordings, enter this same ID. For the next bird in DHB001, label selections as LAI_2015_05_30_DHB001B, and so on.
comments ☐*	Very concise comment with no spaces (<20 characters), as this field is automatically incorporated into the exported song selection filename. Used as a visual marker when browsing in File Explorer. Examples: PossMultiBirds [possibly multiple birds in the selection], ResponseToPlaybk [response to playback], MChasingF [male chasing female], AlarmMobbingMorepork [alarm calling while mobbing morepork]...
calling_bird_seen ☐	Y [yes] / N [no]. Used as an additional measure of certainty of sex / ID. If the calling bird was seen, then sex / ID assessment is more likely to be accurate.
free_format_comments ☐*	Any important comments about the selection. What happens during the selection? What was the social context of the selection? Any remarks on syntax? On gender certainty? Expand on the comments.song comment if applicable.
species_if_not_bb ☐	Fill with the species name for songs from non-bellbird species. If bellbird, leave blank.

3.4.3 Song naming scheme for WAV exports

I gave *Raven*-exported WAV files the naming scheme illustrated in Figure 3.4, constructed from selection table fields.

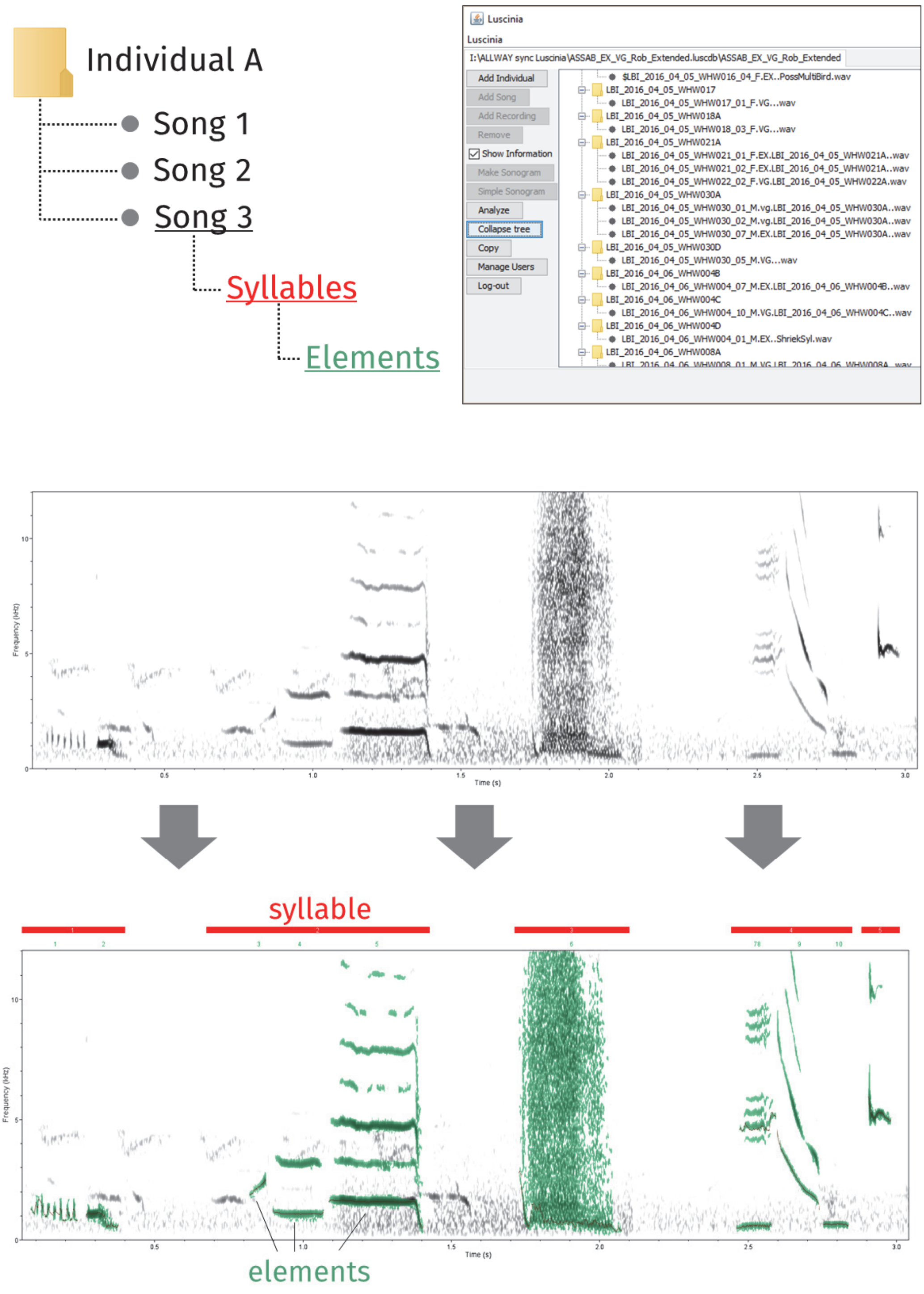
Figure 3.4—Selected song naming scheme for *Raven* WAV file exports.

This is an example song filename as exported from *Raven*. The source track is LAI_2015_05_27_WHW020.wav. The song is the third selection (03) on the source track and was sung by a male (M). The song selection is of Very Good (VG) quality. The ID of the singing bird is LAI_2015_05_27_WHW020B. The comment ‘CntrSinging’ indicates the bird was counter-singing with a background bird.

3.5 Segmenting and measuring syllables

To segment song recordings into their constituent acoustic units, I used the bioacoustics software *Luscinia* (https://github.com/rflachlan/Luscinia/wiki).

A *Luscinia* database consists of hierarchically nested information. At the highest level are *individuals* (birds), which own *songs* (vocalisation bouts), which are comprised of *syllables* (acoustic units), which in turn are comprised of *elements* (uninterrupted traces on the spectrogram).



```
I:\ALLWAY sync Luscinia\ASSAB_EX_VG_Rob_Extended.luscdb\ASSAB_EX_VG_Rob_Extended
    ● $LBI_2016_04_05_WHW016_04_F.EX..PossMultiBird.wav
⊟ LBI_2016_04_05_WHW017
    ● LBI_2016_04_05_WHW017_01_F.VG...wav
⊟ LBI_2016_04_05_WHW018A
    ● LBI_2016_04_05_WHW018_03_F.VG...wav
⊟ LBI_2016_04_05_WHW021A
    ● LBI_2016_04_05_WHW021_01_F.EX.LBI_2016_04_05_WHW021A..wav
    ● LBI_2016_04_05_WHW021_02_F.EX.LBI_2016_04_05_WHW021A..wav
    ● LBI_2016_04_05_WHW022_02_F.VG.LBI_2016_04_05_WHW022A.wav
⊟ LBI_2016_04_05_WHW030A
    ● LBI_2016_04_05_WHW030_01_M.vg.LBI_2016_04_05_WHW030A..wav
    ● LBI_2016_04_05_WHW030_02_M.vg.LBI_2016_04_05_WHW030A..wav
    ● LBI_2016_04_05_WHW030_07_M.EX.LBI_2016_04_05_WHW030A..wav
⊟ LBI_2016_04_05_WHW030D
    ● LBI_2016_04_05_WHW030_05_M.VG...wav
⊟ LBI_2016_04_06_WHW004B
    ● LBI_2016_04_06_WHW004_07_M.EX.LBI_2016_04_06_WHW004B..wav
⊟ LBI_2016_04_06_WHW004C
    ● LBI_2016_04_06_WHW004_10_M.VG.LBI_2016_04_06_WHW004C..wav
⊟ LBI_2016_04_06_WHW004D
    ● LBI_2016_04_06_WHW004_01_M.EX..ShriekSyl.wav
⊟ LBI_2016_04_06_WHW008A
```

Figure 3.5—Hierarchical structure of a *Luscinia* database (top) and demonstration of the segmentation process (middle-bottom).

A *Luscinia* database contains hierarchically-nested information: elements (and their measurements) within syllables, within songs, within individuals. On a song spectrogram, the user 'paints over' elements to segment them. Selected areas of the spectrogram appear green, so the user can inspect whether the desired areas of signal are included, and erase regions of accidentally-selected background noise.

Once the user has added individuals and their songs to the database, spectrograms are generated for each song. The user marks the elements on the spectrogram by ‘painting a blob’ over each element to highlight the signal. As a blob is painted, *Luscinia* adds the element to the database, including start/end times and spectral measurements (e.g. fundamental frequency, entropy) from the painted area. The user then groups elements together into syllables by marking syllable start/endpoints on the spectrogram. The spectrogram and selection settings I employed are given in Table 3.4 below.

Table 3.4—*Luscinia* spectrogram and segmentation settings.

These are the settings I applied in *Luscinia* for producing spectrograms and segmenting units. The first eight settings relate to the appearance of the spectrogram. The remaining five have to do with controlling how elements are segmented when painted over by the brush tool.

Parameter	Settings
Windowing function	Gaussian
Max frequency	12,000 Hz
Frame Length	10.667 ms
Time step	1
Spectrograph points	512
Spectrogram overlap (%)	90.62
High-pass threshold	300 Hz
Time zoom	100%
Dynamic range	Settings adjusted dynamically for clean selections
Brush size	
Dereverberation	
Fundamental Frequency jump suppression	
Fundamental Frequency bias	

Additionally, the ‘View fundamental frequency’ setting was applied to display the fundamental frequency trace in painted elements on the spectrogram, to help ensure that fundamental frequency measurements were accurate and free from artefacts.

In this way I segmented approximately 21,500 syllables in *Luscinia*, over a period of approximately three months. I grouped elements together as a syllable if they were separated by less than 15 ms of silence. For the most part, bellbirds sing in discrete units that are easy to demarcate as syllables based on a substantial silent gap either side. Occasionally there are fast-paced bursts of sounds where it is not clear where one unit ends and another begins; in these cases the threshold of 15 ms provided a consistent rule for demarcating closely-spaced syllables. This value was chosen through trial and error for the best consistency of syllable groupings.

3.6 Classifying syllables

Classifying acoustic units involves both identifying unit classes (i.e. producing a 'catalogue' of classes) and assigning units to those classes. Classification is a key step in studying repertoire size and sequence structure in animal communication, and to date relies primarily on manual methods involving human visual and auditory perception (Kershenbaum et al., 2016). Despite the recognition of unit classification as a key step, I could find no tools suitable for facilitating classification of our large (21,500-syllable), diverse dataset. Existing acoustics software offered visualisation and labelling functionality only within one or few files at a time, not on the large comparative scale required. Accordingly, I collaborated with computer scientist Yukio Fukuzawa, a fellow PhD candidate at Massey University, to develop new software—*Koe*—to meet the need for expedited, large-scale classification. For details and justification of the program workflow and methodology, see Chapter 4.

A *Koe* database was created, and a custom script used to import songs and syllable segmentation from the *Luscinia* database. The syllables were then available in *Koe* as an interactive unit table, with a syllable per row containing spectrograms, audio and syllable metadata. In *Koe* I extracted acoustic features from each syllable to produce a similarity index (UPGMA method; Sokal, 1958), which allowed syllables to be ranked in the table by acoustic similarity. As *Koe* was developed in parallel to this classification task, the technical details of feature extraction morphed over time. See *Koe* documentation (https://github.com/fzyukio/koe/wiki) for the present list of extracted features and details of similarity index calculation.

In *Koe*'s unit table view, I sorted the unit table by similarity and compared the syllables visually and aurally, at varying playback speeds, to discern fine-scale classes and their boundaries. Groups of syllables I deemed to be the same class (see below) were assigned a memorable class label, named according to my subjective impression of contour, timbre and pitch. *Koe*'s 'View exemplars' view was used on a second monitor to keep track of existing syllable classes, as a reference catalogue. Periodically, I sorted the unit table alphabetically by label to see the syllables constituting each class. I identified nonconforming syllables and relabelled them to maintain class consistency. Classification was a highly iterative process, with groups frequently split and merged until within-class consistency and between-class distinctness were satisfactory (see below).

3.6.1 Defining classes

Fine-scale syllable classes were identified based on 'just-noticeable differences' in pitch, timbre, and duration of the playback and visual appearance of the spectrogram. In other words, if I could detect a visual-acoustic difference compared to existing classes, I allocated a new label type to that syllable. After all syllables had been classified at the fine-scale, syllable classes were nested into broad-scale ‘families’ based on salient features of contour and/or timbre. For example, classes characterised by squeaky timbre and upward contour form the *Upsqueak* family; flat-contour pure-tone classes form the *Pipe* family; high-entropy (i.e. like white noise) classes form the *Cough* family, and so on.

The rationale for this approach to classification is as follows. First, human perception is known to be excellent at the ‘multivariate processing’ task of distinguishing acoustic differences between units (Sayigh et al., 2007; Duda et al., 2012). As a trained musician, my aural skills were well-suited for detailed classification, and consistency was maximised by classifying the entire dataset myself.

Second, one cannot know *a priori* how much acoustic difference is enough to constitute a new meme type from the perspective of the bird. Bird hearing is believed to be at least as acute as human hearing for most species, both in sensitivity to pitch and temporal properties (Dooling, 2004). Therefore, if I can hear a difference between two sounds, it is reasonable to assume that a bellbird can too. Of course, how bellbirds mentally categorise the sounds is a separate question to whether they can discern acoustic differences. Because I could not make assumptions about the perceptual categories of bellbirds, I classified at multiple levels of label granularity as described above: at a fine scale (possibly near the limits of the bird’s discrimination) and at the broader scale of class families, enabling analyses at different levels and increasing the robustness of results.

Third, perception-based classification proved more practical than other potential classification criteria. Clustering—automatically defining classes based on statistical separation in acoustic feature-space—is one such alternative. I initially tried clustering using multi-dimensional scaling (MDS) and principal component analysis (PCA) to identify classes, but these failed to distinguish syllables that were (to human perception) very distinct. I reasoned that, given basic similarities in sound processing and cognition between humans and birds (Dooling, 2004), human perception would likely approximate bird perception more closely than biologically-ungrounded statistical clustering based on an arbitrary thresholds.

Another alternative classification method is to use syntactical context to assist with defining classes; that is, labelling syllables based on which syllables precede and follow in a song. I rejected this approach for two reasons: firstly, because the syntactical context could bias classification towards the 'typical' sequence. For example, if syllable A is usually followed by syllable B, a syllable following A is likely to be subconsciously shoehorned into the B class even if it sounds quite different. Secondly (and crucially), syntax-assisted classification would bias the results of syntax analysis. The syntax of songs is based on their constituent units, so the units themselves cannot be defined based on syntax without a degree of circular reasoning. Classifying syllables based on acoustic properties alone removes this possibility of bias.

One final rationale for our classification approach: having classified at a fine scale, it is relatively straightforward to merge similar types together into broader groupings if desired—easier than splitting classes that have been defined too broadly (see

Chapter 6). The chosen approach thus provides the most versatility for future work on this dataset.

Special case: *Pipe* syllable classification

Bellbirds are named after the pure-tone, bell-like notes that are so characteristic of their song (Andersen, 1926). On the spectrogram they appear as a narrow straight horizontal lines (—) reminiscent of a pipe, and so I named this family of syllables *Pipe*.

Unlike other syllable classes, *Pipe*-family syllables (hereafter ‘pipes’) exhibit a wide range of pitches (i.e. fundamental frequencies) and durations, giving an overall impression of continual gradation rather than discrete categories. Yet bellbirds often sing pipes in sequences of highly specified melodic contour, keeping consistent musical intervals between pipes each time they sing the sequence (Andersen, 1926). In fact, site-specific melodic sequences of pipes are to my ears one of the most noticeable features distinguishing the bellbird song profile of each site. In other words, although the *Pipe* family as a whole may show continual gradation, within songs the pipes appear to function categorically, like notes making up a musical phrase.

Because of the musical nature of pipe sequences, I chose to classify pipe syllables with a musical notation based on the nearest note on the piano keyboard, according to the naming scheme in Figure 3.6. To find the closest match I aurally compared playback of each pipe syllable in *Koe* to an online sine wave tone generator (http://www.szynalski.com/tone-generator/).

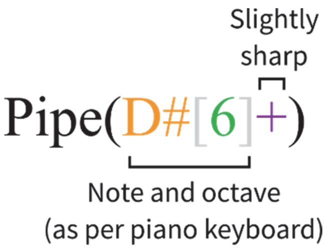


Figure 3.6—Naming scheme for *Pipe* syllables.

Pipes were named based on nearest note on the piano keyboard, consisting of note name (in this case D#) and octave (in this case, octave 6; i.e. the 6th D# from the left end of the standard piano keyboard). If a pipe fell in-between two notes on the piano keyboard, it was named according to the lower of the two, with a plus symbol (+) to indicate 'sharpness' (slightly higher pitch).

Pitch, the perceived highness or lowness of a sound, does not scale linearly with frequency (Hz) but logarithmically. Musical notation (that is, naming according to notes on a piano keyboard) takes this into account and provides a convenient method for binning pipes into perceptually-appropriate categories based on pitch. It has the additional benefit of assisting future investigations of bellbird song *musicality* (Rothenberg et al., 2014); that is, the idea that song potency may depend on its melodic contour affecting the listener, in an analogous way to how human music affects our emotions. I decided not to classify pipes based on duration, as duration may be artificially lengthened by echo in a forest environment and is thus a less reliable feature.

The bellbird dawn chorus, which seems to be performed only by males, is typically sung in a continuous looping sequence of 3–7 syllables (depending on the site). Pipes are often a prominent part of this dawn looping song.

A description from 1843 recounts:

> *"In the morning, I woke early; and, as the dawn first peeped forth, was deafened by the sound of the bell-birds. The woods which were close by seemed to be thronged with them. Never before had I heard so loud a chorus* [...] *They commence at dawn of day their chime of four notes which, repeated independently by a thousand throats, create the strangest melody." (Shortland, 1851)*.

After dawn, additional syllable classes are incorporated into male song, and females begin singing. Towards dusk, again males sing the dawn-pattern cyclical song. The function of dawn and dusk song, why it typically includes pipes, and why it is sung only by males, are fascinating avenues for future research.

Most of the recordings in the present study are from daylight hours, and so for consistency of analysis I have excluded dawn and dusk song from the database.

3.7 Validating classification

To validate my classification, I examined inter-observer reliability in a labelling experiment using 74 judges (Massey University Ornithology students). I found an average label match of 89.6% (median 95.6%), which is a high degree of agreement for inter-observer reliability studies (Nelson et al., 2004; Parker et al., 2012), lending validity to my manual classification. As this was an untrained and comparatively unskilled cohort, validation done with experienced analysts would probably have given even higher matching. See Chapter 4 (Section 4.4.1) for more details.

3.8 References

Andersen, J. C. (1926). *Bird-song and New Zealand song birds*. Auckland [N.Z.] : Whitcombe & Tombs Available at: https://trove.nla.gov.au/version/32884821 [Accessed July 31, 2018].

Anderson, S. H., and Craig, J. L. (2003). Breeding biology of bellbirds (*Anthornis melanura*) on Tiritiri Matangi Island. *Notornis* 50, 75–82.

Atkinson, I. A. E. (1968). An ecological reconnaissance of coppermine island, Hen and Chickens group. *N. Z. J. Bot.* 6, 285–294. doi:10.1080/0028825X.1968.10428814.

Baillie, S. M. (2011). Population genetics, biogeography and ecological interactions of the New Zealand bellbird (*Anthornis melanura*) and their avian malaria parasites: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph. D.) in Zoology at Massey University, Albany, New Zealand. Available at: http://muir.massey.ac.nz/handle/10179/2863 [Accessed July 18, 2014].

Baillie, S. M., Ritchie, P. A., and Brunton, D. H. (2014). Population genetic connectivity of an endemic New Zealand passerine after large-scale local extirpations: a model of re-colonization potential. *Ibis* 156, 826–839. doi:10.1111/ibi.12182.

Bartle, J. A., and Sagar, P. M. (1987). Intraspecific variation in the New Zealand bellbird *Anthornis melanura*. *Notornis* 34, 253–306.

Bellingham, P. (1984). Forest regeneration on Lady Alice Island, Hen and Chickens Group. *Tane* 30, 31–42.

Bellingham, P. J., Lovegrove, T. G., and Pitt, S. E. (1981). Birds of Cuvier Island. *Tane* 27, 23–32.

Bellingham, P. J., Towns, D. R., Cameron, E. K., Davis, J. J., Wardle, D. A., Wilmshurst, J. M., et al. (2010). New Zealand island restoration: seabirds, predators, and the importance of history. *N. Z. J. Ecol.* 34, 115.

Brook, F. J. (1999). Changes in the landsnail fauna of Lady Alice Island, northeastern New Zealand. *J. R. Soc. N. Z.* 29, 135–157. doi:10.1080/03014223.1999.9517588.

Brunton, D. H., Evans, B. A., and Ji, W. (2008). Assessing natural dispersal of New Zealand bellbirds using song type and song playbacks. *N. Z. J. Ecol.* 32, 147–154.

Buddle, M. G. A. (1941). Birds of the Poor Knights. *Emu - Austral Ornithol.* 41, 56–68. doi:10.1071/MU941056.

Court, D. (1978). Forest regeneration on Hen Island. *Tane* 24, 103–118.

Craig, J. L., and Douglas, M. E. (1984). Bellbirds in Auckland and Northland. *Notornis* 31, 82–86.

Cranwell, L. M., and Moore, L. B. (1935). Botanical notes on the Hen and Chickens Islands. *Rec. Auckl. Inst. Mus.* 1, 301–318.

Department of Conservation (2013). Rodent eradication successful on Taranga (Hen) Island. Available at: https://www.doc.govt.nz/news/media-releases/2013/rodent-eradication-successful-on-taranga-hen-island/ [Accessed February 6, 2019].

Dooling, R. (2004). "Audition: can birds hear everything they sing?," in *Nature's Music* (Elsevier), 206–225. doi:10.1016/B978-012473070-0/50010-4.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Duncan, R. P., and Blackburn, T. M. (2004). Extinction and endemism in the New Zealand avifauna. *Glob. Ecol. Biogeogr.* 13, 509–517. doi:10.1111/j.1466-822X.2004.00132.x.

Fraser, W. M. (1925). *The Poor Knights Islands: a brief account of the Maori occupation*. Government Printer.

Galbraith, M., and Cooper, H. (2013). Tiritiri Matangi: an overview of 25 years of ecological restoration. in (New Zealand Ecological Society).

Galbraith, M. P., and Hayson, C. R. (1995). Tiritiri Matangi Island, New Zealand: public participation in species translocation to an open sanctuary. *Reintroduction Biol. Aust. N. Z. Fauna*, 149–154.

Girardet, S. A. B., Veitch, C. R., and Craig, J. L. (2001). Bird and rat numbers on Little Barrier Island, New Zealand, over the period of cat eradication 1976–80. *N. Z. J. Zool.* 28, 13–29. doi:10.1080/03014223.2001.9518253.

Graham, M. F., and Veitch, C. R. (2002). Changes in bird numbers on Tiritiri Matangi Island, New Zealand, over the period of rat eradication. *Turn. Tide Erad. Invasive Species*, 120–123.

Hayward, B., and McCallum, J. (1984). *Offshore Islands Research Group trip to the Chickens (Marotere) Islands, north-east New Zealand, New Year 1981—1982. Introduction and acknowledgements*. Tane.

Hayward, B. W. (1993). Prehistoric archaeology of the Poor Knights Islands, northern New Zealand. *Tane* 34, 89–105.

Hayward, B. W., Moore, P. R., and Newman, M. (1978). Archaeological site survey of Hen Island (Taranga). *Tane* 24, 137–157.

Heather, B. D., and Robertson, H. A. (2000). *The field guide to the birds of New Zealand.* Viking\Penguin Books.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev.* 91, 13–52. doi:10.1111/brv.12160.

Lovegrove, T. G. (1996). Island releases of saddlebacks *Philesturnus carunculatus* in New Zealand. *Biol. Conserv.* 77, 151–157.

McLintock, A. H., Leslie Owen Kermode, B. A., and Taonga, N. Z. M. for C. and H. T. M. (1966). “Little Barrier Island.,” in *An encyclopaedia of New Zealand, edited by A. H. McLintock, 1966.* Available at: https://teara.govt.nz/en/28198 [Accessed January 26, 2019].

Merton, D. (1972). Cuvier — an island restored. *For. Bird* 184, 7–9.

Nelson, D. A., Hallberg, K. I., and Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology* 110, 879–908.

Parker, K. A., Anderson, M. J., Jenkins, P. F., and Brunton, D. H. (2012). The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecol. Lett.* 15, 778–785. doi:10.1111/j.1461-0248.2012.01797.x.

Prickett, N. (1984). An archaeological survey of the Chickens Islands (Marotere), New Zealand. *Tane* 30, 177–197.

Roper, M. M. (2012). Resource partitioning between two competitive species, the hihi (*Notiomystis cincta*) and bellbird (*Anthornis melanura*), during the non-breeding season on Tiritiri Matangi Island. Masters Thesis, Massey University, New Zealand.

Roper, M. M. (2018). Sexual dimorphism of song and life history trade-offs in the New Zealand bellbird: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph. D.) in Ecology at Massey University, Albany, New Zealand.

Rothenberg, D., Roeske, T. C., Voss, H. U., Naguib, M., and Tchernichovski, O. (2014). Investigation of musicality in birdsong. *Hear. Res.* 308, 71–83. doi:10.1016/j.heares.2013.08.016.

Sayigh, L. S., Esch, H. C., Wells, R. S., and Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Anim. Behav.* 74, 1631–1642. doi:10.1016/j.anbehav.2007.02.018.

Shortland, E. (1851). *The southern districts of New Zealand: a journal, with passing notices of the customs of the aborigines*. Longman, Brown, Green & Longmans.

Sokal, R. R. (1958). A statistical method for evaluating systematic relationship. *Univ. Kans. Sci. Bull.* 28, 1409–1438.

Stringer, I. A. N., Brennan, K. E. C., Moir, M. L., Parrish, G. R., Majer, J. D., and Sherley, G. H. (2003). Population structure, growth and longevity of *Placostylus hongii* (Pulmonata: Bulimulidae) on Tawhiti Rahi Island, Poor Knights Islands, New Zealand. *Pac. Conserv. Biol.* 9, 241–247. doi:10.1071/pc040241.

Taylor, R. B., Morrison, M. A., and Shears, N. T. (2011). Establishing baselines for recovery in a marine reserve (Poor Knights Islands, New Zealand) using local ecological knowledge. *Biol. Conserv.* 144, 3038–3046. doi:10.1016/j.biocon.2011.09.009.

The Cornell Lab of Ornithology (2014). *Raven Pro: Interactive Sound Analysis Software*. Ithaca, NY Available at: http://www.birds.cornell.edu/raven.

Towns, D., McFadden, I., and Thomson, P. (1995). *Offshore islands co-operative conservation project with ICI Crop Care Division: Phase three (Cuvier Island)*. Department of Conservation.

Towns, D. R., Parrish, G. R., Tyrrell, C. L., Ussher, G. T., Cree, A., Newman, D. G., et al. (2007). Responses of Tuatara (Sphenodon punctatus) to Removal of Introduced Pacific Rats from Islands. *Conserv. Biol.* 21, 1021–1031. doi:10.1111/j.1523-1739.2007.00742.x.

Veitch, C. R. (2001). The eradication of feral cats (*Felis catus*) from little Barrier Island, New Zealand. *N. Z. J. Zool.* 28, 1–12.

Veitch, C. R., and Newton, A. (1989). Management of a largely intact natural area: the rehabilitation of Cuvier Island. *Manag. New Zealand's Nat. Estate*, 56.

Waugh, S. M., Tennyson, A. J., Taylor, G. A., and Wilson, K.-J. (2013). Population sizes of shearwaters (*Puffinus* spp.) breeding in New Zealand, with recommendations for monitoring. *Tuhinga* 24, 159–204.

4 *Koe*: web-based software to classify acoustic units and analyse sequence structure in animal vocalisations

4.1 Abstract

Classifying acoustic units is often a key step in studying repertoires and sequence structure in animal communication. Manual classification by eye and ear remains the primary method, but new tools and techniques are urgently needed to expedite the process for large, diverse datasets. Here we introduce *Koe*, an application for classifying and analysing animal vocalisations. *Koe* offers bulk-labelling of units via interactive ordination plots and unit tables, as well as visualisation and playback, segmentation, measurement, data filtering/exporting and new tools for analysing repertoire and sequence structure—in an integrated environment. We demonstrate *Koe* with a real-world case study of New Zealand bellbird *Anthornis melanura* songs from an archipelago metapopulation. Having classified 21,500 units in *Koe*, we compare repertoires and sequence structure between sites and sexes. *Koe* is web-

based (koe.io.ac.nz) and easy to use, making it ideal for collaboration, education and citizen science. By enabling large-scale, high-resolution classification and analysis of animal vocalisations, *Koe* expands the possibilities for bioacoustics research.

4.2 Introduction

Acoustic communication is important for many species across the animal kingdom, used in contexts such as territorial defence, mate attraction, group coordination, and recognition of individuals and species (Peters et al., 1980; Catchpole and Slater, 2008; Soltis, 2009; Janik and Sayigh, 2013; Starnberger et al., 2014; Sperber et al., 2017). Often, acoustic communication is structured as a sequence of distinct acoustic units, where information is encoded in the types of units and sometimes their sequence structure (i.e. temporal arrangement) (Kershenbaum et al., 2016). In animals with simple, stereotyped songs, it may be possible to derive useful measures of similarity by comparing entire songs. However, many species have complex, flexible vocalisations, where unit-level (e.g. syllable) analysis is essential, for example to assess unit sharing or evaluate sequence structure. A common workflow for studying such cases is summarised in Figure 4.1 (see Kershenbaum et al., 2016 for a review).

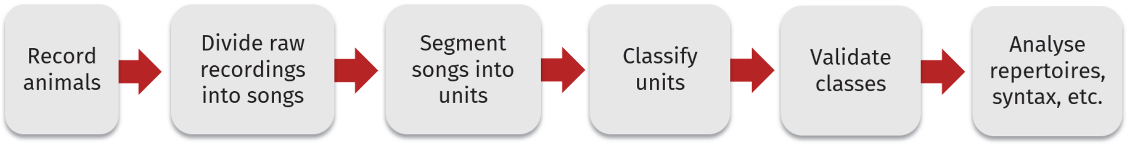


Figure 4.1—A typical process for studies in bioacoustics.

Classifying units involves identifying unit classes (i.e. producing a ‘catalogue' of class types) and then assigning units to those classes; only once units have been classified can repertoires and sequence structures be analysed and compared. Classification is a challenge, however; thus far, attempts at automation have not proven generalisable (Priyadarshani et al., 2018) and the primary approach for most species remains manual classification based on human visual and auditory perception (Kershenbaum et al., 2016).

Although human perception-based methods excel at judging unit similarity (Sayigh et al., 2007; Duda et al., 2012), they can nonetheless be very time consuming. The researcher needs to view spectrograms and hear playback of units, comparing similarity of many units at a time, in order to discern and assign classes. We know of no tools designed for this purpose; existing acoustics software offers visualisation and labelling functionality only within one or few files at a time, not on the large comparative scale required (Table 4.2). Researchers currently resort to cumbersome and non-integrated workflows that severely limit the scope of bioacoustics research.

Here we introduce *Koe*, new web-based software designed to expedite large-scale classification and analysis of acoustic units. *Koe* offers bulk-labelling of units via novel interactive ordination plots, as well as visualisation and playback, segmentation, measurement, data filtering/exporting and new tools for analysing sequence structure, in an integrated environment.

Koe was developed in *Python 3* (Rossum, 1995), which provides flexible server-side computation (Oliphant, 2007; Pedregosa et al., 2012; Jones et al., 2014). For the web framework, we chose *Django* (Django Software Foundation, 2013). Being web-based gives *Koe* numerous benefits: it is accessible from any device, requires no setup, supports multiple simultaneous users, and can run in multiple browser windows at once (useful for viewing different program views concurrently). This makes it ideal for collaboration, teaching and citizen science. The official application, maintained by the authors, is available at koe.io.ac.nz. Source code is available at github.com/fzyukio/koe.

4.3 *Koe* workflow

We designed *Koe* to provide end-to-end functionality in an intuitive and flexible workflow (Figure 4.2). Each step in the workflow has a program view tailored to that task. Here we highlight significant features of each program view; for more detail, see *Koe* documentation (github.com/fzyukio/koe/wiki).

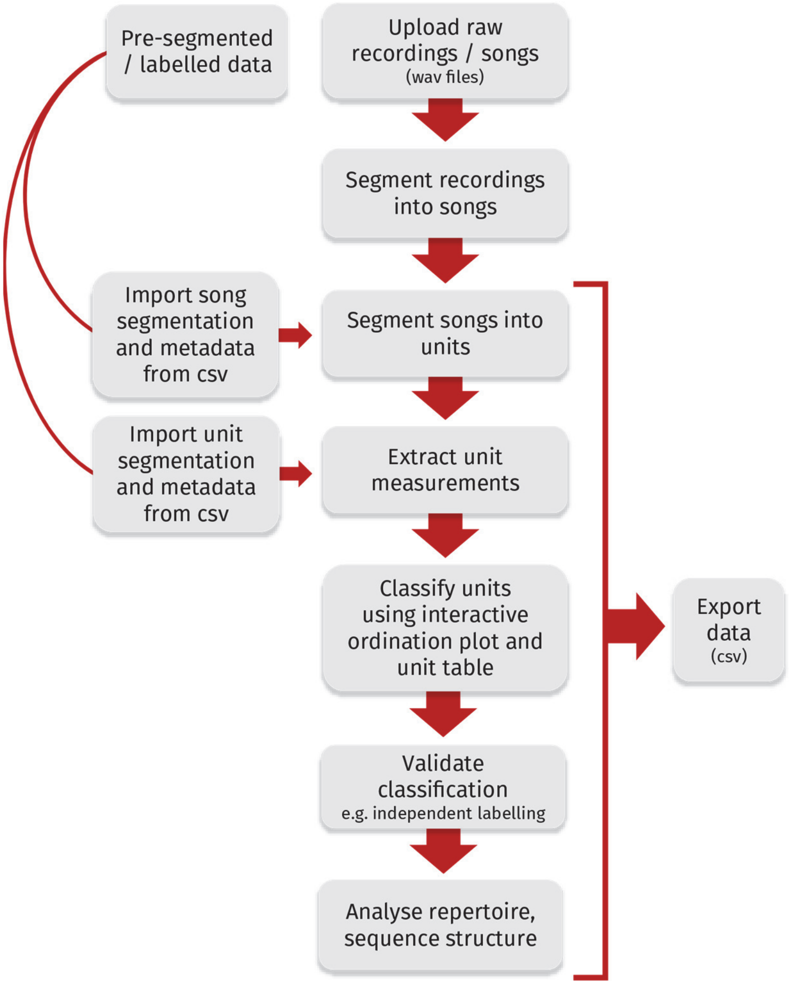


Figure 4.2—Suggested *Koe* workflow.

Koe data is integrated; any step in the workflow can be revisited, and modifications will dynamically update throughout the program. Raw recordings / songs (in wav format) are uploaded. Recordings are divided into songs, which are segmented into units. *Koe* extracts unit features to produce similarity indices and interactive ordination plots, which help a user to rapidly and accurately classify units. Once complete, classification can be validated by independent labelling. *Koe* provides tools for analysing repertoires and sequence structure. Data can also be exported for external use.

4.3.1 Upload and segment recordings

In *Koe*, the user can partition raw recordings into individual vocalisation bouts, termed 'songs' (*Upload & split raw recordings* view), and segment songs into their constituent acoustic units (*Segment songs into units* view). The tasks are similar; they involve the user selecting start/endpoints of songs or units, respectively, on a spectrogram, then committing the selections to the database (stored securely on the *Koe* server). It is up to the user to decide on segmentation criteria appropriate to their species/question. The user can adjust playback speed, spectrogram contrast and time-axis zoom. Saved selections become available in other program views. Acoustic units are not stored as audio segments, but as start/endpoint information referencing the source song; the user can freely readjust unit segmentation and program views will update dynamically. For any songs already segmented into units in other software, start/endpoints can be imported as a csv file. Pre-partitioned song files can also be uploaded directly to the database in *View all songs.* Users can grant database access to other users, with custom permission levels.

4.3.2 Extract acoustic features from units

In quantitative analysis, raw signals are often replaced by their compact representations using a specific set of features, extracted from the signals. Analyses performed on compact representations can be more effective and computationally efficient. In *Extract unit features* view, *Koe* can extract a wide array of spectral-, tempo- and chroma-related features for bioacoustic analyses (see github.com/fzyukio/koe/wiki). Extracted features are utilised by *Koe* to construct ordinations and similarity indices (see below), and can be exported as csv files for analysis in other software.

Attribute	Value
Length	3.79 secs
Name	LBI_2016_04_05_BAE003_02_F.EX.LBI_2016_04_05_BAE003A.fullSongBySameBird.wav
Species	Anthornis melanura
Date	2016-04-05
Quality	EX
Track	LBI_2016_04_05_BAE003
Individual	LBI_2016_04_05_BAE003A
Sex	F

☐	ID	Start	End	Duration	Label	Family	Subfamily	Note
☐	7795	27	287	260	PipeOverlap(A#[6]-C[6])	Stepdown		
☐	7796	489	675	186	LBIShriek	Shrieky		
☐	7797	803	977	174	Dinnerplate(1)	Pipe_Down		
☐	7798	1074	1229	155	Pipe(A#[6]+click+)	Pipe		
☐	7799	1318	1505	187	Pipe(C#[7]+)	Pipe		
☐	7800	1565	1743	178	Down_Pipe(G[6])	Down_Pipe		

Figure 4.3—*Segment songs into units* view, showing a song being segmented into units.

Units are manually segmented by dragging over the spectrogram; unit endpoints can be readjusted at any time. A selection box can be clicked for playback. Spectrogram zoom, contrast and colourmap can be adjusted. Units can be labelled and commented. This example is a female NZ bellbird song from Hauturu. The interface for partitioning recordings into songs is similar.

4.3.3 Classify units

Manual classification requires the ability to visually/acoustically compare and label large numbers of units quickly. *Koe* offers interactive ordination plots, unit tables and class exemplars as complementary tools for this purpose.

Interactive ordination plots

A previously unexplored potential of ordination is to expedite the manual classification of units. *Koe*'s interactive ordination plots (Figure 4.4) incorporate audio playback, spectrograms and classification functionality, so that a user can simultaneously use their audio-visual perception of unit similarity and the structure of the data to rapidly and robustly classify units.

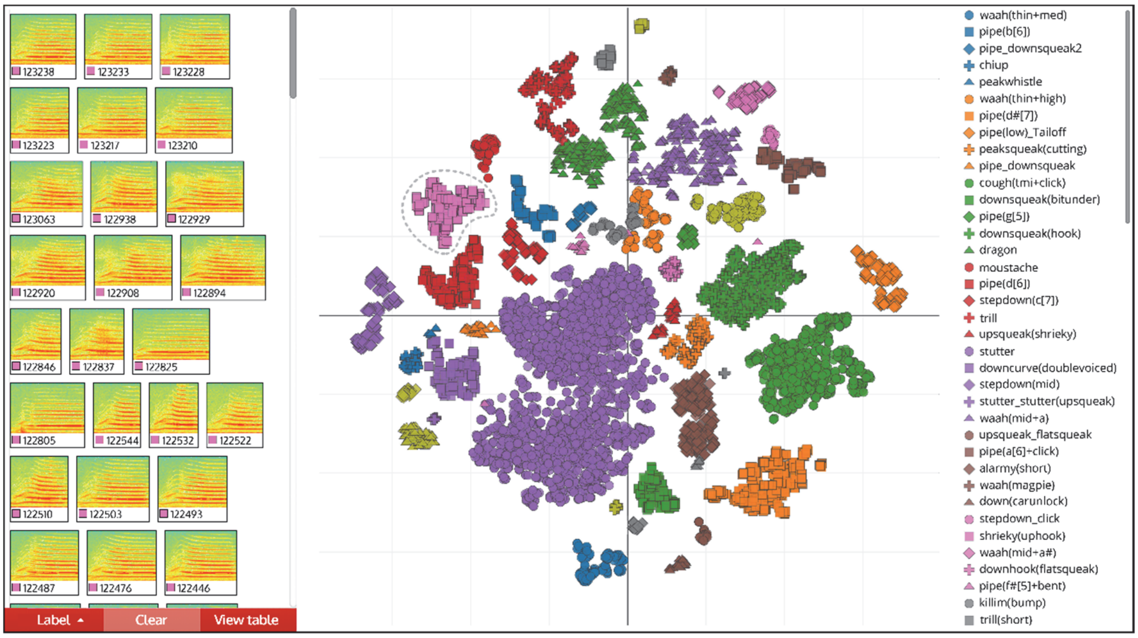


Figure 4.4—*Koe*'s interactive *Ordination* view.

Interactive ordination view allows the user to encircle groups of points on the plot with the lasso tool, to view their spectrograms and hear their audio. Mousing over a point in a selection highlights the corresponding spectrogram in the left-hand panel. Selections can be labelled in bulk directly on the plot or opened as a unit table to view detailed unit information. The user can zoom, toggle the visibility of classes, and export the plot as a vector graphic. This example shows a t-SNE ordination of 7189 syllables of male and female NZ bellbird song.

Koe implements three ordination techniques: Principal Component Analysis (PCA; Pearson, 1901), Independent Component Analysis (ICA; Comon, 1994), and t-distributed Stochastic Neighbour Embedding (t-SNE; Van Der Maaten and Hinton, 2008). t-SNE aims to preserve local structure in the data and is particularly effective for defining and discriminating between clusters.

The user encircles groups of points on the plot to see spectrograms and hear playback of units; if a selection appears acoustically consistent, it is classified in bulk. The user can toggle visibility of each class independently and can zoom to examine structural detail.

Unit tables

As a complement to the ordination plot, units can be viewed as an interactive table (Figure 4.5). Each unit is represented by a row containing spectrogram, audio and associated information (class label, unit duration, song ID, individual ID, date, etc.). The table can be sorted/filtered by any column. A notable feature is the *similarity index*, which ranks units based on acoustic similarity. The index is produced as follows: from the raw feature measurements or from the ordination, *Koe* calculates pairwise Euclidean distance between each pair of units, then constructs a ladderised dendrogram using agglomerative hierarchical clustering (UPGMA) (Sokal, 1958). The order of the dendrogram leaf nodes becomes the similarity index. Sorting by the *similarity index* column orders the table so that similar units arrange together, allowing them to be selected and labelled in large batches.

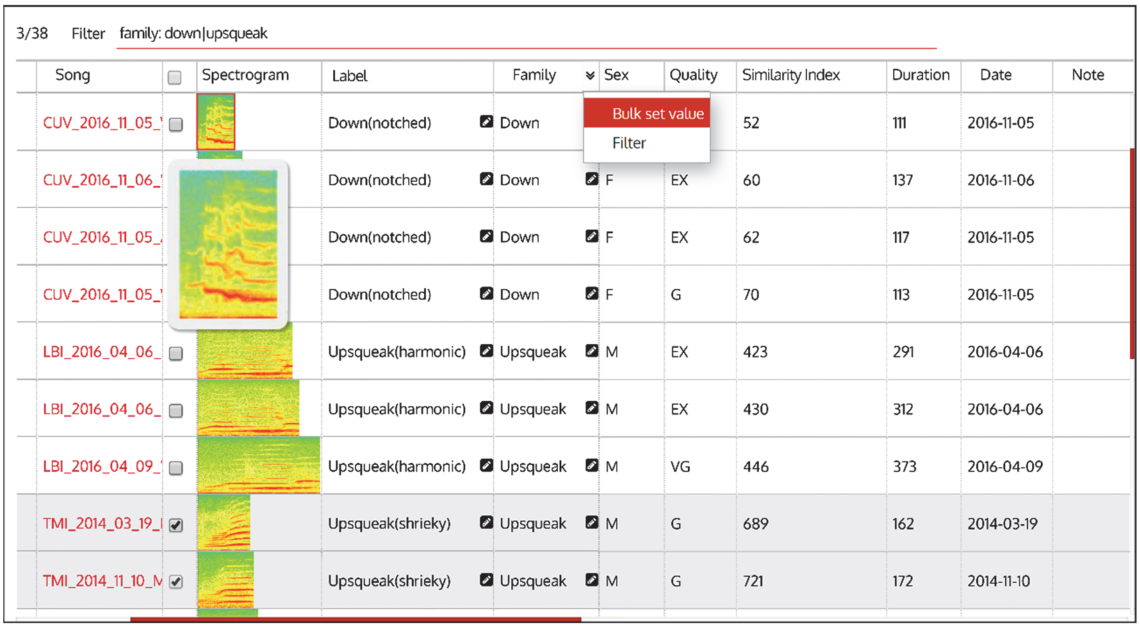
Class exemplars

A class catalogue is a useful reference during classification. *Koe* produces one automatically in *Exemplars* view, with exemplar spectrograms and playback for every class, making visual/acoustic comparison easy (Figure 4.6). The catalogue updates dynamically as classification progresses, displaying 10 randomly-chosen exemplars per class to reflect within-class variability.

Classification granularity

Without *a priori* knowledge of the animal's perception of units, a researcher must identify classes and assign units based on their own perception. Classifying units at multiple hierarchical levels of granularity increases robustness by enabling analyses at different scales. *Koe* offers up to three granularity levels (fine, medium, and broad-scale) for labelling in the *Ordination* and *Unit table* views. For example, the units in

Figure 4.5 are labelled at two granularity levels: the broad-scale *Upsqueak* family is subdivided at the fine-scale into *Upsqueak(harmonic)* and *Upsqueak(shrieky)*.

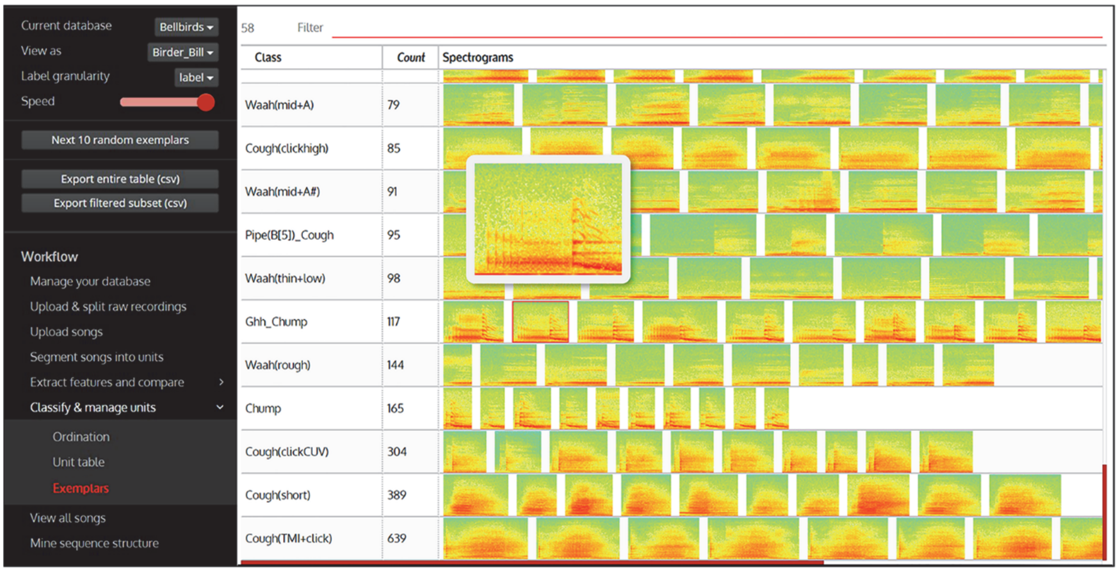


3/38 Filter family: down|upsqueak

Song	☐	Spectrogram	Label	Family	Sex	Quality	Similarity Index	Duration	Date	Note
CUV_2016_11_05_	☐		Down(notched)	Down			52	111	2016-11-05	
CUV_2016_11_06_			Down(notched)	Down	F	EX	60	137	2016-11-06	
CUV_2016_11_05_			Down(notched)	Down	F	EX	62	117	2016-11-05	
CUV_2016_11_05_			Down(notched)	Down	F	G	70	113	2016-11-05	
LBI_2016_04_06_	☐		Upsqueak(harmonic)	Upsqueak	M	EX	423	291	2016-04-06	
LBI_2016_04_06_	☐		Upsqueak(harmonic)	Upsqueak	M	EX	430	312	2016-04-06	
LBI_2016_04_09_	☐		Upsqueak(harmonic)	Upsqueak	M	VG	446	373	2016-04-09	
TMI_2014_03_19_	☑		Upsqueak(shrieky)	Upsqueak	M	G	689	162	2014-03-19	
TMI_2014_11_10_	☑		Upsqueak(shrieky)	Upsqueak	M	G	721	172	2014-11-10	

Figure 4.5—*Unit table* view in *Koe*.

Koe's unit table is designed for classifying, annotating and filtering units. Each unit row contains a spectrogram which becomes enlarged during mouse-over. Unit audio plays when a spectrogram is clicked. The table can be sorted/filtered by any columns. Sorting by the *similarity index* column arranges units by spectral similarity for expedited labelling.



Current database Bellbirds
View as Birder_Bill
Label granularity label

58 Filter

Class	Count	Spectrograms
Waah(mid+A)	79	
Cough(clickhigh)	85	
Waah(mid+A#)	91	
Pipe(B[5])_Cough	95	
Waah(thin+low)	98	
Ghh_Chump	117	
Waah(rough)	144	
Chump	165	
Cough(clickCUV)	304	
Cough(short)	389	
Cough(TMI+click)	639	

Figure 4.6—*Exemplars* view in *Koe.*

Exemplars view displays one unit class per row, with up to 10 randomly-selected exemplars per class, serving as a reference catalogue. Lists of classes and their counts can be exported as a CSV file. The navigation panel on the left is visible in all views; the top portion of the panel displays controls applicable to the active view, while the lower half is used to navigate between views. Example data are NZ bellbird song syllables.

4.3.4 Validate classification through independent labelling

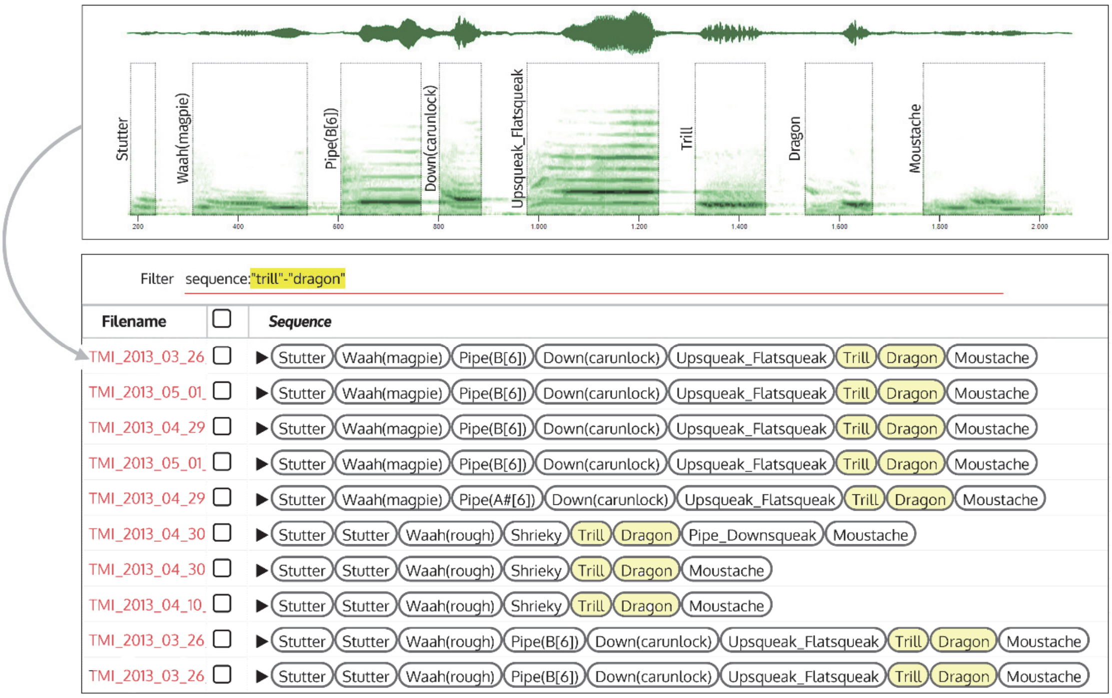
To ensure robustness of manual classification, a common validation method is to have several judges independently label the dataset and calculate their degree of agreement; high agreement lends credibility to the classification (Nelson et al., 2004; Parker et al., 2012). Typically, 2–5 judges are used, but robustness can improve substantially with more judges (Jones et al., 2001). *Koe* facilitates independent labelling experiments. A subset of songs/units can be selected and copied to a new database. The database owner grants labelling access to participating judges, who label the dataset online. The labels of participants are automatically saved to the server and are compiled to evaluate concordance (see Case Study below for a real-world example using 74 judges).

4.3.5 Analyse sequence structure

For many animals, songs are comprised of acoustic units ordered into a sequence (Kershenbaum et al., 2016). The aim of sequence analysis is to reveal patterns in the sequence structure of songs, which *Koe* can do in three ways, outlined below.

Filter songs by subsequence

In *View all songs* a user can filter for songs that contain a certain subsequence of unit labels (Figure 4.7). This could be used to identify all instances of a certain song type, for example.



Filter sequence:"trill"-"dragon"

Filename	☐	Sequence
TMI_2013_03_26	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_05_01_	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_29	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_05_01_	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_29	☐	▶ Stutter Waah(magpie) Pipe(A#[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_30	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Pipe_Downsqueak Moustache
TMI_2013_04_30	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Moustache
TMI_2013_04_10_	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Moustache
TMI_2013_03_26	☐	▶ Stutter Stutter Waah(rough) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_03_26	☐	▶ Stutter Stutter Waah(rough) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache

Figure 4.7—*Koe*'s songs list view.

View all songs displays songs in an interactive table, with one song per row. Segmented songs (like the example in the top panel) are represented as a sequence of unit labels. Entire songs can be played, or individual labels clicked for unit playback and spectrograms. Here the filter has returned all songs containing *Trill* followed by *Dragon* (highlighted).

Discover patterns using sequence rule mining (cSPADE)

As one way of exploring rules that may govern song structure (i.e. syntax; Lachlan et al., 2013), *Koe* uses the cSPADE (constrained Sequential Pattern Discovery using Equivalence classes) algorithm (Zaki, 2001) to discover commonly-occurring sequences in a set of songs.

We consider a sequence to be an ordered list of *x* acoustic units denoted as A→B→C→ ...→ *x*.

A sequence rule has two parts: a left and right side. The rule states that when the left side occurs, the right side follows: [Left side] ⇒ Right side. For example, the rule

[A→B] ⇒ C states that when the sequence [A→B] occurs, C comes next. The left side can be a sequence of any length, but in our implementation the right side is a single unit (see https://github.com/fzyukio/koe/wiki, *Mine Sequence Structure*).

The cSPADE algorithm calculates the credibility of sequence rules. Credible rules have a large confidence factor, a large level of support and a value of lift greater than one, as defined below using the example rule, [A→B] ⇒ C.

Support

The proportion of songs in the database that contain the entire sequence [A→B→C] at least once.

Confidence

The proportion of those songs containing [A→B] that also contain [A→B→C].

Lift

A measure of the strength of the association relative to chance. Lift is equal to the confidence of the rule, divided by the proportion of songs containing the right side. Thus it gives the ratio of (i) the proportion of songs with [A→B] that transition to C, versus (ii) the proportion of songs expected to contain [A→B→C] by chance association of [A→B] and C.

To demonstrate cSPADE, consider the ten songs shown in Figure 4.7 to be a population of songs. The rule [*Stutter→Waah(magpie)*] ⇒ *Pipe(B6)* has a **support** of 0.4 since the entire sequence occurs in four of the 10 songs. The rule has a **confidence** of 0.8, because in four of the five songs that contain *Stutter→Waah(magpie)*, the transition to *Pipe(B6)* occurs. The proportion of songs with *Pipe(B6)* is 0.6, so the **lift** of this rule is 0.8/0.6=1.33. That is, the association

[*Stutter*→*Waah(magpie)*] ⇒ *Pipe(B6)* occurs in 1.33 times as many songs as expected by chance association of [*Stutter*→*Waah(magpie)*] and *Pipe(B6)*.

4.3.5.1 Visualise sequence structure as a network

Two-unit associations from cSPADE can be visualised using a *directed network*. The network models the direction and strength of association between pairs of units, across a population of songs. Units are represented by nodes which are joined by lines (edges) if the units occur consecutively. The order of units is indicated by arrow direction, and strength of association between units (lift) is represented by edge thickness. Visually cluttered networks can be simplified using the filter, e.g. to show only associations with high lift.

Thus, *Koe* offers pattern recognition and sequence visualisation. If a user desires to make formal inferences, such as assessing the influence of experimental factors (fixed effects) and random effects of individuals on sequence structure, sophisticated Markovian frameworks exist (e.g. Sarkar et al., 2018). A user can export sequence data from *Koe* for external analysis.

4.4 Case study: New Zealand bellbird song

Here we showcase *Koe*'s analysis features with a large (21,500-unit) dataset of New Zealand bellbird *Anthornis melanura* (hereafter ‘bellbird’) songs, manually classified in *Koe* using the methods described above. The songs were recorded at six sites in an archipelago metapopulation (Figure 4.8). Bellbird songs are complex and diverse, both in units (i.e. syllables) and sequence structure (Heather and Robertson, 2000; Brunton and Li, 2006). Furthermore, both sexes of bellbird sing prolifically with sexually dimorphic songs and site-specific dialects. This presents a challenging test case for classification and for analysing repertoires and sequence structure. For a step-by-step case study tutorial see the *Koe* documentation (https://github.com/fzyukio/koe/wiki/Analysis-Tutorial%3A-Case-Study-of-NZ-bellbird-song).

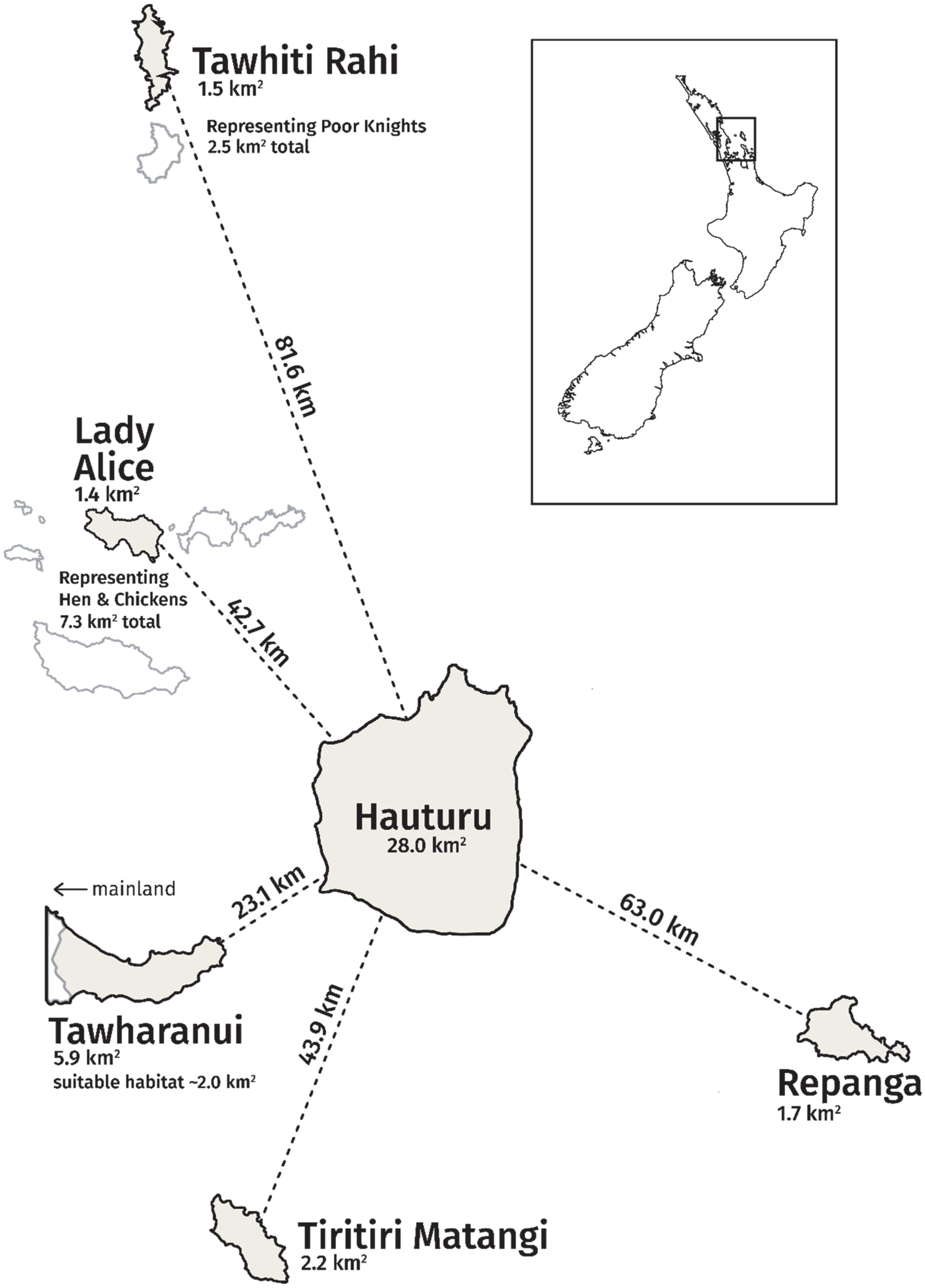


Figure 4.8—The six sites where NZ bellbird *Anthornis melanura* populations were recorded in the Hauraki Gulf, New Zealand.

Between 2013–2017, songs were recorded from each site using portable recorders and shotgun microphones. For simplicity, only distances from Hauturu are shown (distances not to scale).

4.4.1 Validating our classification with independent labelling

We harnessed the citizen science potential of *Koe* to evaluate inter-observer reliability using 74 judges (to our knowledge the largest number of judges yet used). From our labelled dataset of 21,500 bellbird syllables (units), we constructed a subset database of 200 syllables in *Koe*. The 200 syllables consisted of 18 label classes from two populations—Tawharanui and Tiritiri Matangi (Tiri)—with 3–20 of each class, including 4 “other” syllables that did not match the exemplars (the classes and numbers of each class were chosen randomly). The 74 judges (all naïve to *Koe*, spectrograms and bellbird song) created *Koe* accounts and were granted online access to the database. Judges were asked to independently label the syllables using the unit table by comparing against labelled class exemplars. To reduce bias, the judges were not given the exact number of each class, but rather a range that included the true value. Hence they were told there were 3–20 of each class, and 3–20 “other” syllables that did not match the exemplars. For each syllable, the percentage of judges whose labels matched our own label was calculated: average 89.6%; median 95.6%. This is a high degree of agreement for inter-observer reliability studies (Nelson et al., 2004; Parker et al., 2012), lending validity to our manual classification.

4.4.2 Analysing syllable diversity

Data can be explored quickly using *Koe*'s filter, which combines different criteria specified by a regular expression. For example, `duration:<100; sex:F; label:Downsqueak` returns only units with duration shorter than 100 ms, vocalised by female birds, and label containing *Downsqueak*. Below, we use the filter to compare male and female unit diversity across the whole metapopulation, then between males at two sites.

What is the overall unit diversity for male and female bellbirds across all six sites?

Exemplars view displays the catalogue and count of unit classes, revealing 754 fine-scale classes in all. In *Unit table* view we made subset collections by sex. Viewing the subsets in *Exemplars* view revealed 517 fine-scale classes sung by males, 415 sung by females and 178 (24%) sung by both.

How many male unit types are now shared between Tawharanui and source population Hauturu?

In 2005, bellbirds from Hauturu naturally recolonized Tawharanui Peninsula following pest control. At that point, repertoires of source and founder populations were indistinguishable (Brunton et al., 2008). After a decade of potential cultural divergence, what is the current degree of male repertoire overlap between the sites?

In *Unit table* view we made subsets containing units sung by males from Tawharanui (`sex:M; song:TAW`) and Hauturu (`sex:M; song:LBI`). These subsets showed 141 male fine-scale unit classes on Hauturu, and 104 at Tawharanui. By examining the intersect of the two class lists in *R Studio* (RStudio Team, 2015), we found 35 Tawharanui unit classes shared with Hauturu (34%)—since colonization 66% of the unit classes sung by males at Tawharanui appear to have been innovated.

4.4.3 Evaluating song structure in bellbirds

To analyse sequence structure in detail, *Mine sequence structure* shows the strength of association between unit classes using the cSPADE algorithm (Zaki, 2001), and produces network visualisations. This can be used to visually compare sequence structure between any data subsets, e.g. individuals, sexes, sites or seasons. Here we compare the sequence structure of male and female bellbird song on Tiri.

Male and female song structure as revealed by cSPADE

One-unit 'associations' show prevalence of syllable classes in the population. The three most prevalent classes for males are *Dragon*, *Cough(TMI+click)* and *Downsqueak(hook)*, occurring in 52%, 40% and 35% of male songs, respectively (**Error! Reference source not found.**). The three most prevalent female classes are *Stutter*, *Chiup* and *Downcurve(doublevoiced)*, occurring in 93%, 59% and 36% of female songs, respectively.

The most prevalent male two-unit association, *Cough(TMI+click)* ⇒ *Downsqueak(hook)*, occurs in 32% of songs. It has a confidence of 0.81; i.e. when *Cough(TMI+click)* occurs, *Downsqueak(hook)* comes next in 81% of songs. It has a lift of 2.33; i.e. *Downsqueak(hook)* follows *Cough(TMI+click)* in 2.33 times the number of songs as expected by chance. Male song most often starts with *Stutter* (29% of songs) or *Cough(TMI+click)* (23%). It most often ends with *Dragon* (30%), *Moustache* (12%) or *Cough(TMI+click)* (12%). By contrast, for females**,** the importance of *Stutter* is apparent. Not only does *Stutter* start in 91% of female songs, it also repeats (i.e. *Stutter* ⇒ *Stutter*) in 85% and ends in 38%. Furthermore, *Stutter* is commonly repeated many (≤11) times.

Table 4.1—Key features of male and female song structure as extracted by the cSPADE algorithm.

Results for NZ bellbird males (389 songs comprised of 5305 syllables) and females (138 songs comprised of 1884 syllables) recorded on Tiri between February 2013 and December 2015. Rule length is how many units constitute the association. Song count is how many songs the association occurs in. See text for definitions of support, confidence and lift.

Rule length	Association rule	Song count	Support	Confidence	Lift
Male population					
1	Dragon	201	0.52	—	—
1	Cough(TMI+click)	156	0.40	—	—
1	Downsqueak(hook)	135	0.35	—	—
…	…	…	…	…	…
2	Cough(TMI+click) ⇒ Downsqueak(hook)	126	0.32	0.81	2.33
2	START ⇒ Stutter	113	0.29	0.29	0.95
2	START ⇒ Cough(TMI+click)	90	0.23	0.23	0.57
2	Dragon ⇒ END	115	0.30	0.57	0.57
2	Moustache ⇒ END	46	0.12	0.81	0.81
2	Cough(TMI+click) ⇒ END	45	0.12	0.29	0.29
…	…	…	…	…	…
3	[Trill → Dragon] ⇒ Moustache	32	0.08	0.86	5.9
…	…	…	…	…	…
Female population					
1	Stutter	129	0.93	—	—
1	Chiup	82	0.59	—	—
1	Downcurve(doublevoiced)	50	0.36	—	—
…	…	…	…	…	…
2	START ⇒ Stutter	125	0.91	0.91	0.97
2	Stutter ⇒ Stutter	117	0.85	0.91	0.97
2	Stutter ⇒ END	52	0.38	0.40	0.40
2	Downcurve(doublevoiced) ⇒ Chiup	49	0.36	0.98	1.65
…	…	…	…	…	…

9	[Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter] ⇒ Stutter	21	0.15	0.75	0.80
...	...	...	...	...	...
11	[Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter → Stutter] ⇒ Stutter	10	0.07	0.77	0.82

Contrasting male and female song structure with network diagrams

Sexual dimorphism in two-unit sequences was examined at the population level with network diagrams with a threshold of support>0.05 (**Error! Reference source not found.**). The songs of both sexes contain some nodes with many connections; these unit classes (e.g. *Dragon* in males, *Stutter* in females) can be preceded/followed by many different classes. Both sexes also contain nodes which are always preceded/followed by a certain class; for example, in males *Waah(mid+A)* is always followed by *Dragon*, and in females *Peakwhistle(TMI)* is always preceded by *Stutter*. Both sexes sing repeated units, represented by looped lines. For males: *Dragon*, *Stutter*, and *Cough(TMI+click)*. For females: *Stutter*, *Stutter(wavey)*, *Pipe(D#[7]+)*, and *Pipe(E[7])*. A notable difference in the female network (**Error! Reference source not found.**B) is the chain of strongly associated nodes: *Killem* → *Pipe(C#[7]+)* → *Upsqueak_Flatsqueak(1)* → *Downsqueak(twopart)* → *Down_Click_Stepdown* → *Stepdown(C[7]-C[6])*.

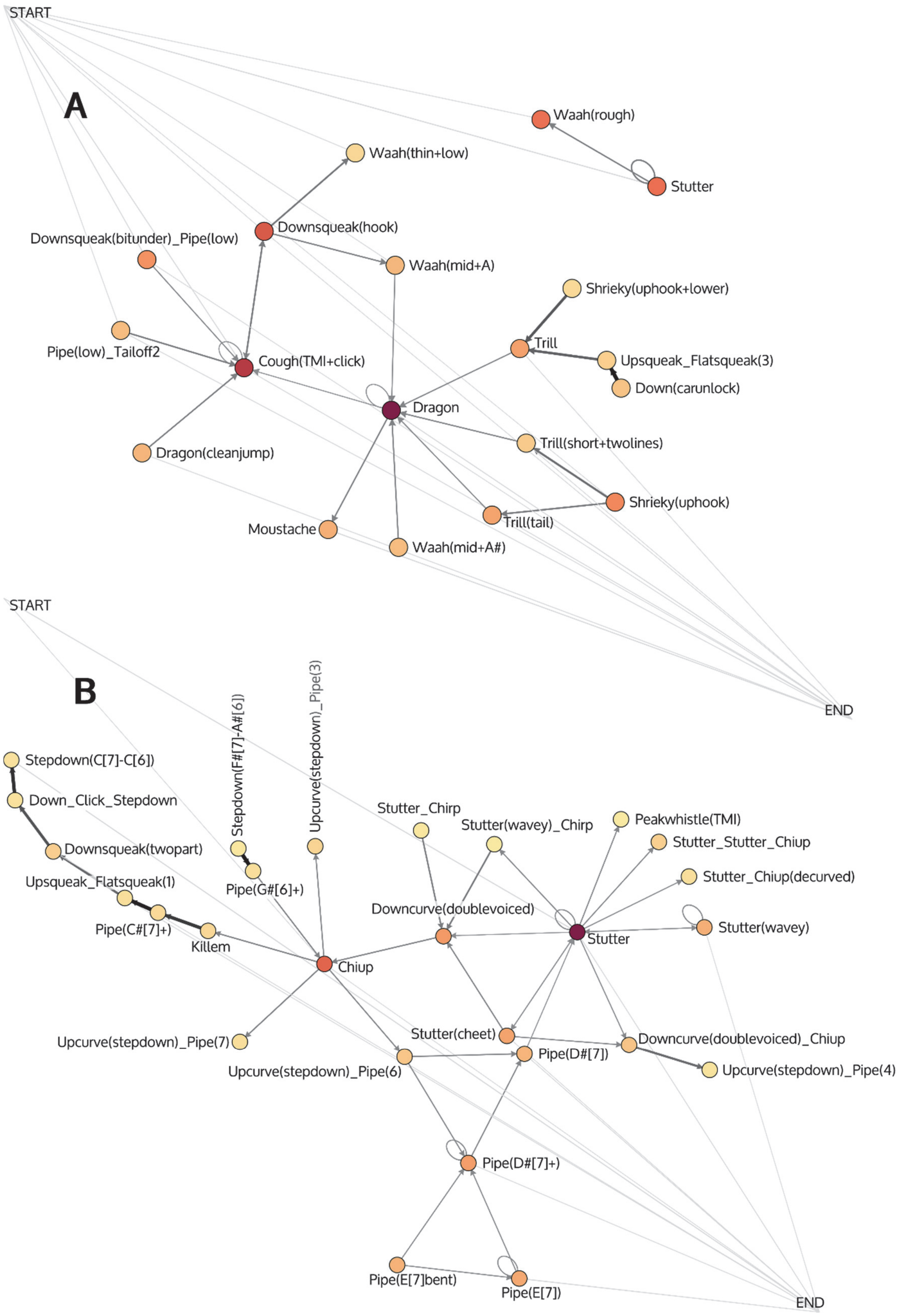

A
START
Waah(rough)
Waah(thin+low)
Stutter
Downsqueak(hook)
Downsqueak(bitunder)_Pipe(low)
Waah(mid+A)
Shrieky(uphook+lower)
Trill
Pipe(low)_Tailoff2
Cough(TMI+click)
Upsqueak_Flatsqueak(3)
Down(carunlock)
Dragon
Trill(short+twolines)
Dragon(cleanjump)
Shrieky(uphook)
Trill(tail)
Moustache
Waah(mid+A#)
END
B
START
Stepdown(F#[7]-A#[6])
Upcurve(stepdown)_Pipe(3)
Stepdown(C[7]-C[6])
Down_Click_Stepdown
Stutter_Chirp
Stutter(wavey)_Chirp
Peakwhistle(TMI)
Stutter_Stutter_Chiup
Downsqueak(twopart)
Stutter_Chiup(decurved)
Upsqueak_Flatsqueak(1)
Pipe(G#[6]+)
Downcurve(doublevoiced)
Pipe(C#[7]+)
Stutter
Stutter(wavey)
Killem
Chiup
Stutter(cheet)
Downcurve(doublevoiced)_Chiup
Upcurve(stepdown)_Pipe(7)
Pipe(D#[7])
Upcurve(stepdown)_Pipe(4)
Upcurve(stepdown)_Pipe(6)
Pipe(D#[7]+)
Pipe(E[7]bent)
Pipe(E[7])
END

Figure 4.9—Network visualisations of male and female song sequence structure.

The networks show two-unit associations for (A) male and (B) female NZ bellbird song on Tiri. The networks are produced from the cSPADE data in Table 4.1. Each node represents a unit class. Colours of the nodes represent frequency of occurrence, with darker, redder nodes being more frequently recorded classes. The arrows show directions of the association. Thickness and darkness of the arrows represent the strength of the association. The networks have been filtered to show only those associations that occur in >5% of songs. Classes that start or end in at least one instance are tethered to the START/END nodes, respectively, with thin grey lines.

A complete set of network visualisations for all six sites in the Hauraki Gulf archipelago is provided in Appendix 7.2, with a comparative analysis of male and female networks.

4.5 Conclusion

Koe is an end-to-end solution for large-scale, high-resolution classification and analysis of animal vocalisations. It features tools for segmenting, comparing and classifying acoustic units and analysing sequence structure. Interactive ordination and unit tables provide a major advance in classification speed and robustness over existing methods. We have demonstrated the power of *Koe* for large-scale analyses with a case study of 21,500 syllables of bellbird song, classified in *Koe* into over 700 classes and analysed in terms of repertoires and sequence structure. *Koe* was designed primarily for taxa/questions requiring acoustic classification; however, *Koe* is not limited to such cases. Where classification is not the aim, e.g. in animals where vocalisations show graded (non-discrete) variation, *Koe* is still a powerful tool for efficient extraction of measurements, which can be exported for analysis. Being web-based and accessible from any device, *Koe* is ideal for collaboration, teaching and citizen science.

4.6 Data availability

Koe source code is available on GitHub (github.com/fzyukio/koe) and is archived on Zenodo (https://doi.org/10.5281/zenodo.3538887). *Koe* users have access to the *Bellbirds_Case_Study* database analysed in this paper; instructions are available in the user manual at koe.io.ac.nz, which contains a step-by-step tutorial to reproduce our analyses. Label data for reproducing our analyses, and a set of bellbird song recordings, are also available on Dryad (https://doi.org/10.5061/dryad.h44j0zpfg).

4.7 Author contributions

Wesley Webb and Yukio Fukuzawa conceived of the software and developed the project with guidance from the other authors. Yukio Fukuzawa coded the software. Andrew Gilman arranged the hosting server. Matthew Pawley, Andrew Gilman, and Stephen Marsland gave valuable insights into the algorithms *Koe* utilizes. Wesley Webb led the writing of the manuscript and user manual with contributions from the other authors. Wesley Webb, Michelle Roper, and Dianne Brunton contributed song data for the development of *Koe* and for the case study.

4.8 Program performance

A comparison of *Koe* functionality against existing bioacoustics software is given in Table 4.2.

Table 4.2—Comparison of *Koe* functionality with existing bioacoustics software.

		Koe	***Raven Pro*** [A]	***Sound Analysis Pro 2011*** [B]	***Luscinia*** [C]	***WarbleR*** [D]	***Avisoft*** [E]
1	Date of latest release	2019	2017	2011	2016	2019	2019
2	Cost	Free	100 USD p.a, or 400–800 USD permanent licence.	Free	Free	Free	1800–2400 EUR permanent licence
3	Platform	Any web-enabled device	Windows/Mac	Windows	Windows/Mac	Windows/Mac (R package)	Windows
4	Open source	✓	✗	✗	✓	✓	✗
5	Metadata integrated as a database	✓	✗	✓	✓	✗	✓
6	Filter/search metadata	✓	✗	✓	✗	✗	✓
7	Multiple simultaneous users (collaboration)	✓	✗	✗	✗	✗	✗
8	Segmentation start/end points of units saved to database	✓	✗	✗	✓	✗	✓

9	Interactively visualise many units at once (spectrograms / playback)	✓	✗	✗	✗	✗	✗
10	Manually classify units in bulk	✓	✗	✗	✗	✗	✗
11	Ordination-assisted manual classification	✓	✗	✗	✗	✗	✗
12	Interactive catalogue of class exemplars	✓	✗	✗	✗	✗	✗
13	Visualise feature measurement accuracy	✗ (not yet)	✓	✓	✓	✓	✓
14	Visualise sequence structure across multiple songs	✓	✗	✗	✗	✗	✗
15	Visualise sequence structure as label strings	✓	✗	✗	✗	✗	✗
16	Visualise sequence structure as networks	✓	✗	✗	✗	✗	✗
17	Filter songs by sequence	✓	✗	✗	✗	✗	✗
18	Bulk export song selections as audio files	✓	✗	✗	✗	✗	✗

A: The Cornell Lab of Ornithology. (2017). Raven Pro: Interactive Sound Analysis Software (Version 1.5). http://ravensoundsoftware.com/software/raven-pro/

B: Ofer Tchernichovski, http://soundanalysispro.com

C: Robert Lachlan, http://rflachlan.github.io/Luscinia/

D: Araya-Salas, M., & Smith-Vidaurre, G. (2017). warbleR: An R package to streamline analysis of animal acoustic signals. *Methods in Ecology and Evolution*, 8(2), 184–191. https://doi.org/10.1111/2041-210X.12624

E: Specht, R. (2004). Avisoft-SASLab Pro. *Avisoft, Berlin*. http://www.avisoft.com/sound-analysis.htm

4.9 References

Brunton, D. H., Evans, B. A., and Ji, W. (2008). Assessing natural dispersal of New Zealand bellbirds using song type and song playbacks. *N. Z. J. Ecol.* 32, 147–154.

Brunton, D. H., and Li, X. (2006). The song structure and seasonal patterns of vocal behavior of male and female bellbirds (*Anthornis melanura*). *J. Ethol.* 24, 17–25. doi:10.1007/s10164-005-0155-5.

Catchpole, C. K., and Slater, P. J. B. (2008). *Bird song: Biological themes and variations*. 2nd ed. New York, NY, US: Cambridge University Press doi:10.1017/CBO9780511754791.

Django Software Foundation (2013). Django: The Web framework for perfectionists with deadlines. *Djangoproject.Com*.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Heather, B. D., and Robertson, H. A. (2000). *The field guide to the birds of New Zealand.* Viking\Penguin Books.

Janik, V. M., and Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A* 199, 479–489. doi:10.1007/s00359-013-0817-7.

Jones, A., and Bijleveld, C. C. J. H. (2001). The interobserver reliability of scoring sonagrams by eye: A study on methods, illustrated on zebra finch songs. *Anim. Behav.* 62, 791–801. doi:10.1006/anbe.2001.1810.

Jones, E., Oliphant, T., and Peterson, P. (2014). SciPy: open source scientific tools for Python.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev.* 91, 13–52. doi:10.1111/brv.12160.

Lachlan, R. F., Verzijden, M. N., Bernard, C. S., Jonker, P.-P., Koese, B., Jaarsma, S., et al. (2013). The progressive loss of syntactical structure in bird song along an island colonization chain. *Curr. Biol.* 23, 1896–1901. doi:10.1016/j.cub.2013.07.057.

Nelson, D. A., Hallberg, K. I., and Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology* 110, 879–908.

Oliphant, T. E. (2007). Python for scientific computing. *Comput. Sci. Eng.*

Parker, K. A., Anderson, M. J., Jenkins, P. F., and Brunton, D. H. (2012). The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecol. Lett.* 15, 778–785. doi:10.1111/j.1461-0248.2012.01797.x.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2012). Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.*

Peters, S. S., Searcy, W. A., and Marler, P. (1980). Species song discrimination in choice experiments with territorial male swamp and song sparrows. *Anim. Behav.* 28, 393–404. doi:10.1016/S0003-3472(80)80048-0.

Priyadarshani, N., Marsland, S., and Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. *J. Avian Biol.* 49, jav-01447. doi:10.1111/jav.01447.

Rossum, G. (1995). Python Tutorial. Amsterdam, The Netherlands, The Netherlands: CWI (Centre for Mathematics and Computer Science).

RStudio Team (2015). *RStudio: integrated development environment for R*. Boston, MA: RStudio, Inc. Available at: http://www.rstudio.com/.

Sayigh, L. S., Esch, H. C., Wells, R. S., and Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Anim. Behav.* 74, 1631–1642. doi:10.1016/j.anbehav.2007.02.018.

Soltis, J. (2009). Vocal production, affect expression, and communicative function in African elephant rumble vocalizations. *J. Acoust. Soc. Am.* 126, 2210–2210. doi:10.1121/1.3248716.

Sperber, A. L., Werner, L. M., Kappeler, P. M., and Fichtel, C. (2017). Grunt to go—Vocal coordination of group movements in redfronted lemurs. *Ethology* 123, 894–905. doi:10.1111/eth.12663.

Starnberger, I., Preininger, D., and Hödl, W. (2014). From uni- to multimodality: towards an integrative view on anuran communication. *J. Comp. Physiol. A* 200, 777–787. doi:10.1007/s00359-014-0923-1.

Van Der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.*

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Mach. Learn.* 42, 31–60.

5 Cultural evolution of male and female song diversity in an archipelago metapopulation

5.1 Abstract

Cultural evolution is the gene-like transmission of cultural units or *memes* through populations. A significant example is birdsong, which in songbirds is learnt (i.e. culturally transmitted) rather than innate. The song profile of a songbird population changes as new memes (syllables) are introduced through immigration, copying errors, and innovation, while other syllables fall out of use. This leads to a diversification of the syllable pool across the species, much like the diversification and spatial patterns of human language. Dialects have been well studied in male songbirds but have been overlooked in females—how cultural evolution of female song compares to that of conspecific males is largely unknown. In this chapter, I undertake one of the first comparisons of male and female song culture in birds, analysing five years of song data from New Zealand bellbirds *Anthornis melanura*, spanning an archipelago metapopulation of six islands. Having classified 21,500 syllables, I compare population syllable repertoire sizes, prevalence, and overlap

between sites and sexes. I show that males and females—both with complex songs—have distinct song cultures, sharing only 6–26% of syllable types within each site. Furthermore, syllables shared between sexes can be statistically discriminated from sex-specific types based on acoustic properties. Most syllable types were site-specific, but for types shared between sites, sharing decreased with distance only for males. Also, despite females moving between sites more frequently, males had greater site–site repertoire overlap overall. These results provide some of the first evidence for female dialects, and show different cultural processes at play for the two sexes. I discuss the implications for future research on female culture.

5.2 Introduction

5.2.1 Cultural evolution

Culture is shared information or behaviour acquired through social learning from conspecifics (Dawkins, 1976). A unit of culture is an idea or behaviour that replicates through imitation by other individuals—for instance, a skill, melody, image, gesture, word or vocalisation (McNamara, 2011). Cultural units are termed *memes* (Dawkins, 1976), and share many parallels with genes (transmissible units of genetic information) in that both are replicated, can spread through a population, both are subject to mutation and are operated on by natural selection (Aplin, 2019). Thus the three prerequisites for a Darwinian evolutionary system—replication, variance and selection—are met.

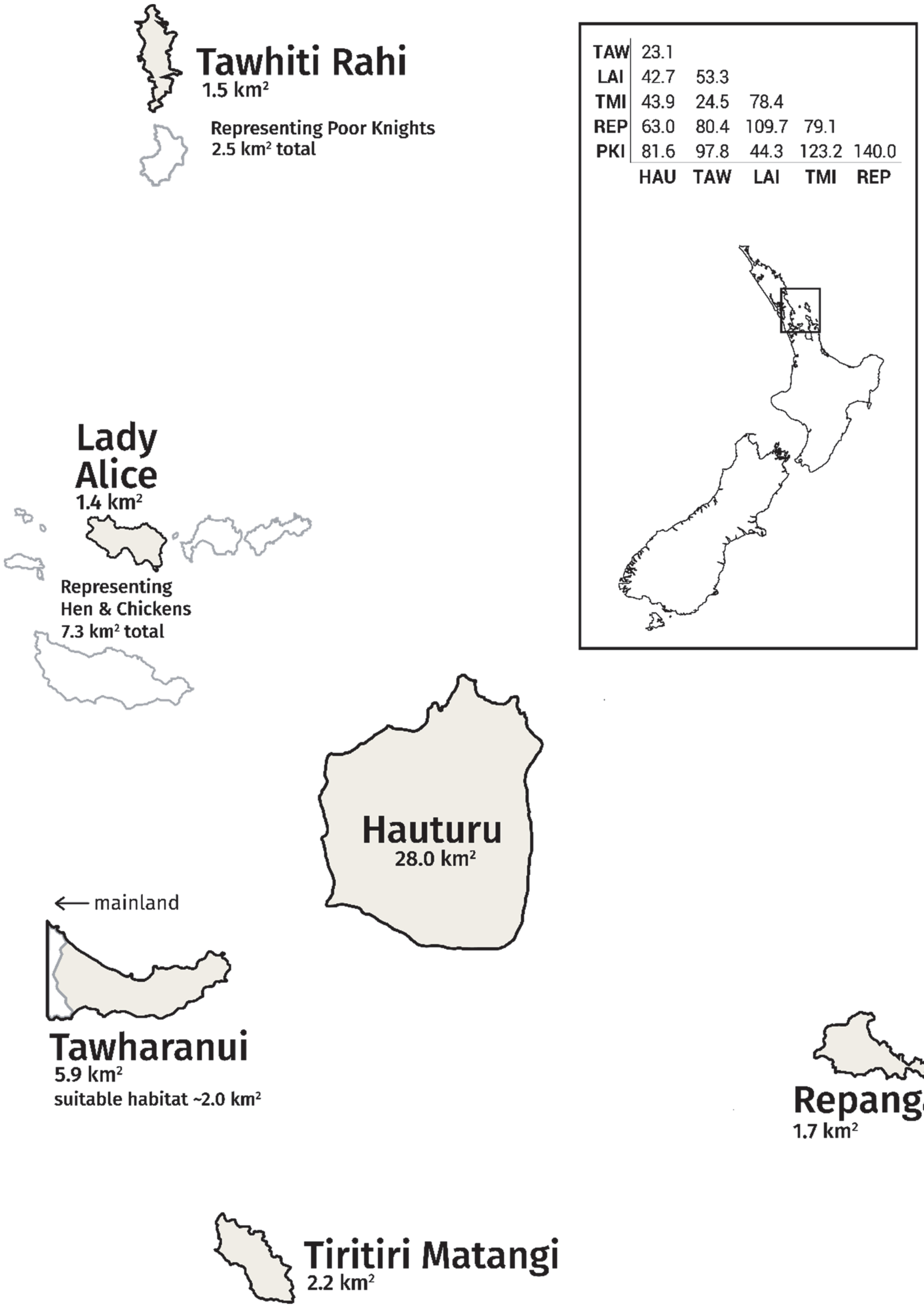
However, there are several notable differences between genes and memes (Claidière and André, 2011; Oren et al., 2018). First, in contrast to genetic transmission, which is vertical and bi-parental, cultural transmission can theoretically take place between any individuals (Payne et al., 1988; Lynch et al., 1989; Baker and Gammon, 2008). Memes are also far more prone to copy errors, and are subject to deliberate modifications and learning biases (Shennan, 2011). Because a behavioural innovation may quickly spread to be adopted by a whole population, memes frequently have much higher fecundity than genes (Whiten et al., 2011). Thus, cultural evolution is complex, and can be a rapid and powerful means of adaptation to environments, sometimes more so than genetic evolution (Perreault, 2012). It can be a key factor in speciation through behavioural isolation. For example, in birds, the cultural evolution of local song dialects and mating preferences for the local dialect can form mating barriers leading to speciation (Wilkins et al., 2013), which may have contributed to the radiation of the passerines (Podos and Warren, 2007). Because cultural evolution can profoundly affect population change and survival (Parker et al., 2012), understanding the dynamics of memes is of great importance for ecology and evolutionary biology.

Non-human animal culture is ubiquitous and taxonomically widespread (Brakes et al., 2019). However, *vocal* culture—the social learning of acoustic memes—has been observed only in the following taxa: songbirds (oscines; Passeri), parrots, hummingbirds, cetaceans, elephants, seals, bats, and humans (Paton et al., 1981; Baptista and Schuchmann, 1990; Janik and Slater, 1997; Poole et al., 2005; Sanvito et al., 2007; Catchpole and Slater, 2008). In these taxa, the song profile of a population changes as new memes are introduced through immigration, copying errors, and innovation, while other memes fall out of use (Catchpole and Slater, 2008). This leads to a diversification of the meme pool across the species, much like the diversification of human language—and often results in *dialects*. Dialects are characterised by uniformity of signal forms within given localities, with sharp geographic boundaries and discrete differences in signal forms between localities (Podos and Warren, 2007). Geographic dialects have been documented in most vocally learning taxa, but very extensively in songbirds (Jenkins, 1978; Whitehead and Rendell, 2014; Aplin, 2019).

Despite the high volume of studies on male song culture and dialects, virtually nothing is known about female song culture. This is partly due to a northern-hemisphere-biased view of sexual selection (Darwin, 1871) that emphasises male-male competition and female choice in driving elaborate traits in males, but which does not provide a framework for understanding elaborate female traits (Riebel et al., 2019). Female song (and other elaborate female traits) have been overlooked as non-functional aberrations, perhaps resulting 'accidentally' from shared genetic architecture with males (Darwin, 1871; Lande, 1980; see Tobias et al., 2012 for review). This view has now been roundly discounted. Female song is present in 64% of surveyed songbird species (Webb et al., 2016), has been recovered as the ancestral state (Odom et al., 2014), can evolve independently of the male song phenotype (Price, 2015) and has female-specific functions in territory and resource defence, mate attraction, mate defence, and pair bonding (Searcy and Yasukawa, 1995; Langmore et al., 1996; Rogers et al., 2007; Brunton et al., 2008b; Templeton et al., 2011; Hall et al., 2015).

Therefore, the paucity of information about female song culture is a major oversight and significant knowledge gap. The focus on males has impeded development of a more general theory that accounts for both sexes. Do female birds have dialects, and how does female song culture—and cultural evolution—compare to that of conspecific males? It is conceivable that sexual differences in dispersal and song learning modes could result in quite different geographic patterns for males and females, yet this possibility has hardly been investigated (Graham et al., 2017a).

The New Zealand bellbird, *Anthornis melanura*—hereafter ‘bellbird’—is an endemic honeyeater (Family Meliphagidae) with complex, geographically diverse song in both sexes. I have identified an archipelago metapopulation of bellbirds—spanning a network of six island and peninsula sites in the Hauraki Gulf, northeastern New Zealand—as an ideal system for studying cultural evolution of female song in comparison to males (Figure 5.1). With sites differing in size, bellbird population histories and connectivity to each other (see Chapter 3; Baillie et al., 2014), the influence of various spatial and social parameters governing song culture can be explored. Furthermore, because males and females share genetic and environmental backgrounds yet differ in dispersal behaviour, the system makes it theoretically possible to isolate the effect of dispersal frequency on meme dynamics.



	HAU	TAW	LAI	TMI	REP
TAW	23.1				
LAI	42.7	53.3			
TMI	43.9	24.5	78.4		
REP	63.0	80.4	109.7	79.1	
PKI	81.6	97.8	44.3	123.2	140.0

Figure 5.1—Simplified map of the Hauraki Gulf archipelago, home to a metapopulation of bellbirds.

Tawhiti Rahi represents the larger Poor Knights Islands group; likewise, Lady Alice Island represents the larger Hen and Chickens Islands group. The other islands of these wider groups are shown in white. The distance matrix between sites (measured in km) is provided (HAU, Hauturu; TAW, Tawharanui; LAI, Lady Alice; TMI, Tiritiri Matangi; REP, Repanga, PKI, Tawhiti Rahi). Distances not to scale on simplified map.

5.2.2 Sexual dimorphism in songbird song

Songbird species differ widely in their degree of sexual song dimorphism. In some species, repertoires show no sex-specificity; for example, males and females have identical repertoires in the forest weaver *Symplectes bicolor* (Wickler and Seibt, 1980), magpie lark *Grallina cyanoleuca* (Hall, 2000), and northern cardinal *Cardinalis cardinalis* (Yamaguchi, 1998). This lack of sex-specificity can be maintained if birds learn songs indiscriminately from males and females, as has been shown for cardinal chicks (Yamaguchi, 1998).

In other species, males and females are capable of singing the same song types, but are not constrained to be identical (Rose et al., 2018). For example, male and female North Island kokako *Callaeas wilsoni*, appear to share a common repertoire of phrases, but flexibly divide themes between mated pair members such that the male tends to sing some phrase types and the female other phrase types (Molles et al., 2006). Similarly, in rufous-naped wrens *Campylorhynchus rufinucha*, male and female phrase types sound similar and repertoires are similarly sized, but there are still sex-specific characteristics of singing behaviour (Bradley and Mennill, 2009). Yet other species—many duetting wrens, for instance—show completely non-overlapping repertoires between sexes (Brown and Lemon, 1979; Levin, 1996; Mann et al., 2009). Some species where pair-members duet have sex-specific contributions to the coordinated song. For instance, in the chirruping wedgebill *Psophodes cristatus*, the female inserts her single rising syllable into the gaps of the male's trisyllabic motif (Austin et al., 2019). In the closely related eastern whipbird *P. olivaceus*, males sing a pure-tone whistle culminating in a 'whip crack', to which females rejoin with two syllables, "*chew chew*" (Mennill and Rogers, 2006).

There have been very few studies comparing spatial variation of male and female songs (Graham et al., 2018a). Mennill and Rogers (2006) found that male whipbird song was highly consistent across the species' geographic range, whereas the female song showed more pronounced variation, falling into multiple distinct song types. The authors suggest that eastern whipbirds have undergone a decoupling of male and female song learning strategies in response to different sex-specific selection pressures.

In rufous-and-white wrens, song cultures of males and females appear to evolve in similar directions via acoustic adaptation and cultural drift (Graham et al., 2018a). However, taking a closer view of spatial-acoustic structure within a population, males and females show quite different processes (Graham et al., 2017a). The further males dispersed, the lower the song-sharing with their fathers, but there was no relationship between female dispersal distance and song sharing with their mothers—suggesting the importance of song matching differs between sexes (Graham et al., 2017a).

This growing body of research shows that female song may be similar to males in some species, but can sometimes be totally distinct and may sometimes be driven by different functions/selective pressures. It certainly is not typically 'simplified male song' as documented in some early accounts (Arcese et al., 1988). However, the knowledge is still fragmentary, and studies have mostly focused on within-pair repertoires in duetting species, or species with simple song. Outside of whipbirds and rufous-and-white wrens, I know of no studies to examine sexual dimorphism in song culture at a population level. The archipelago metapopulation of bellbirds described above provides a prime opportunity to continue filling this knowledge gap.

5.2.3 Distinctness of male and female repertoires in bellbirds

It has long been noted that bellbirds have sexually dimorphic song (Potts, 1882), but formal characterisation of bellbird song and sexual differences in singing behaviour was first undertaken as recently as 2006 (Brunton and Li, 2006). Examining bellbird song recordings taken March–December 2000 from Tiri, Brunton and Li (2006) found a number of differences in male and female song, as summarised in Table 5.1. Both sexes were found to sing year-round, with different seasonal peaks in singing activity. Males and females had different singing patterns and song structures and had distinct but partially overlapping syllable repertoires of comparable size. Though 'syllables' were classified differently and more coarsely than in the present thesis, the study nonetheless gives interesting insight into the sexual dimorphism of song characteristics in the Tiri population, and represents a quantitative and early contribution to the topic of sexual song dimorphism in birds. More research is needed to determine if Brunton and Li's picture of sexual dimorphism in repertoire holds when syllables are classified at a fine scale, and whether it remains consistent between sites and years.

Table 5.1—Key features of male and female bellbird song on Tiri, March–December 2000.

Results from Brunton and Li (2006). Note that song units were classified differently and at a coarser granularity than in this thesis.

Song feature	Male		Female
Seasonality	Sing year-round, but highest in March (end of breeding season) and August (start of breeding season).		Sing year-round, but highest in breeding season (August—December).
Song structure	10 song types, sung in repeated series (AAAA…) or with continuous variety (ABCADEFBD…)		5 song units, arranged variously to form songs (21 song types with 9 frequently used).
Singing pattern	Prolonged bouts up to 5 min without gaps between songs.		Discrete songs, with at least 3 sec between songs.
Daily variation	Repeated-series singing more frequent in morning, Continuous-variety singing more frequent in afternoon.		None noted.
Counter-singing	Yes, synchronised with other males (repeated-series song).		Yes, with other females, but not synchronised.
Number of ‘syllable’ types	24	7 of these types shared	19*
Songs sung per 30 min	27.6 ± 13.9	Significantly different $P < 0.01$	11.8 ± 12.9
Song duration (s)	2.17 ± 0.09	Significantly different $P < 0.04$	1.93 ± 0.03
Mean pitch (kHz)	14.25 ± 1.04	Significantly different $P < 0.001$	17.38 ± 1.88
Mean entropy (randomness)	-3.44 ± 0.07	Significantly different $P < 0.001$	-4.07 ± 0.04

*Corrected from 12 to 19 by authors

5.2.4 Bellbird population genetics and dispersal

Genetic connectivity

Dispersal is a critical aspect of ecology because it has profound effects on the genetic and phenotypic structure of animal populations (Bohonak, 1999; Graham et al., 2017a). Studies on bellbird gene flow (Baillie, 2011; Baillie et al., 2014) show ongoing dispersal on an 'isolation-by-distance' basis at the scale of the entire country, with higher gene flow between closer sites (Figure 5.2). This implies that bellbirds would have dispersed continuously through the New Zealand landscape, at least in recent history (Baillie et al., 2014).

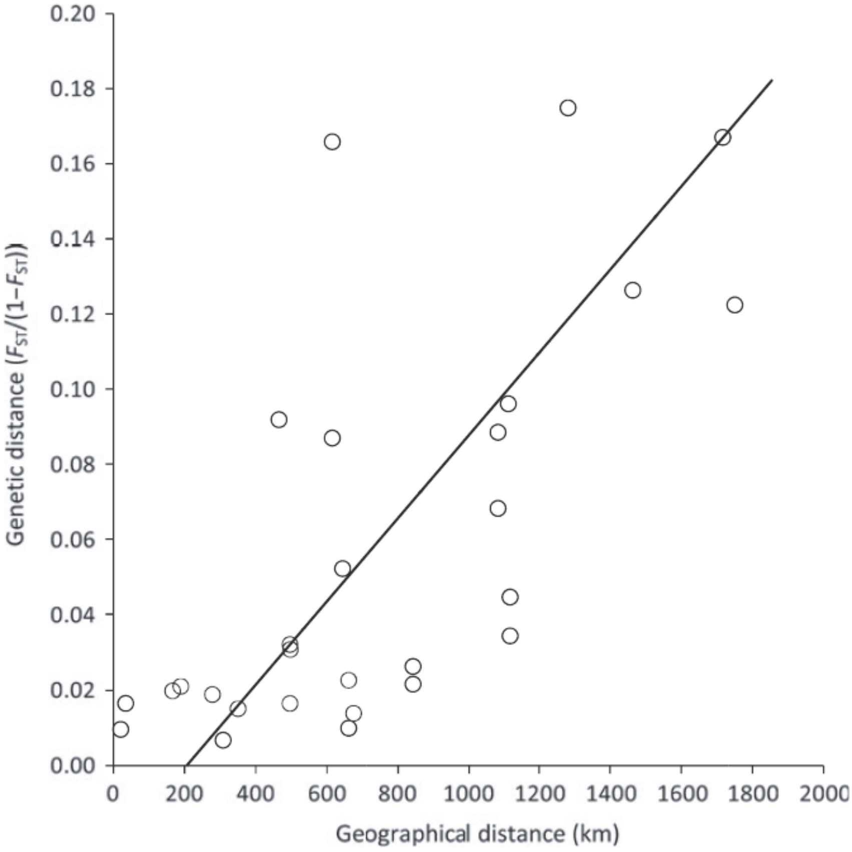


Figure 5.2—The relationship between genetic distance (Rousset 1997) and geographical distance for New Zealand bellbirds, at the scale of the entire country.

From Baillie et al. (2014); details therein. Strong evidence of 'isolation-by-distance' was found. Note that distances between the 6 sites in our Hauraki metapopulation are all <150 km (bottom left-hand corner of the graph).

Genetic sampling on islands also showed that distances of open water up to *at least* 35 km did not pose strong barriers to gene flow, implying frequent open-water traversal. Genetic mixing is apparent for most of Baillie's sites in Figure 5.3. Three sites in Baillie's sample that are genetically more distinct are the Poor Knights Islands, Tiri, and the Auckland Islands. For the Poor Knights, genetic distinctness is likely due to geographic isolation from other bellbird populations, and possibly an historic founder event (Baillie, 2011) or drought-induced population bottleneck, given the harsh climate and paucity of fresh water (Buddle, 1941; Bartle and Sagar, 1987). The genetic distinctness of Tiri results mostly from deforestation reducing the population to a bottleneck prior to restoration efforts in 1978 (Anderson and Craig, 2003; see Chapter 3, Section 3.2 for population histories). Now that Tiri is reforested and sugar-water feeders have propelled the population to carrying capacity, it may be difficult for immigrants to the island to establish territories and mates; failure to integrate into the population could maintain genetic differentiation even with high levels of dispersal to Tiri (Baillie, 2011, p. 56). The genetic distinctness of the subantarctic Auckland Islands is best explained by initially small founder population(s) of vagrant bellbirds, which then expanded (Baillie, 2011), but with strong genetic drift due to geographic isolation (465 km south of mainland New Zealand).

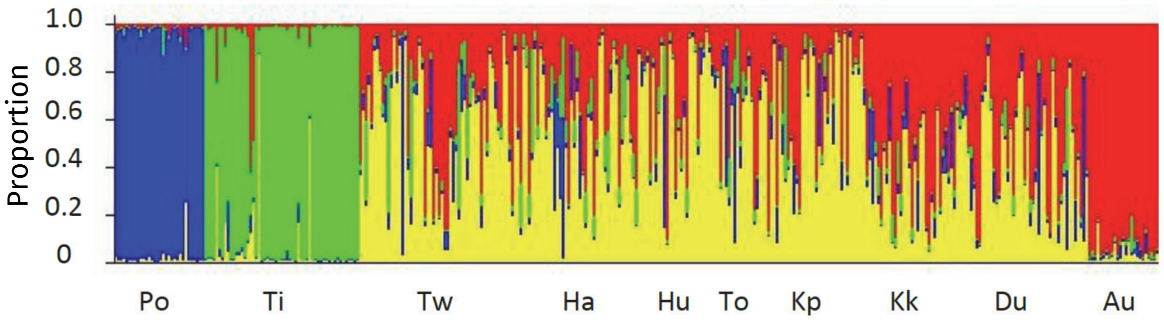


Figure 5.3—Genetic profiles of *A. melanura* populations as based on eight microsatellite markers.

From Baillie (2011); details therein. *A. melanura* populations from throughout their geographical range, from Poor Knights Islands in the north to the Auckland Islands in the south, genetically cluster into four 'haplogroups'. Note that most sites are strongly mixed genetically, except for Poor Knights (Po), Tiri (Ti), and Auckland Islands (Au). Other sampling locations are abbreviated as follows: Tawharanui, Tw; Hauturu, Ha; Hunua, Hu; Tongariro, To; Kapiti, Kp; Kaikoura, Kk; Dunedin, Du.

Examining dispersal within the Hauraki Gulf region, Baillie found that 3–10% of each population were first generation migrants, except for Poor Knights Islands and Hauturu, where no first-generation migrants were detected (Baillie, 2011 p. 52). Lack of detected migrants in the Poor Knights is unsurprising given the relative geographic isolation. However, in the large, genetically diverse population of Hauturu, gene frequencies of migrants and natives may well have been too similar to discern between them, rather than reflecting a true lack of immigration (Baillie, 2011 p. 58).

Together, then, all six sites in the Hauraki Gulf archipelago appear within range of feasible dispersal for bellbirds, with geographic isolation regulating the degree of connectivity.

Evidence supporting female-biased dispersal

Bellbird social ecology involves vigorous resource defence (Craig and Douglas, 1984, 1986; Rasch and Craig, 1988). Typically, bird species with resource defence systems have philopatric males defending local resources in order to attract mates and breed successfully, whereas unpaired females are free to disperse (Greenwood, 1980; Clarke et al., 1997; Paris et al., 2016). Given the general female-biased dispersal pattern seen in songbirds it is parsimonious to assume female-biased dispersal for bellbirds.

Empirical evidence strengthens this assumption. In her study of bellbird population connectivity across New Zealand, Baillie (2011 p. 71) looked for first-generation migrants within five sites in the Hauraki Gulf and found a notable female bias: a total of 12 first-generation migrants were detected, of which eight were female, one was male, and three were of unknown sex.

What form does female-biased dispersal take in bellbirds—a bias in dispersal distance, dispersal frequency, or both? Bellbirds are present on remote islands such as the Auckland Islands, 465 kilometres south of the New Zealand's South Island. The Auckland Islands are of volcanic origin (Vitt, 1979) and have never been closer to mainland NZ; thus, direct flight of 465 kilometres from the mainland is the only plausible explanation for bellbird occurrence there. Clearly, then, both sexes of bellbird are strong fliers capable of crossing hundreds of kilometres of open water, even if rarely. Furthermore, in an unsuitable environment *both* sexes quickly disperse away, as has occurred after all nine attempts at bellbird translocations 1931–2012 (Miskelly, 2013). Together, these facts suggest that female-biased dispersal in bellbirds results not from males being limited in dispersal ability or dispersal distance compared to females, but rather from a higher *frequency* of dispersal for females. That is, at a given site with suitable bellbird habitat (like the sites in our sample), a female has a higher probability of emigrating than a male.

Age and timing of dispersal

In February 2005, Tawharanui (hitherto bellbird-free) was colonised by a flock of *c.* 100 bellbirds from Hauturu, at the end of the breeding season (Brunton et al., 2008a). The flock was comprised of a mixture of males and females, some with juvenile plumage and others with adult plumage. Song recordings confirmed that many were indeed adults with crystallised song closely matching that of the Hauturu population (Brunton et al., 2008a). Aside from this dispersal event, single bellbirds have also been observed flying over open water between islands in the Hauraki Gulf (pers. obs. Dianne Brunton). These observations support the prior conjecture that both adults and juveniles can disperse, and may do so as a flock or as individuals. Together, these facts represent all that is currently known about the age and timing of bellbird dispersal.

5.2.5 Bellbird song learning strategies

Vocally-learning birds show a diversity of song learning strategies (Beecher and Brenowitz, 2005) but can be broadly categorised into two groups with respect to the timing of their song learning. Species with 'closed-ended' learning have a single sensitive period (critical period of learning; Nottebohm, 1969) in the first year of life to acquire acoustic memes, after which their repertoire is fixed (e.g. song sparrows *Melospiza melodia*; Nordby et al., 2002). By contrast, 'open-ended' learners can learn new acoustic memes throughout their lives (e.g. nightingales *Luscinia megarhynchos*; Todt and Geberzahn, 2003). There is much yet unknown about bellbird life history and ecology (Roper, 2018) and it has not been experimentally tested whether bellbirds are open-ended learners, able to adopt new memes into adulthood. Tutoring experiments would be ideal for testing this question, but currently we must rely on anecdotes and observational evidence, summarised as follows.

Dianne Brunton played bellbird song to a young male bellbird in the aviary at Auckland Zoo. The male was the only bellbird in the aviary and had hitherto sung only tui-like song. Despite being past the typical age of bellbird song 'crystallisation' (Table 5.2), he successfully learnt the playback song and sang only bellbird song from then on (pers. comm. Dianne Brunton). This gives an indication that receptivity to acquire new memes may persist into adulthood, but for what timeframe is unclear, especially given that this bird was in unnatural circumstances and so may not reflect species-typical behaviour. In her study of bellbird song development on Tiri, Michelle Roper found that adult individuals can drop syllable types from the repertoire for a year or two, then re-incorporate those types (Roper, 2018). This shows that the repertoire is under dynamic control into adulthood (the adult repertoire is not 'fixed'), but again is not strong evidence that adults acquire *new* types; they may be cycling already-learnt types.

Table 5.2—The timings of song development phases for male and female NZ bellbirds. From Roper (2018).

Song developmental phase	Male	Female
Onset age of subsong	3 weeks	3 weeks
Onset age of plastic song	13 weeks	15 weeks
Onset age of crystallisation	24 weeks	24 weeks

However, given that tui *Prosthemadera novaeseelandiae* are open-ended vocal learners capable of mimicry throughout their lives (Robertson, 1996; Zwartz, 2018), it does increase the likelihood that bellbirds, as their closest relatives (Driskell et al., 2007), are open-ended too. Furthermore, bellbird social ecology indicates there may be strong selective pressure for open-ended learning. Both males and females frequently engage in chorusing with conspecifics of the same sex. For males this involves highly synchronised group singing, where males may attempt to sing the same syllables at precisely the same time, sometimes synchronising syntax and tempo so closely that they sound like a single bird. Females also chorus with each other using matching song types (Brunton et al., 2008b), and on some islands have been observed matching syllables in synchrony, as males do (e.g. Poor Knights Islands; pers. obs.). Considering the prevalence and aggression of chorusing behaviour, the ability to match local syllable types seems to be of great importance to bellbirds; a bird unable to learn local syllable types would be unable to engage in the synchronised chorusing bouts that commonly occur at foraging sites and presumably suffer social costs as a result. Dispersing adults would likely be unable to integrate into new populations if not for the ability to quickly learn all or part of the local dialect, and so the social ecology of the species makes open-ended learning plausible.

Predicted patterns of meme sharing under open-ended versus closed-ended learning scenarios

Given that there is detectable migration occurring between populations, certain patterns can be expected under open-ended or closed-ended learning scenarios.

(1) Closed-ended learning

If bellbirds are closed-ended learners, then any adults dispersing to a new site will continue to sing the memes of their previous population (and will be unable to learn the memes of the new population). The greater the population connectivity between the two sites, the more the two population repertoires should resemble each other. Thus, we would expect an acoustic isolation-by-distance scenario, where closer (higher-connectivity) sites share more memes than distant sites. Juvenile immigrants may sing some combination of memes from the previous population and from the new, depending on the age of dispersal relative to the phase of song learning. In this closed-learner scenario, the song culture within sites could evolve through two means: (i) influx of new memes carried by immigrant birds (these memes may remain ‘sterile’ or may be transmitted to learning-phase juveniles), and (ii) innovation by learning-phase juveniles, with the new memes retained into adulthood (and eventually possibly transmitted to juveniles). Adults would retain fixed repertoires throughout their lives (up to ~8 years; Spurr, 2008) and so the rate of cultural change would be relatively slow, especially when emigration/immigration is low.

(2) Open-ended learning—singing only local memes after immigrating

Under an open-ended learning scenario, a range of outcomes is possible. If immigrants completely cease expressing the memes of their source population and adopt the memes of the new site instead, then over time population repertoires will become sharply distinct. There should be very low meme repertoire overlap between sites, with no relationship between site connectivity and degree of meme repertoire overlap between sites, as dispersers effectively leave their memes behind when they move between sites.

(3) Open-ended learning—singing a combination of source and local memes after immigrating

Less extremely than (2), it is possible that immigrants continue to sing *some* of their source population memes but *augment* their repertoire with local population memes upon arrival. This could occur if there are 'local only' memes that incur high aggression at new sites, and other 'widely acceptable' memes that can be sung freely at multiple sites, for example. In this case we would expect an acoustic isolation-by-distance pattern as for (1) above, but weaker because only some source-population memes are retained after arrival.

In open-learner scenarios, the song culture within sites could evolve through two means: (i) innovation (purposeful or accidental) of new memes by both juveniles and adults, with possible spread through the population (transmission to both juveniles and adults); (ii) influx of new memes carried by immigrant birds, with possible spread through the population. Adults would be free to alter their repertoires throughout their lives and so the rate of cultural change could be rapid. The rate of change would be driven mostly by how strongly the culture is enforced through aggression and sexual selection, rather than turnover of individuals.

5.2.6 Aims and questions

Cultural evolution is both an individual-level and population-level phenomenon, and the two scales provide information on different mechanisms. This chapter focuses on population-level dynamics of cultural evolution using multiple islands within an archipelago as a model system. The aim of this chapter is to address three questions regarding population-level differences between male and female song:

> (1) Are there sexual differences in population syllable repertoire size or meme prevalence?
>
> (2) How much do male and female population repertoires overlap?
>
> (3) What is the pattern of meme sharing between sites, and does it differ for males and females?

These questions are detailed below.

(1) Are there sexual differences in population syllable repertoire size or meme prevalence?

Sexual dimorphism in repertoire size is of considerable interest, because it reflects (at least in part) the differences in selection pressures for the sexes—a topic that remains severely understudied (Tobias et al., 2012; Graham et al., 2018a). Species where both sexes sing complex song provide important opportunities for examining the drivers of sexual dimorphism in song structure, singing behaviour and song complexity (Graham et al., 2017a; Riebel et al., 2019). How does the syllable repertoire size of the female population compare to that of males, at each site?

Secondly, repertoire size is not the whole story; dimorphism in type *prevalence* is also important because it provides information about how the repertoire is used (Laland, 2004). Are all types sung equally often, or do we find a bias with some types prevalent and others rare? And do the prevalence distributions differ between the sexes?

(2) How much do male and female population repertoires overlap?

While many studies have evaluated male and female repertoires in duetting pair members, few have yet characterised the degree of male and female repertoire overlap at a population level (Bradley and Mennill, 2009; Graham et al., 2017a) or in non-duetting species. A detailed analysis of male and female repertoire overlap in bellbirds would provide a foundation for understanding how cultural evolution operates in the two sexes. Therefore, here I test the following:

(i) to what extent do the syllable repertoires of male and female population sectors overlap within sites? In other words, is the female set of types separate from, partially overlapping with, or a fully overlapping subset of male types? I examine this at both fine- and broad-scale classification levels across sites and on Tiri across years.

(ii) Are there features of shared types (i.e. common to both sexes) that distinguish them from sex-specific types? I test two potential predictors of sharing: syllable type prevalence in the population, and acoustic properties. The two hypotheses are not mutually exclusive and are detailed as follows.

Hypothesis 1: Prevalence in the population

It is possible that there is a link between the prevalence of syllable types and whether they are common to both sexes. For example, females may be more likely to learn syllable types from males (or males from females) because of frequent exposure (Mennill et al., 2018)—or, conversely, they may prefer to learn rare types (Laland, 2004). Prevalence in the population could be measured in several ways; here I use the proportion of songs in which the type occurs (i.e. *support*, ranging from 0 to 1; a support of 0.06 means it occurs in 6% of songs in the focal population). I predict that if there is a link between intersexual sharing and type prevalence, then the mean support value of types common to both sexes should be significantly different from that of sex-specific types.

Hypothesis 2: Acoustic properties

It is possible that the acoustic properties of a syllable type determine its likelihood of being sung by both sexes. If this is so, we would expect to be able to distinguish sex-specific types from types common to both sexes on the basis of acoustic properties, e.g. contour, pitch range, timbre, or some other combination of acoustic traits. Visualising male and female syllables as an ordination in acoustic space, we would expect male and female syllables to occupy largely distinct ranges with shared syllables lying in between (Figure 5.4A). Alternatively, if the acoustic properties do not determine likelihood of intersexual sharing, we should not be able to distinguish sex-specific and shared types; in an ordination, male and female syllables might occupy the same ranges of acoustic space, with pockets of sex-specific and shared syllables interspersed randomly throughout (Figure 5.4B).

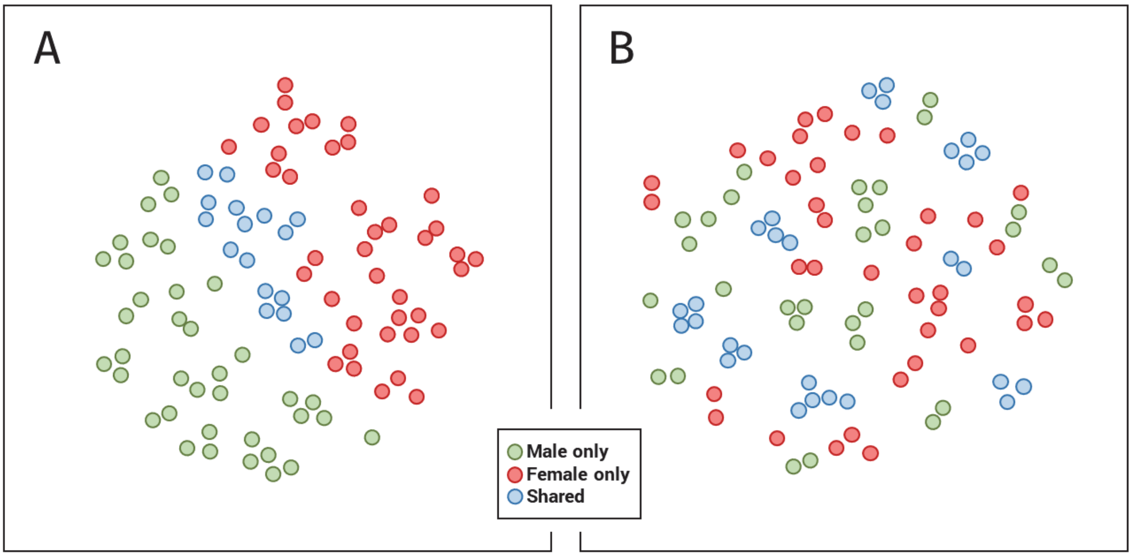


Figure 5.4—Hypothesised ordinations of song syllables in acoustic space.

There are two general possibilities for where shared syllables fall in acoustic space: (**A**) Male and female syllables occupy broadly separate ranges of acoustic space, with shared syllables in a constrained region somewhere between the two clusters. (**B**) male and female syllables occupy the same overall range of acoustic space, with localised separation; shared syllables are interspersed throughout.

(3) What is the pattern of meme sharing between sites?

Given the predicted patterns of meme sharing under open-ended vs closed-ended learning scenarios outlined earlier (p. 134), I examine which of the hypothesised patterns best fits the data.

(i) The large size and geographic centrality of the Hauturu population make it likely to be a substantial source of dispersing bellbirds to other sites in the Hauraki Gulf archipelago. Do bellbird populations at greater distances from Hauturu have lower repertoire overlap with Hauturu? If so, are syllable types lost *progressively* with distance from Hauturu, suggesting that islands radiating from Hauturu act as a chain of 'stepping stones' for dispersal?

(ii) Does geographic proximity drive patterns of meme sharing between islands, regardless of island size? If so, I predict that sites that are closer together will share more memes (i.e. an acoustic isolation-by-distance pattern).

(iii) Females have higher dispersal frequency between sites than males. Based on the assumption that this results in higher flow of female memes between sites, I predict higher inter-site sharing for females than males.

5.3 Methods

Recording trips were conducted to six island and mainland peninsula sites in the Hauraki Gulf 2013–2017, involving substantial logistical challenges (recruiting volunteers, obtaining iwi consent, research and landing permits, bio-quarantine, transport, food and shelter arrangements). Additional sites were explored but had unsuitably low bellbird population densities and so were excluded. A team of 1–5 recordists per site spent a grand total of over 1,000 recordist-hours actively tracking and recording wild bellbirds, with Marantz PMD661 portable solid-state recorders paired with handheld Sennheiser ME-66 shotgun microphones. Recordists coordinated movement to sample each site systematically and with maximal coverage, gathering a grand total of 2,137 high-quality recorded songs (discrete vocalisation bouts by individuals). I processed these into a database of 21,500 song syllables. Syllables were classified into 780 fine-scale syllable types, which were then grouped into 39 broad-scale ‘families’ using new bioacoustics software, *Koe* (introduced in Chapter 4). Refer to Chapter 3 for more details on field methods and creation of the song database.

5.3.1 Sexual differences in population syllable repertoire size and meme prevalence

Assessing overall male and female syllable diversity

A complete catalogue of male and female types was produced by exporting the *Koe* unit table for the *Bellbirds* dataset and arranging in a *Microsoft Excel* (Office 365) Pivot Table to show the list of types with spectrograms, and counts from each site and sex (Appendix 7.3.1).

Extrapolating diversity estimates for each site

The raw number of types found at each site are not directly comparable, due to the inevitable confounding effects of different sampling effort. To account for this, I used the statistical software *EstimateS* (Colwell, 2013) to produce syllable-type accumulation curves and estimate syllable-type richness, the true number of syllable types at each site (*S*). I did the following for both fine-scale and broad-scale classification data.

First, unit table data were exported from *Koe*. The data were arranged in an *Excel* Pivot Table to show counts of each syllable type, with syllable types as 'species' (columns) and songs as 'samples' (rows). In *EstimateS* I chose "sample-based incidence or abundance data", with "one set of replicated sampling units (classic EstimateS input)"—samples as rows, with 'species' (syllable types) as columns. Under "sample order randomization for estimators and indices" I used the default of 100 runs (the number of randomizations). Under "Extrapolation of rarefaction curves" I chose to extrapolate to a total of 1000 samples (songs). Under "Estimation points (knots) for rarefaction and extrapolation" I chose to estimate at every sample. Other settings were left at their default values. Rarefaction curve graphs were then produced in *Excel* from the *S(est)* values column, with *S(est) 95% CI Lower Bound* and *S(est) 95% CI Upper Bound* columns providing the 95% confidence interval for the estimate, with the *x* axis displaying values for the first 500 songs.

Calculating prevalence of syllable types

Support is a measure of prevalence (Zaki, 2001). It is simply the proportion of songs in which the syllable type occurs (e.g. support of 0.06 means it occurs in 6% of songs in the population). There are other ways of measuring prevalence, such as the absolute number of occurrences of a syllable—but support is a more appropriate measure of population-level prevalence because it is not unduly influenced by a syllable type dominating one or a few songs. Song sequence data for all songs were exported from *Koe*. For each site-and-sex group, the support of each type was calculated by dividing the number of songs containing the type by the total number of songs, using *Excel*. Distributions of support were then plotted using *RStudio.*

5.3.2 Repertoire overlap of male and female population sectors

Extracting lists of syllable types

The lists of syllable types sung by each sex at each site were extracted by exporting the *Koe* unit table for the complete *Bellbirds* dataset and filtering by site and sex in *Excel.* For robustness I filtered the resultant lists to exclude syllable types with fewer than 3 occurrences within that site and sex. I did this at the fine-scale and broad-scale classification level.

Finding syllable types sung by both sexes at each site

In *R Studio* (RStudio Team, 2015) I used the *intersect* function to find the intersection of male and female lists at each site—that is, the types sung by both sexes. To visualise the degree of overlap I plotted the data as horizontal bars with length proportional to repertoire size and degree of bar overlap proportional to the number of types shared (Figure 5.7). For each site I calculated the percentage repertoire overlap according to the Jaccard Similarity Index (Hamers et al., 1989): the number of shared types divided by the total number of types.

Characterising shared versus sex-specific memes: Testing the prevalence hypothesis

For each site-and-sex group, at both the fine and broad scales of classification, I generated lists of shared (i.e. common to both sexes) and sex-specific syllable types and their support values. For robustness, for each site-and-sex group I first removed all syllable types with only one occurrence in the group. I then performed a two-tailed t-test for a difference in mean support between sex-specific and shared types. The tests assumed equal variances between the two groups and were conducted using the Analysis Toolpak in *Excel.* I chose a two-tailed test because I didn't have a hypothesised direction of difference *a priori*; a significant difference in means in either direction would indicate a link between prevalence and intersexual sharing.

Characterising shared versus sex-specific memes: Testing the acoustic properties hypothesis

In *Koe* I extracted the entire available suite of acoustic features for units. Feature measurements were exported, combined with corresponding *Koe* unit table metadata in *Excel*, and normalised. For each site, units were labelled as shared (sung by both sexes), male-only or female-only, and a two-dimensional Linear

Discriminant Analysis (LDA) was performed on the normalised measurements to test whether these three groups can be statistically discriminated. I performed this for each site separately and for the entire metapopulation pooled, at the fine-scale classification level.

5.3.3 Assessing the pattern of meme sharing between sites

Calculating site–site repertoire overlaps

Similarly to finding male–female repertoire overlaps above, I used the *intersect* function in *R Studio* (RStudio Team, 2015) to find repertoire overlaps between sites, for all pairwise combinations of site lists—comparing males against males, females against females. Each list excluded types with fewer than three occurrences for that sex at that site. This filtering was applied independently at the fine scale and broad (family-level) scale of classification; that is, broad-scale families with three or more occurrences were included even if comprised of fine-scale types with fewer than three occurrences each. I produced four matrices: for males at fine-scale classification, females at fine-scale, males at broad-scale, females at broad-scale. The matrices show the degree of overlap for each pairwise comparison of sites. In each cell of the matrix, I plotted the data as horizontal bars with length proportional to repertoire size and degree of bar overlap proportional to the number of types shared, displaying overlap percentage.

Examining repertoire overlap versus distance

To examine whether there is an isolation-by-distance pattern in repertoire overlaps between sites, I first plotted population repertoire overlap with Hauturu against geographic distance from Hauturu and calculated Spearman rank correlation using the spreadsheet available at:

www.biostathandbook.com/spearman.html.

In *Excel* I displayed the lists of types shared with Hauturu, to examine whether sites at increasing distances from Hauturu shared the same subset of syllable types progressively diminishing with distance.

I then plotted percentage of site–site repertoire overlap against site–site geographic distance for *all* pairwise combinations of sites in the archipelago, and calculated Spearman rank correlation to test for an isolation-by-distance pattern. In this case, to account for non-independence of points (being pairwise combinations), I used the RELATE routine in *PRIMER* (version 7, Clarke and Gorley, 2015) with 9999 permutations.

Testing for sex differences in degree of site–site repertoire overlap

I used the Sign Test (Dixon and Mood, 1946) to test for a sex difference in degree of site–site repertoire overlap, as a non-parametric alternative to a paired t-test. To do this, for each pair of sites I compared the degree of male–male repertoire overlap with the degree of female–female repertoire overlap. If the male–male overlap was greater, the comparison was marked as ‘+’. If the female–female overlap was greater, the comparison was marked as ‘–’.

If overall there is no difference in the degree of site-site repertoire overlap between sexes (the null hypothesis), then the counts of positive and negative signs are expected to be equal. If males tend to have greater overlap, there will be significantly more positive signs than negative, whereas if females tend to have greater overlap, there will be significantly more negative signs than positive.

5.4 Results

5.4.1 Sexual differences in population repertoire size and meme prevalence

Male and female syllable diversity

My dataset of acoustic recordings was collected from six locations. However, by considering the three years on Tiri both separately and together, there are nine groups for analysis. These are: Tawhiti Rahi, Lady Alice, Hauturu, Tawharanui, Repanga, Tiri 2013, Tiri 2014, Tiri 2015, and Tiri with all years pooled.

Syllable type accumulation curves at fine-scale and broad-scale classification levels show that by 500 songs, the estimated number of syllable types appears to level off nearly or completely for all groups (i.e. approximating the asymptotic syllable richness estimate) and is thus comparable between groups. The one exception is the female Tiri 2013 group at broad-scale classification level, which had insufficient data for correct calculation and thus is excluded from comparisons. An example accumulation curve is shown in Figure 5.5; see Appendix 7.3.2 for all accumulation curves.

At fine-scale classification, within sites, male point estimates are higher than females in seven out of nine cases, with the exceptions of Hauturu and Repanga where estimates are equal between sexes (Table 5.3). However, 95% confidence intervals overlap for males and females in all cases other than Tiri 2015, giving no evidence of a difference in repertoire size for males versus females within sites.

There appears to be high variability between sites for both sexes, with male point estimates ranging between 80 (Repanga) and 186 (Hauturu) and almost identically for females—between 81 (Repanga) and 186 (Hauturu). However, between *years* within one site (Tiri) we find higher variation for males: 92 (2014) to 168 (2015), compared to females: 81 (2014) to 93 (2015). Pooling all years, the estimated Tiri male population diversity was 165, versus 145 for females. We get an indication of the rate of type turnover by comparing the repertoire estimates for each year on Tiri with the estimate for all years pooled. We can infer low turnover of fine-scale types between years for males, because 165 for all years pooled is no bigger than the estimate for Tiri 2015 alone. We can infer higher turnover for females, as 145 for all years pooled is considerably bigger than the estimate for any single year on Tiri.

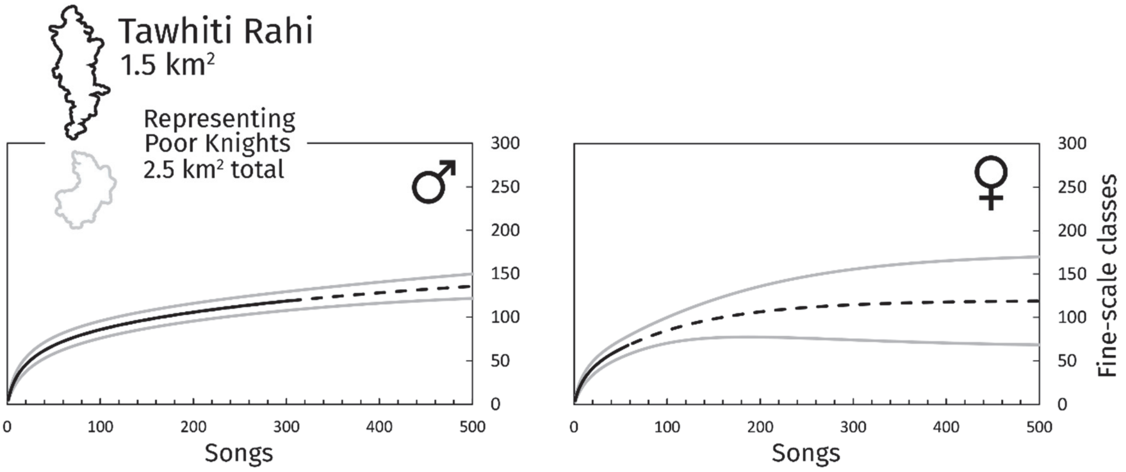


Figure 5.5—An example syllable type accumulation curve, for males and females at Tawhiti Rahi (fine-scale classification).

The black inner line shows the estimated true number of fine-scale types (*S*) for the number of songs on the *x* axis. The dashed region of the line indicates the extrapolated estimate beyond the sample data. The grey lines show the 95% confidence interval for the estimate. Calculated with the statistical software *Estimate S* (Colwell, 2013) using 100 runs.

Table 5.3—Extrapolated estimates of male and female syllable diversity within sites.

Estimated true number of fine-scale and broad-scale types at 500 songs, for males and females at each site. Calculated with the statistical software *Estimate S* (Colwell, 2013) using 100 runs.

Site	Sample size (songs)		Fine-scale types (estimate and CI)		Broad-scale types (estimate and CI)	
	Male	Female	Male	Female	Male	Female
Tawhiti Rahi	306	58	**135** (121–149)	**119** (68–169)	**27** (22–32)	**23** (10–37)
Lady Alice	181	125	**143** (118–169)	**128** (93–163)	**22** (22–22)	**21** (9–32)
Hauturu	67	98	**186** (139–234)	**186** (151–220)	**24** (20–29)	**18** (18–18)
Tawharanui	204	37	**116** (103–129)	**89** (48–130)	**27** (20–33)	**38** (8–68)
Repanga	328	196	**80** (72–89)	**81** (65–97)	**16** (13–18)	**14** (11–17)
Tiri 2013	206	34	**119** (102–137)	**91** (59–124)	**21** (19–24)	*Insufficient data**
Tiri 2014	70	37	**92** (76–108)	**81** (53–110)	**20** (20–20)	**22** (8–36)
Tiri 2015	113	66	**168** (126–210)	**93** (66–120)	**21** (21–21)	**20** (10–30)
Tiri 2013, 2014, 2015	389	138	**165** (152–178)	**145** (115–175)	**21** (21–21)	**32** (18–47)

*This estimate could not be correctly calculated due to insufficient sample size.

Syllable family accumulation curves (i.e. broad-scale classification; see Appendix 7.3.2) have much more variable confidence intervals than the fine-scale estimates, with five groups showing complete convergence (e.g. females at Hauturu) and others very broad intervals (most extremely, females at Tawharanui). Within sites, male point estimates are higher than females in 5 out of 8 cases: Poor Knights, Lady Alice, Hauturu, Repanga, Tiri 2015. Point estimates were higher for females at Tawharanui, Tiri 2014, and Tiri all years combined. However, 95% confidence intervals overlap for males and females in all cases other than Hauturu, giving no evidence of a difference in syllable family repertoire size for males versus females (within sites). There appears to be high variability between sites for both sexes, with male point estimates ranging between 15 (Repanga) and 25 (Tawharanui) and female estimates between 14 (Repanga) and 37 (Hauturu). Between *years* within Tiri, male syllable family point estimates were stable, ranging from 20 families (2014) to 21 (2013 and 2015). Furthermore, for all three years combined the estimated Tiri male diversity remained at 21 families, suggesting low turnover of male families between years. Female family variability cannot be accurately ascertained, as the 2013 estimate is invalid (due to insufficient data for calculation) and estimates for the other two years have wide confidence intervals.

Prevalence of types

Overall, male and female prevalence distributions were similar in shape. At the fine scale of classification, all male and female distributions resembled a negative exponential distribution, with many rare types and progressively fewer types at higher levels of prevalence (Figure 5.6). For all distributions, the large number of rare types was comprised mostly of types from the *Pipe* family, and the *Pipe*-related *Stepup* and *Stepdown* families. By contrast, types of higher prevalence were from various families. At several sites, females had 1–4 types of exceptionally high prevalence, extending the upper tail beyond that of the male distribution. These outlier female types were as follows:

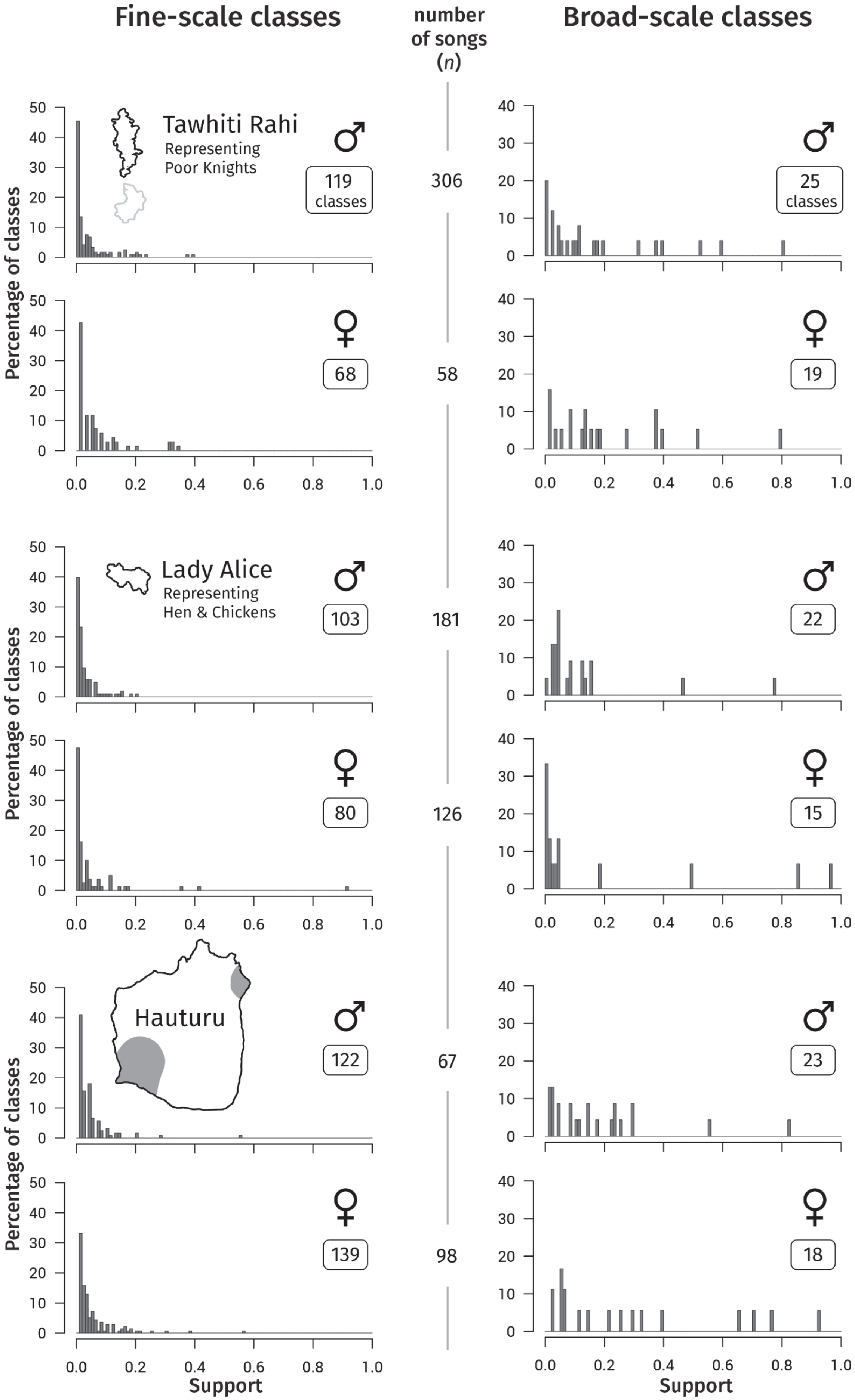
Lady Alice: *Downsweep(roughtosmooth)* (support = 0.91).

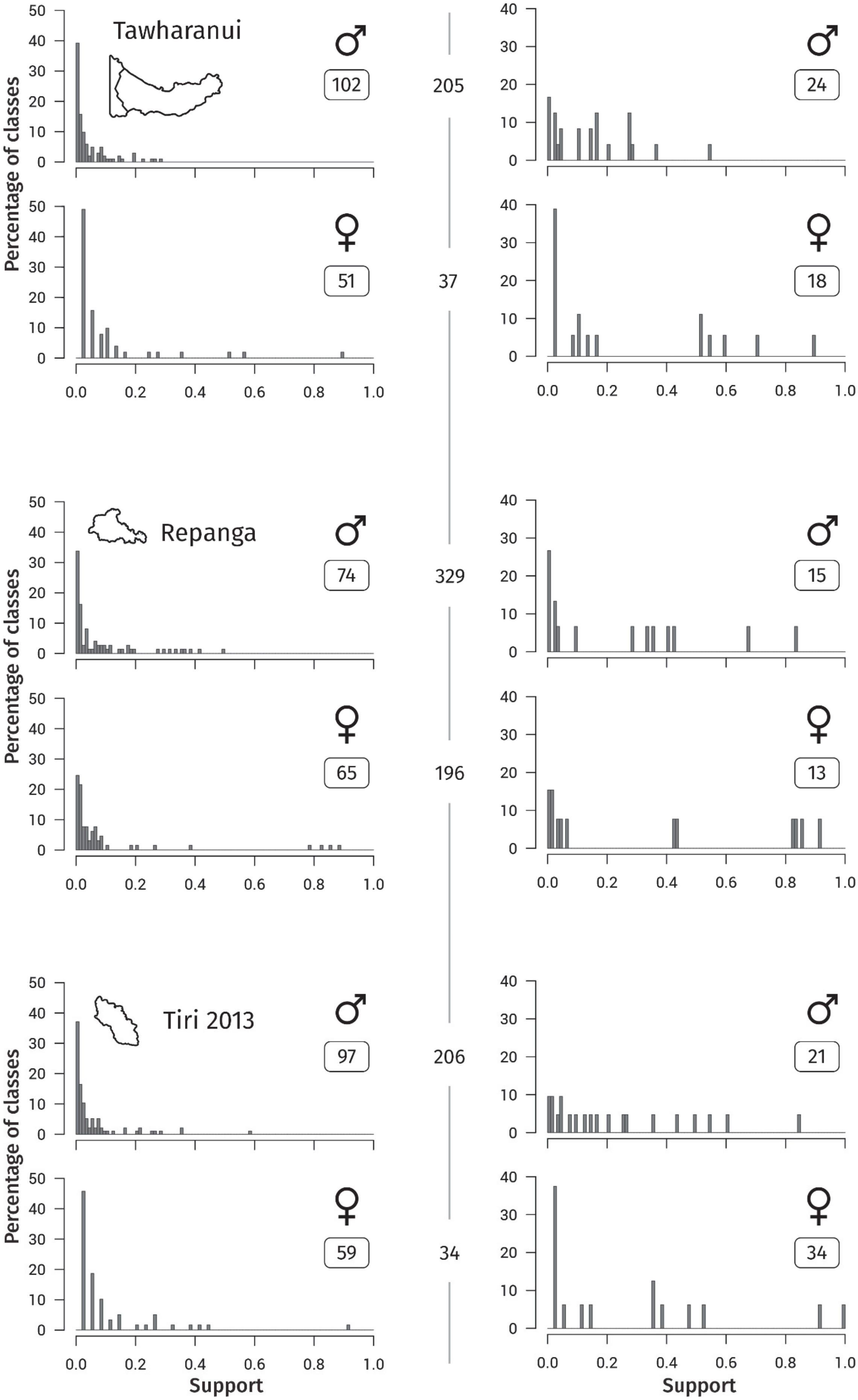
Tawharanui: *Stutter(TAW)* (0.89), *Dinnerplate(2)* (0.57), *Stutter_Downsqueak(yshape)* (0.51).

Repanga: *Down(notched)* (0.89), *Nailpolish* (0.85), *Stutter(decurved)* (0.83), *Down(CUV1)* (0.78).

Tiri: in all years, *Stutter* (>0.90); *Chiup* was an outlier in 2014 (0.78) and 2015 (0.61) after being only moderately prevalent in 2013 (0.38).

The distribution of syllable families (broad-scale classification) was generally less biased towards low prevalence, more discontinuously distributed and more variable between distributions.


Fine-scale classes
number of songs (n)
Broad-scale classes
Tawhiti Rahi
Representing Poor Knights
♂
119 classes
306
25 classes
♀
68
58
19
Lady Alice
Representing Hen & Chickens
♂
103
181
22
♀
80
126
15
Hauturu
♂
122
67
23
♀
139
98
18
Percentage of classes
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0
Support


Tawharanui
♂
102
205
♂
24
♀
51
37
♀
18
Repanga
♂
74
329
♂
15
♀
65
196
♀
13
Tiri 2013
♂
97
206
♂
21
♀
59
34
♀
34
Percentage of classes
0
10
20
30
40
50
0
10
20
30
40
0.0
0.2
0.4
0.6
0.8
1.0
Support

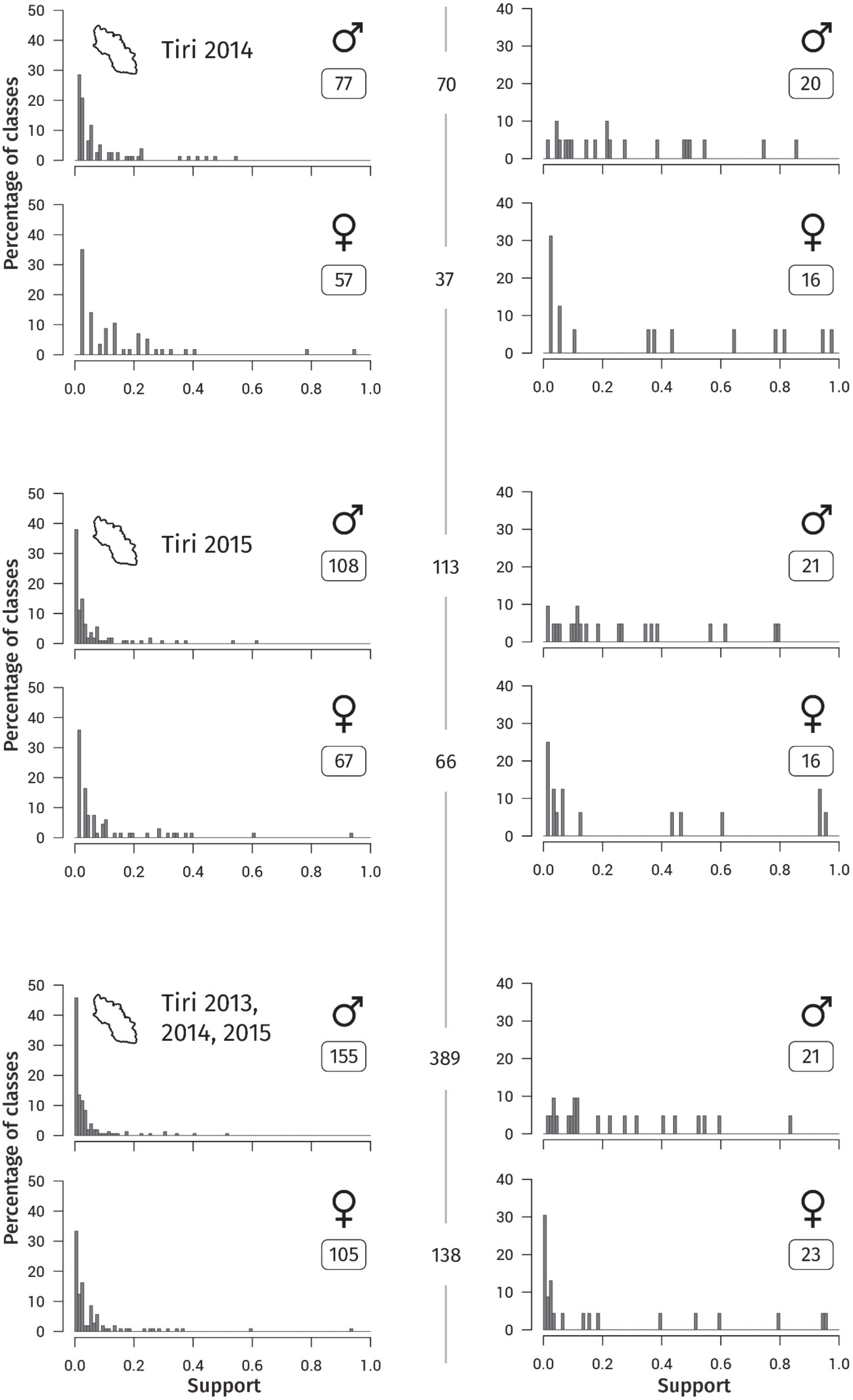

Tiri 2014
♂
77
70
♂
20
♀
57
37
♀
16
Tiri 2015
♂
108
113
♂
21
♀
67
66
♀
16
Tiri 2013, 2014, 2015
♂
155
389
♂
21
♀
105
138
♀
23
Percentage of classes
Support
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0

Figure 5.6—Distributions of syllable type prevalence for each site-and-sex population.

(***Above***) Prevalence is here measured as *support*, the proportion of songs in the population containing the syllable type. Distributions of fine-scale type prevalence are given in the left-hand column, and broad-scale family prevalence in the right-hand column. The number of types in each distribution are shown in rounded boxes. Sample sizes (number of songs) for each population are given between the columns. All types (i.e. with at least one occurrence) have been included.

5.4.2 Repertoire overlap of male and female population sectors

The degree of male–female repertoire overlap within sites is presented in Figure 5.7. At the fine scale, for all sites it is apparent that the male and female sectors had largely separate repertoires, ranging from 6% total units shared at Tawhiti Rahi to 26% shared at Hauturu. In other words, 74–94% of fine-scale syllable types at each site were sung by one sex only. The degree of male–female repertoire overlap stayed constant (6–8%) between years on Tiri, giving an indication of temporal stability.

At the broad scale, sharing was higher, ranging from 28% of syllable families shared at Tawharanui to 59% shared at Hauturu. In other words, 41–72% of families at each site were sung by one sex only. At this scale, the degree of repertoire overlap was slightly more variable between years on Tiri (39–48%).

A catalogue of syllable types and families sung by both sexes within each site is provided in Appendix 7.3.4.

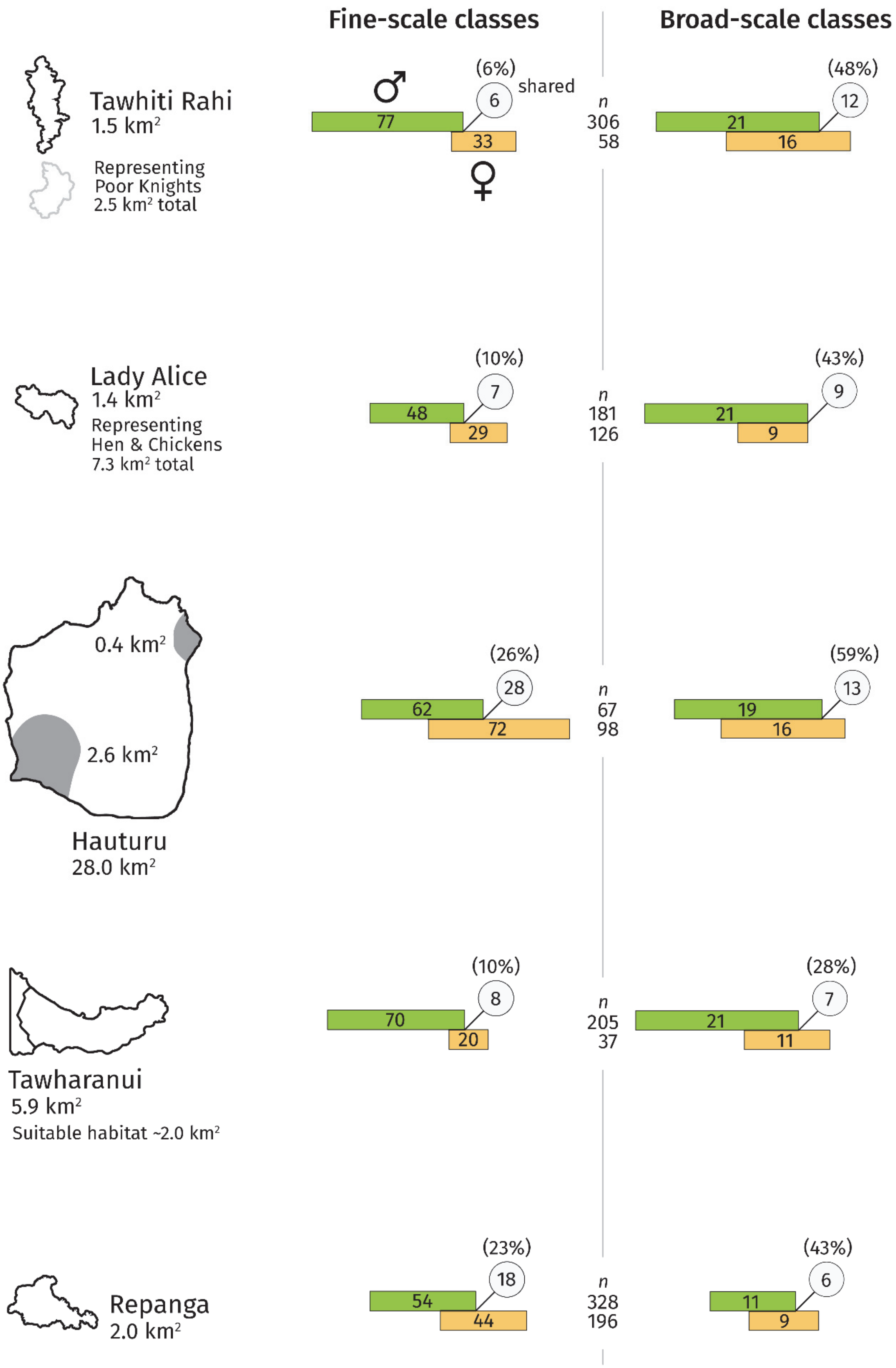

Fine-scale classes
Broad-scale classes
Tawhiti Rahi
1.5 km²
Representing
Poor Knights
2.5 km² total
♂
♀
(6%)
6
shared
77
33
n
306
58
(48%)
12
21
16
Lady Alice
1.4 km²
Representing
Hen & Chickens
7.3 km² total
(10%)
7
48
29
n
181
126
(43%)
9
21
9
0.4 km²
2.6 km²
Hauturu
28.0 km²
(26%)
28
62
72
n
67
98
(59%)
13
19
16
Tawharanui
5.9 km²
Suitable habitat ~2.0 km²
(10%)
8
70
20
n
205
37
(28%)
7
21
11
Repanga
2.0 km²
(23%)
18
54
44
n
328
196
(43%)
6
11
9

Figure 5.7—Overlap of male and female population syllable repertoires.

Each diagram indicates the recorded number of male (green bar) and female (orange bar) syllable types, and the number of types common to both sexes (circle). The percentage overlap is given above the circle, calculated as the Jaccard similarity index: number shared divided by total number of types; e.g. for Tawhiti Rahi at the fine scale, 6 shared / (77+33−6) in total ≈ 6%. For robustness, site-and-sex repertoires exclude types with fewer than three occurrences within that group. The left-column diagrams are based on fine-scale syllable classification, and right-column diagrams are based on broad-scale ('family'-level) classification. Bar length is proportional to the number of syllable types, and overlap length is proportional to the number of shared syllable types. Sample sizes (number of songs) for the male and female bars are given between the columns. Note that repertoire sizes are recorded values, not extrapolated, and therefore are not directly comparable due to differing sampling effort. However, relative overlap percentage is more robust to differing sampling effort and thus meaningful to compare.

Characterising shared versus sex-specific memes: The prevalence hypothesis

There was little evidence that memes sung by both sexes tend to be more (or less) prevalent than sex-specific memes. There are effectively nine groups in my sample when separating Tiri data by year (see above); considering both males and females separately, at both fine and broad scales of classification, there are thus 36 statistical groups for analysis (9×2×2=36). The majority of the site-and-sex groups (31/36) were not significantly different in prevalence between shared and sex-specific syllable types (two-tailed t-tests). Only five comparisons were significantly different: females on Lady Alice, fine-scale classification (p=0.008, t=2.789, df=40); females on Hauturu, fine-scale classification (p<0.001, t=3.871, df=91); males on Tawharanui, broad-scale classification (p=0.025, t=2.431, df=19); males on Tiri in 2013, broad-scale classification (p=0.032, t=2.326, df=18); females on Tiri in 2013, fine-scale classification (p=0.023, t=2.392, df=30). In these few cases of significant differences, the mean support of shared syllable types was greater than that of unshared types.

Characterising shared versus sex-specific memes: The acoustic properties hypothesis

Acoustic properties do indeed predict which syllables are common to both sexes. Linear discriminant analysis (LDA) reveals that of the 1,500 feature measurements extracted for each unit in *Koe* (see feature descriptions in *Koe* user manual: koe.io.ac.nz), a set of 17 features together provide the best separation between shared and sex-specific syllables; adding more features simply adds noise. The ordination of this 17-feature LDA (Figure 5.8) reveals that male-only and female-only syllables form two clear, partially-overlapping clusters, with shared syllables in between, when data from all sites are pooled. For individual sites the configuration of the three clusters varies but shows a tendency for female-only syllables to occur at lower LD1 values, male-only to occur at higher LD1 values, with shared syllables in between.

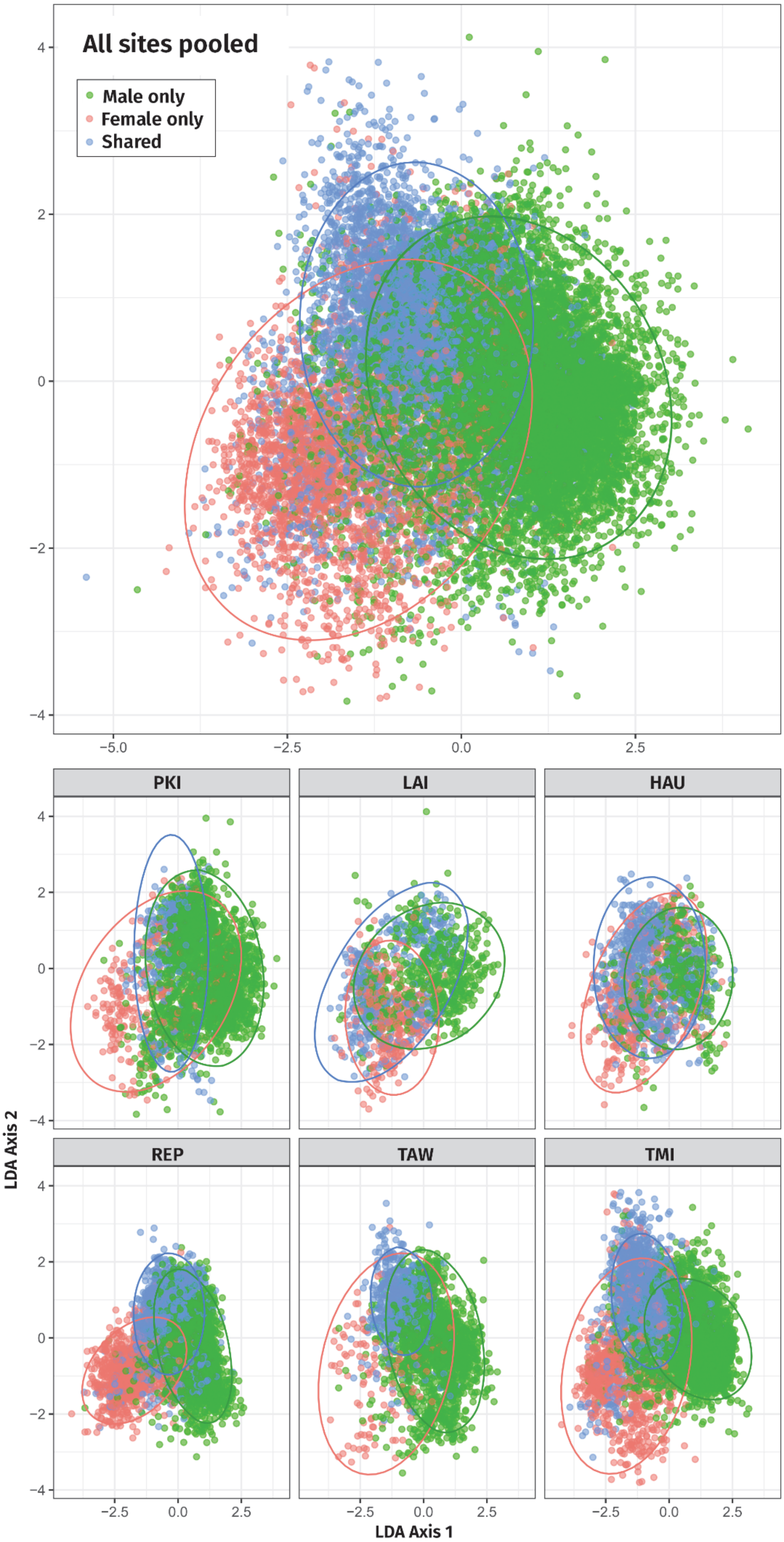

All sites pooled
Male only
Female only
Shared
4
2
0
−2
−4
−5.0
−2.5
0.0
2.5
PKI
LAI
HAU
REP
TAW
TMI
LDA Axis 2
LDA Axis 1

Figure 5.8—Two-dimensional linear discriminant analysis (LDA) of male-specific, female-specific, and shared syllables for all sites combined (top) and each site separately (bottom).

Ellipses represent the regions covering 95% of the data (according to the fitted normal). Each site is plotted using the same axes to be comparable. Site abbreviations are as follows: PKI, Tawhiti Rahi, representing Poor Knights Islands; LAI, Lady Alice Island, representing Hen & Chickens; HAU, Hauturu; REP, Repanga; TAW, Tawharanui; TMI, Tiri (2013, 2014, 2015).

The seven features most important for separating male-only, female-only and shared (common to both sexes) syllables are listed in Table 5.4, consisting of duration, a goodness-of-pitch-derived feature, and five Mel Frequency Cepstral Coefficient (MFCC)-derived features.

Table 5.4—Coefficients of linear discriminants.

Large weightings are in bold. The ‘duration’ feature is the duration of the syllables. Features containing ‘mfcc’ are derived from Mel Frequency Cepstral Coefficients (MFCC), commonly used in human speech recognition. Features containing ‘divcon’ result from a computer science technique called ‘divide and conquer’, which divides the transformed spectrogram into a specified number of segments and measures each separately. The bracketed numbers are an index specifying the transformed spectrogram segment. ‘Goodness of pitch’ is an estimate of harmonic pitch periodicity; high values indicate a harmonic stack, as opposed to a pure tone or white noise. See the *Koe* user manual for more detailed feature descriptions (github.com/fzyukio/koe/wiki).

Feature	LD1	LD2
duration	**0.50574**	**-0.62251**
mfcc_mean[4]	**-0.49722**	0.24487
mfcc_mean[6]	**0.53090**	0.11554
mfcc_divcon_3_mean[6]	**-0.64118**	**0.54774**
mfcc_divcon_5_mean[6]	**0.57327**	0.33713
mfcc_divcon_7_mean[63]	**-0.57809**	-0.17282
goodness_of_pitch_divcon_7_mean[4]	-0.02791	**0.55676**

5.4.3 The pattern of meme sharing between sites

Site–site repertoire overlap

Matrices of site–site repertoire overlap are presented in Figures 5.9–5.12. At the fine scale, the degree of site-site repertoire overlap between male populations was low, ranging 3–18% (median=6%); for females: 1–12% (median=5%). Remarkably, for both sexes, the fine-scale types shared between sites almost all belonged to the *Pipe* and *Stutter* families, with the exception of Hauturu–Tawharanui, which additionally shared types from the *Cough*, *Chump*, *Tink* and *Upsqueak* families for males and *Down*, *Pipe_Down* and *Downsqueak* families for females (see Appendix 7.3.5). This likely reflects the recent Hauturu→Tawharanui founding event. Both male and female repertoires overlapped between pairs of years within Tiri (2013 vs 2014, 2013 vs 2015, 2014 vs 2015) by just over half (49–61% and 50–60%, respectively).

At the broad scale, the degree of site–site repertoire overlap between male populations ranged 33–62% (median=48%); for females: 25–59% (median=32%). Because the lists of broad-scale families were filtered based on three or more occurrences *at the broad scale*, a broad scale family could consist of, for example, three fine-scale types with only one occurrence each. As a result, the diversity of families shared between sites was considerably greater at this scale, for both sexes (see Appendix 7.3.5). Between pairs of years within Tiri, overlap between male population repertoires ranged 95–100%, versus 79–100% for females.

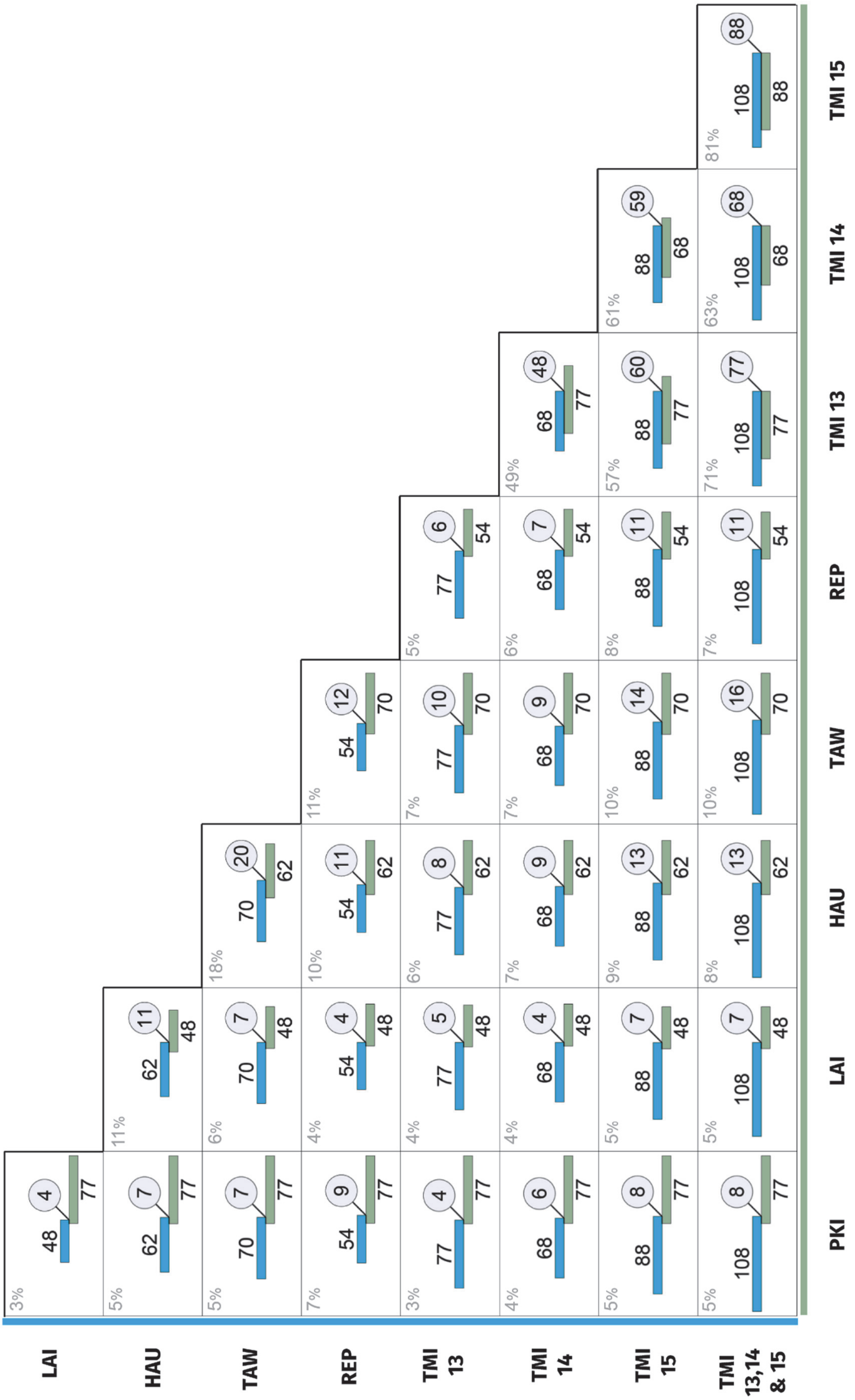

LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
TMI 13,14 & 15
PKI
LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
3%
48
77
4
5%
62
77
7
11%
62
48
11
5%
70
77
7
6%
70
48
7
18%
70
62
20
7%
54
77
9
4%
54
48
4
10%
54
62
11
11%
54
70
12
3%
77
77
4
4%
77
48
5
6%
77
62
8
7%
77
70
10
5%
77
54
6
4%
68
77
6
4%
68
48
4
7%
68
62
9
7%
68
70
9
6%
68
54
7
49%
68
77
48
5%
88
77
8
5%
88
48
7
9%
88
62
13
10%
88
70
14
8%
88
54
11
57%
88
77
60
61%
88
68
59
5%
108
77
8
5%
108
48
7
8%
108
62
13
10%
108
70
16
7%
108
54
11
71%
108
77
77
63%
108
68
68
81%
108
88
88

Figure 5.9—Matrix of male population repertoires (at fine-scale syllable classification) for each site.

Each cell of the matrix indicates the recorded number of male syllable types for the two sites (blue and green bars), and the number of types common to both (circle). The percentage overlap is given at the top left of each cell and is calculated as Jaccard similarity index: Number of types shared / (number of types at Site A + number of types at Site B – number of types shared). For robustness, site repertoires exclude types with fewer than three occurrences within that site and sex. Bar length is proportional to the number of syllable types, and overlap length is proportional to the shared number of syllable types. Site abbreviations are as follows: PKI, Tawhiti Rahi (representing Poor Knights Islands); LAI, Lady Alice Island (representing Hen and Chickens); HAU, Hauturu; TAW, Tawharanui; REP, Repanga; TMI, Tiri.

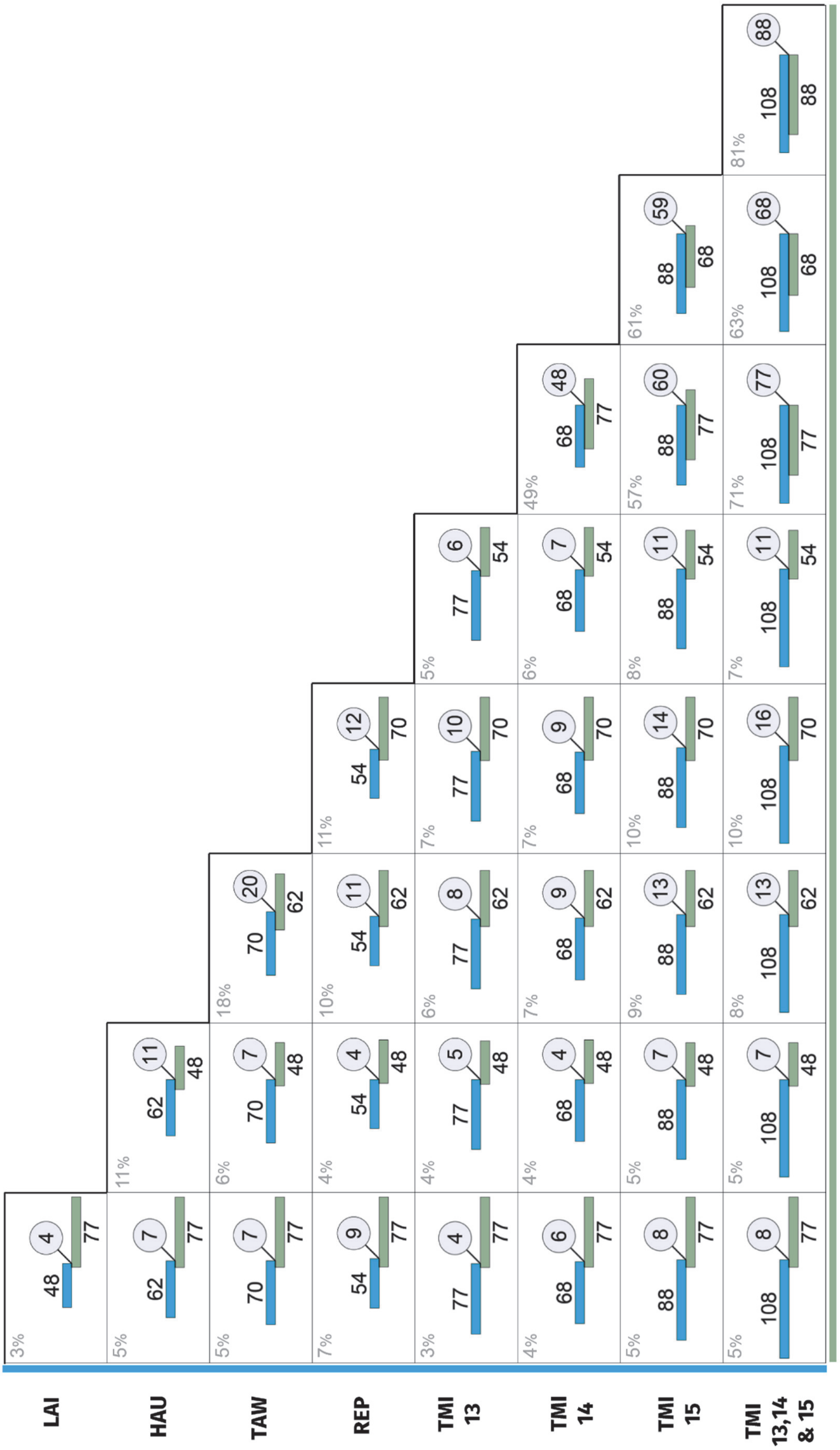

LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
TMI 13,14 & 15
PKI
LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
3%
48
77
4
5%
62
77
7
11%
62
48
11
5%
70
77
7
6%
70
48
7
18%
70
62
20
7%
54
77
9
4%
54
48
4
10%
54
62
11
11%
54
70
12
3%
77
77
4
4%
77
48
5
6%
77
62
8
7%
77
70
10
5%
77
54
6
4%
68
77
6
4%
68
48
4
7%
68
62
9
7%
68
70
9
6%
68
54
7
49%
68
77
48
5%
88
77
8
5%
88
48
7
9%
88
62
13
10%
88
70
14
8%
88
54
11
57%
88
77
60
61%
88
68
59
5%
108
77
8
5%
108
48
7
8%
108
62
13
10%
108
70
16
7%
108
54
11
71%
108
77
77
63%
108
68
68
81%
108
88
88

Figure 5.10—Matrix of female population repertoires (at fine-scale syllable classification) for each site.

Each cell of the matrix indicates the recorded number of female syllable types for the two sites (blue and green bars), and the number of types common to both (circle). The percentage overlap is given at the top left of each cell and is calculated as Jaccard similarity index: Number of types shared / (number of types at Site A + number of types at Site B – number of types shared). For robustness, site repertoires exclude types with fewer than three occurrences within that site and sex. Bar length is proportional to the number of syllable types, and overlap length is proportional to the shared number of syllable types. Site abbreviations are as follows: PKI, Tawhiti Rahi (representing Poor Knights Islands); LAI, Lady Alice Island (representing Hen and Chickens); HAU, Hauturu; TAW, Tawharanui; REP, Repanga; TMI, Tiri.

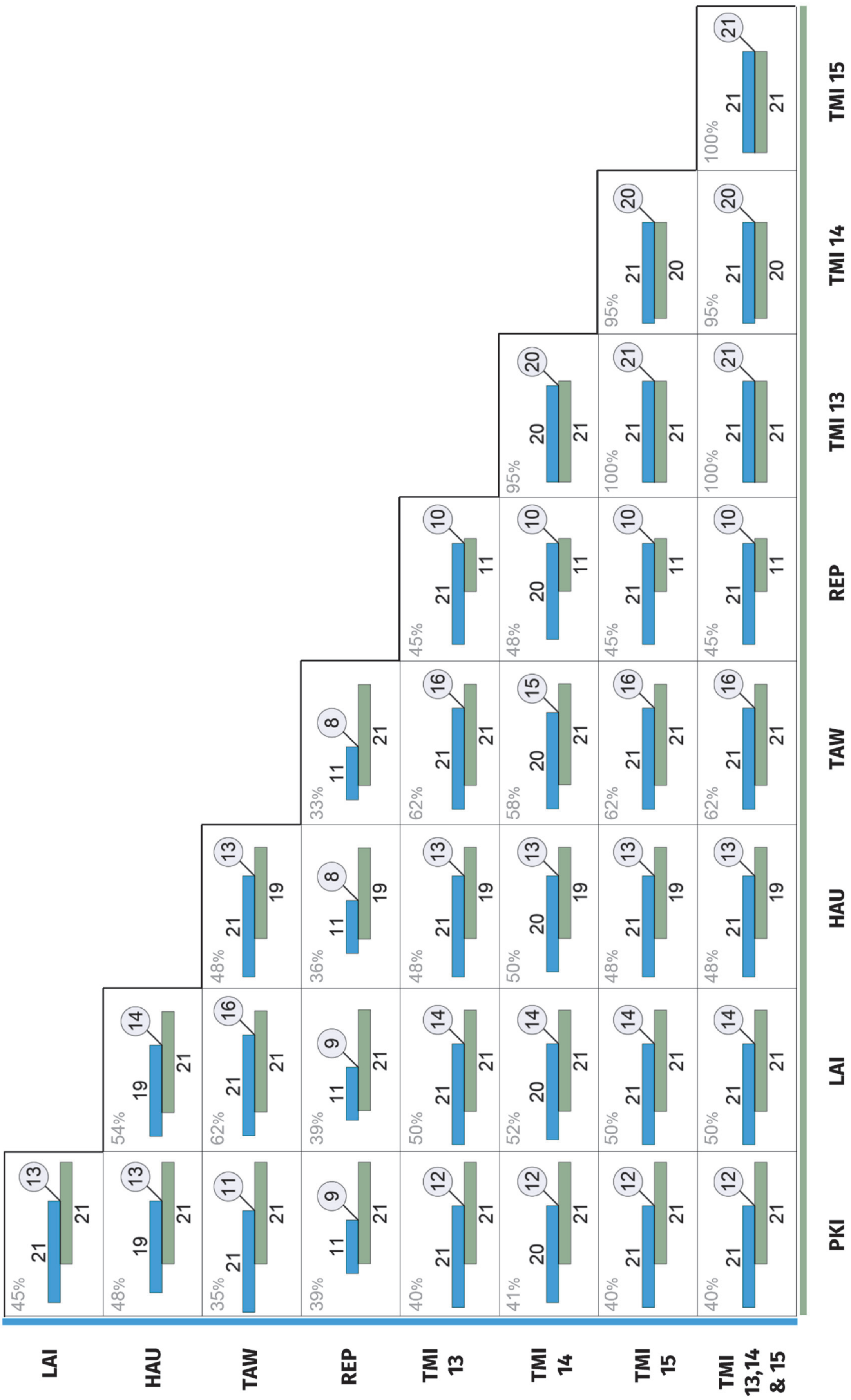

LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
TMI 13,14 & 15
PKI
LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
45% 21 21 13
48% 19 21 13
54% 19 21 14
35% 21 21 11
62% 21 21 16
48% 21 19 13
39% 11 21 9
39% 11 21 9
36% 11 19 8
33% 11 21 8
40% 21 21 12
50% 21 21 14
48% 21 19 13
62% 21 21 16
45% 21 11 10
41% 20 21 12
52% 20 21 14
50% 20 19 13
58% 20 21 15
48% 20 11 10
95% 20 21 20
40% 21 21 12
50% 21 21 14
48% 21 19 13
62% 21 21 16
45% 21 11 10
100% 21 21 21
95% 21 20 20
40% 21 21 12
50% 21 21 14
48% 21 19 13
62% 21 21 16
45% 21 11 10
100% 21 21 21
95% 21 20 20
100% 21 21 21

Figure 5.11—Matrix of male population repertoires (at broad-scale 'family'-level syllable classification) for each site.

Each cell of the matrix indicates the recorded number of male syllable types for the two sites (blue and green bars), and the number of types common to both (circle). The percentage overlap is given at the top left of each cell and is calculated as Jaccard similarity index: Number of types shared / (number of types at Site A + number of types at Site B – number of types shared). For robustness, site repertoires exclude families with fewer than three occurrences within that site and sex. Bar length is proportional to the number of syllable types, and overlap length is proportional to the shared number of syllable types. Site abbreviations are as follows: PKI, Tawhiti Rahi (representing Poor Knights Islands); LAI, Lady Alice Island (representing Hen and Chickens); HAU, Hauturu; TAW, Tawharanui; REP, Repanga; TMI, Tiri.

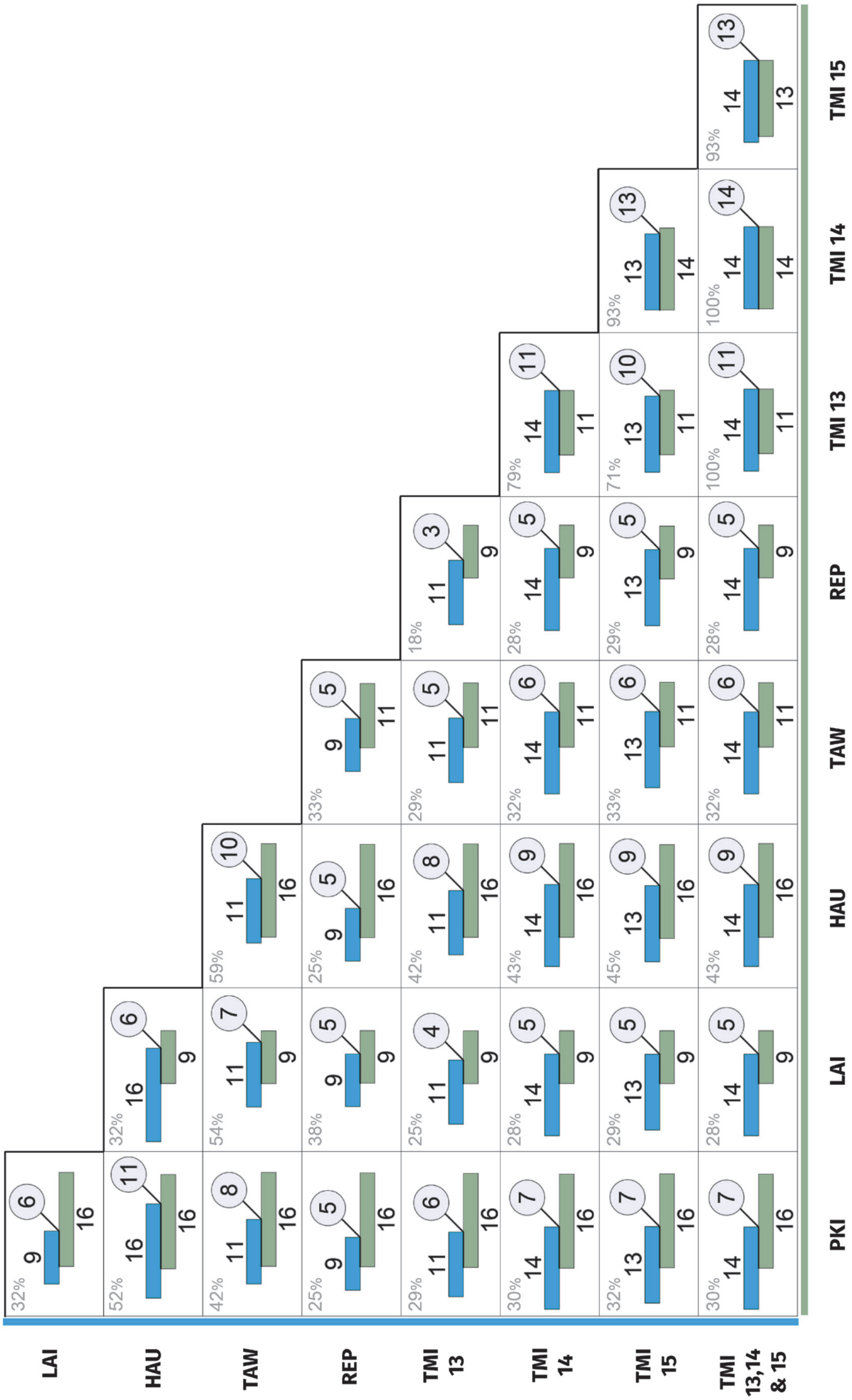

LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
TMI 13,14 & 15
PKI
LAI
HAU
TAW
REP
TMI 13
TMI 14
TMI 15
32% 9 16 6
52% 16 16 11
32% 16 9 6
42% 11 16 8
54% 11 9 7
59% 11 16 10
25% 9 16 5
38% 9 9 5
25% 9 16 5
33% 9 11 5
29% 11 16 6
25% 11 9 4
42% 11 16 8
29% 11 11 5
18% 11 9 3
30% 14 16 7
28% 14 9 5
43% 14 16 9
32% 14 11 6
28% 14 9 5
79% 14 11 11
32% 13 16 7
29% 13 9 5
45% 13 16 9
33% 13 11 6
29% 13 9 5
71% 13 11 10
93% 13 14 13
30% 14 16 7
28% 14 9 5
43% 14 16 9
32% 14 11 6
28% 14 9 5
100% 14 11 11
100% 14 14 14
93% 14 13 13

Figure 5.12—Matrix of female population repertoires (at broad-scale 'family'-level syllable classification) for each site.

Each cell of the matrix indicates the recorded number of female syllable types for the two sites (blue and green bars), and the number of types common to both (circle). The percentage overlap is given at the top left of each cell and is calculated as Jaccard similarity index: Number of types shared / (number of types at Site A + number of types at Site B – number of types shared). For robustness, site repertoires exclude families with fewer than three occurrences within that site and sex. Bar length is proportional to the number of syllable types, and overlap length is proportional to the shared number of syllable types. Site abbreviations are as follows: PKI, Tawhiti Rahi (representing Poor Knights Islands); LAI, Lady Alice Island (representing Hen and Chickens); HAU, Hauturu; TAW, Tawharanui; REP, Repanga; TMI, Tiri.

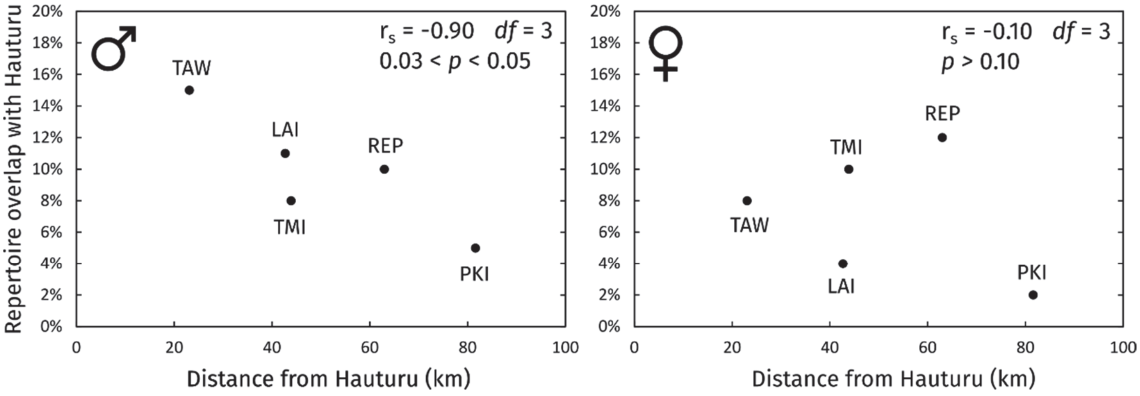
Repertoire overlap versus distance

Repertoire overlap with Hauturu versus geographic distance from Hauturu is shown in Figure 5.13. At fine-scale classification there was a strong negative correlation between sharing and distance for males ($r_S = -0.90$, $df=3$, $0.03<p<0.05$) but no relationship for females ($r_S = -0.10$, $df=3$, $p>0.10$). At broad-scale classification, there was no correlation for either sex (Males: $r_S = -0.45$, $df=3$, $p>0.10$; Females: $r_S = -0.30$, $df=3$, $p>0.10$). Contrary to predictions, sites at increasing distances from Hauturu did not share a progressively diminishing subset of syllable types with Hauturu, but different (apparently unrelated) subsets; this was true for both males and females (Appendix 7.3.5).

When considering sharing between *all* pairwise combinations of sites in the archipelago, males exhibited weak evidence of an isolation-by-distance pattern at fine-scale classification ($\rho=-0.525$, $p=0.064$), and strong evidence at the broad-scale ($\rho=-0.799$, $p=0.003$). For females, there was weak evidence of an isolation-by-distance pattern at fine-scale classification ($\rho=-0.525$, $p=0.076$) and no evidence at broad-scale classification ($\rho=-0.314$, $p=0.200$).

The recent Hauturu→Tawharanui founder event (Brunton et al., 2008a) could possibly inflate the degree of sharing between these two sites, compared to other sites where both populations are long-established. Therefore I tested robustness of the correlations to removal of the Hauturu–Tawharanui datapoint, and found little change in the slope or strength of the correlations, other than for females at the broad-scale classification level in the Hauturu-centric analysis (Figure 5.13); in this case, the slope inverted to a positive correlation.

Fine-scale classes



Broad-scale classes

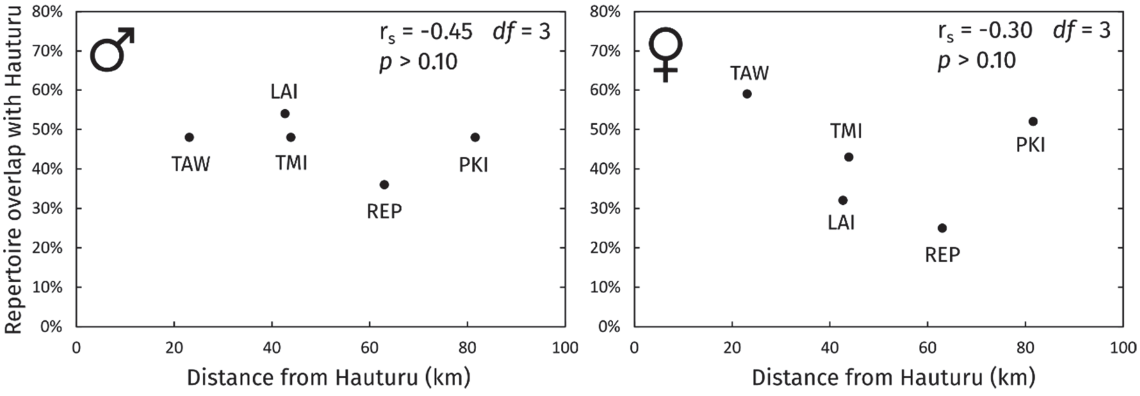


Figure 5.13—Repertoire overlap with Hauturu versus distance from Hauturu.

Overlaps were calculated on repertoire lists which had been filtered to only show types with 3+ occurrences within each site-and-sex group. Site abbreviations are as follows: TAW, Tawharanui;

LAI, Lady Alice Island (representing Hen and Chickens); TMI, Tiri (2013, 2014, 2015); REP, Repanga; PKI, Tawhiti Rahi (representing Poor Knights Islands).

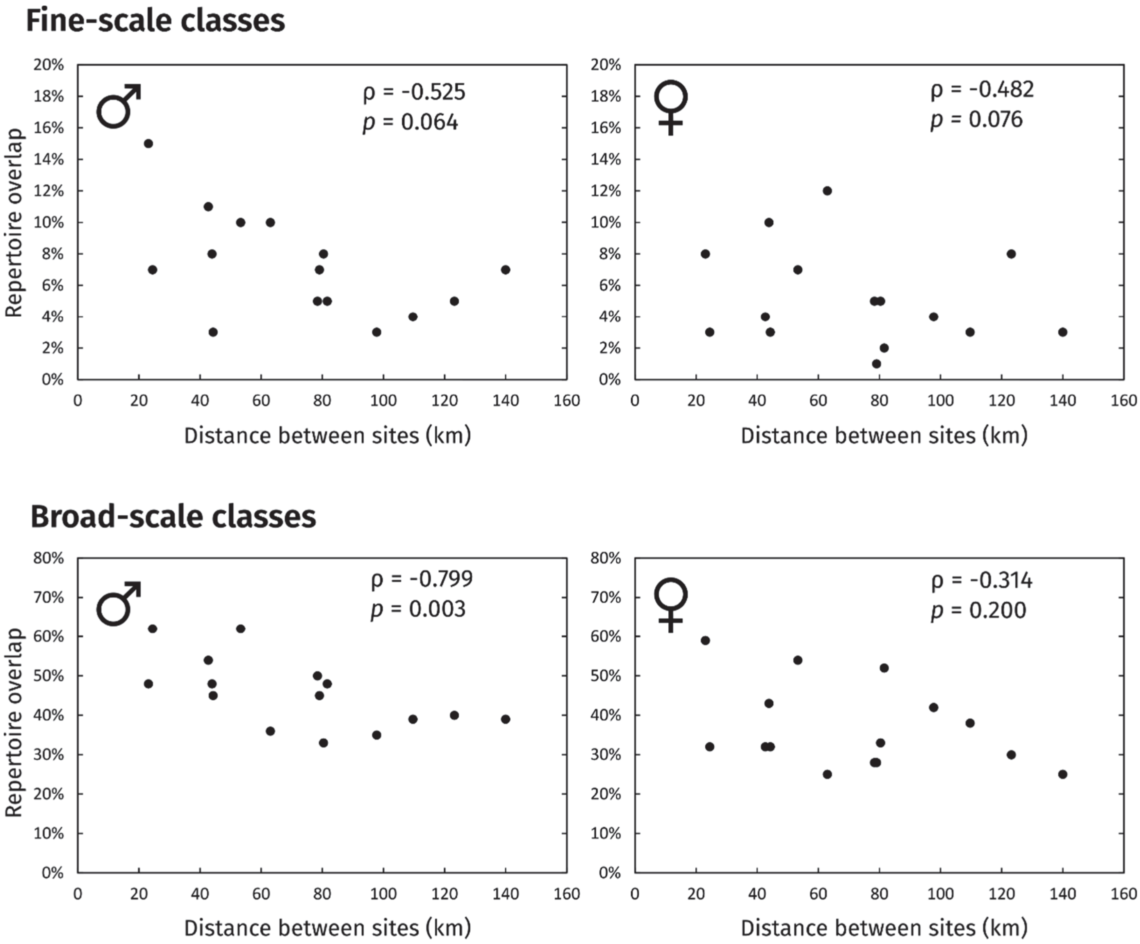


Figure 5.14—Site–site repertoire overlap versus site–site geographic distance, for all pairs of sites.

Overlaps were calculated on repertoire lists which had been filtered to only show types with 3+ occurrences within each site-and-sex group (Tiri considered as one site with all years pooled). Spearman rank correlations were calculated using the RELATE routine in *PRIMER* (version 7) with 9999 permutations, accounting for the non-independence of points.

Sex differences in degree of site-site repertoire overlap

At fine-scale classification, there was no evidence of differences between males and females in the degree of site-site repertoire overlap (Sign Test p=0.11). By contrast, at the broad scale, male populations had higher overlap than female populations in 31 out of 36 cases, providing strong evidence (Sign Test p<0.001) that males share more broad-scale syllable types between sites than do females—the opposite of what I predicted (Table 5.5).

Table 5.5—Sign Test results.

The test was conducted on fine-scale repertoire overlaps (Figure 5.9 and Figure 5.10) as well as broad-scale 'family' overlaps (Figure 5.11 and Figure 5.12).

Label granularity	M>F	M<F	M=F	Sign Test *p*
Fine-scale (n = 36)	22	14	0	0.106
Broad-scale (n = 36)	31	4	1	< 0.001

5.5 Discussion

My study provides some of the first evidence of female song dialects in songbirds. Comparing between the male and female sectors of each population, I found comparable levels of syllable diversity and similar-shaped distributions of syllable type prevalence, but largely distinct syllable repertoires (sharing only 6–26% of syllable types within each site). Furthermore, syllables shared between sexes can be statistically discriminated from sex-specific types based on acoustic properties. Most syllable types were site-specific; despite a large and varied repertoire within sites, generally the only types shared between sites were *Stutter* family syllables and pure-tone whistle notes of various pitches (*Pipe* family syllables). Between-site sharing of these syllable types decreased with distance for males but not females. These contrasting patterns of sharing across the archipelago may result from sex differences in dispersal, meme mutation rates and song-learning modes.

5.5.1 Repertoire sizes and meme prevalence

Repertoire sizes

In many male birds, repertoire size is driven by male–male competition and female choice (Catchpole, 1987; Hill et al., 2018), whereas the drivers of female repertoire size seem more likely to centre on female–female competition for non-sexual breeding resources (Tobias et al., 2012). In species where both sexes sing, comparing repertoire size can give an indication of the relative intensity of selection on the two sexes. Studies have found smaller female repertoires in some cases (e.g. rufous-and-white wrens *Thryophilus rufalbus*, Mennill and Vehrencamp, 2005; banded wrens *Thryophilus pleurostictus*, Hall et al., 2015), equal-sized repertoires in other cases (e.g. bay wren *Thryothorus nigricapillus*, Levin, 1996), and at least one case of larger female repertoire size (stripe-headed sparrow *Peucaea r. ruficauda*, Illes, 2014).

In the present study, I found no difference in male and female population repertoire sizes; in all sites but one (Tiri in 2015), male and female confidence intervals overlapped. Bellbirds are socially monogamous (Roper, 2018), and both sexes are highly social and aggressive in singing interactions. As they likely experience similar levels of social interaction and competition, it is plausible that male and female bellbirds are under similar levels of social pressure for song elaboration and that this results in similar levels of syllable diversity.

Repertoire sizes varied greatly between sites, and between years on Tiri. Adaptation to differing acoustic environments (Potvin and Clegg, 2015; Graham et al., 2017b) seems an unlikely explanation for this, as all sites were coastal, with similar vegetation structure (though soundscape was not measured). Perhaps between-site differences in repertoire size are driven by different levels of competition. For example, sites with higher population density may have elevated competition for food, or sites with high population connectivity may experience increased competitive encounters with migrants—selecting for bigger repertoires. However, this idea cannot be assessed with population-level data alone, because we need to distinguish the contributions of within-individual and between-individual diversity. In other words, larger repertoire sizes (such as on Hauturu) could reflect larger *individual* repertoires, or greater diversity *between* individuals. To resolve these two sources of diversity will require focused recording of known individuals.

Comparing syllable repertoires between years on Tiri indicates a higher rate of turnover for females than for males. If this is true, several possibilities suggest themselves. First, perhaps there is a higher turnover of female individuals through higher immigration/emigration (in line with female-biased dispersal; Baillie, 2011). A second possibility involves repertoire cycling. Analogous to how different genes are turned on or off in different cellular environments, individual bellbirds seem to selectively express different sets of syllables in different years, drawing from their wider latent repertoire (Roper, 2018). Thus, a higher syllable turnover may be due to individual females cycling through their latent repertoires at a faster rate than males. Third, perhaps female memes mutate more quickly than male memes. For example, in some species male song consistency is enforced by sexual selection, whereas female song can diversify free from this constraint (e.g. eastern whipbirds, Mennill and Rogers, 2006). More data on dispersal rates and longitudinal song data from known individuals will illuminate the processes underlying repertoire turnover.

Regarding repertoire size, one more point of interest is the relative influence of culture versus genetics in determining song expression—a topic that has inspired considerable debate (Mundinger, 1980; Cavalli-Sforza and Feldman, 1981; Boyd and Richerson, 1988; Aunger, 2001; Richerson et al., 2010; Whiten et al., 2011; Graham et al., 2018a). If repertoire size is linked to genetic diversity in bellbirds, we would expect the sites with lower genetic diversity (Tawhiti Rahi and Tiri; Figure 5.3) to have smaller repertoires. This was not the case (Table 5.3). Therefore, song culture does not appear tightly constrained by genetic diversity in this species (see also Graham et al., 2018).

Prevalence of syllable types

Overall, male and female prevalence distributions were similar in shape. At fine-scale classification, both sexes had many rare types—predominantly from *Pipe* and closely related *Stepup* and *Stepdown* families—and few prevalent types. At some sites, females had outlier types of especially high prevalence. However, at broad-scale ('family'-level) classification, the distribution of syllable families was generally more variable, more discontinuous, and less biased towards low prevalence.

The high percentage of rare types, most of them from the *Pipe* family, likely results from the very fine granularity of *Pipe* syllable classification. Unlike other types, which do not show high degrees of pitch variability, *Pipe* syllables as a group show continuous pitch variation; I thus employed a different classification method for *Pipe* syllables (described in Chapter 3), classifying based on just-noticeable differences in pitch with reference to a sinewave tone generator. Thus, in these prevalence distributions the bias towards rare types could partly reflect the high pitch variability of the *Pipe* family.

The wide variation in type prevalence is intriguing. Bellbird chorusing behaviour involves aggressive syllable matching using prevalent syllable types. But bellbirds also sing solo songs, where the risks of immediate aggressive contact may be lower, and here it is possible that rare types are favoured. For example, in Darwin's medium ground finches *Geospiza fortis*, males who sang less common song types were more likely to breed and had better survival (Gibbs, 1990). I speculate that different pressures during chorusing and solo song could explain the maintenance of a range of prevalent and rare types.

5.5.2 Repertoire overlap of male and female population sectors

An exciting finding is that male and female bellbirds had largely distinct, yet partly overlapping song cultures. Within sites, males and females shared 6–26% of fine-scale syllable types. Overlap of syllable-family-level repertoires was higher but still partial (except for Lady Alice, which had 100% overlap). The results are intriguing because logically, all sex-specific syllables types must be learnt intra-sexually. That is, females must learn female-specific types from females, males must learn male-specific types from males (as in rufous-and-white wrens, for example; Mennill and Vehrencamp, 2005). But what about the types that are common to both sexes? These must reflect inter-sexual learning in some form, whether accidental or intentional, ongoing or historic, whether male→female or female→male or bi-directional (Evans and Kleindorfer, 2016). Once learnt inter-sexually by one or few individuals, it is conceivable that these memes proceed to be transmitted intra-sexually and become an established part of the repertoire for that sex—though experimental studies are needed to determine the specifics of song learning modes in bellbirds (see Chapter 6).

Why have partially-overlapping song cultures? Roper (2018) found that juvenile female bellbirds on Tiri attempt to sing male-specific syllable types—such as those in the *Cough* family—but drop these from their repertoires before crystallization. Perhaps both sexes of bellbird are physiologically capable of overlapping in acoustic space, but other factors (e.g. learning strategies, sexual/social selection) prevent it in the wild. This appears to be the case in slate-coloured boubous *Laniarius funebris*; male and female boubous do not share any syllable types in the wild, but birds hand-raised under experimental conditions developed syllables of both sexes (Wickler and Seibt, 1988). Wild birds may choose to express sex-specific memes to help avoid being mistaken for the other sex, which risks attracting same-sex rivals or repelling potential mates (Logue et al., 2007). However, the benefit of *shared* syllables is unclear. Playback studies will be needed to investigate whether sex-specific syllables are used for intra-sexual communication, and shared syllables for inter-sexual communication, for example.

I tested two potential predictors of male–female sharing. In my 'prevalence' hypothesis, I reasoned that females may be more likely to learn prevalent syllable types from males (or males from females) because of frequent exposure (Mennill et al., 2018), or conversely would prefer to learn rare types—depending on learning mode (Laland, 2004). Overall, I found no evidence for prevalence (as measured by support) as a driver of intersexual syllable sharing. However, this result is confounded. If the prevalence of a type *at the time of learning* is a key determinant of intersexual transmission, the strong fluctuation of syllable prevalence over time (Tiri data in Appendix 7.3.3)—and the potential for individuals to modify their repertoires dynamically—could easily obscure the connection between intersexual sharing and prevalence at the population level. Therefore, we cannot discount the possibility of a link between prevalence and intersexual transmission until tutoring experiments have elucidated bellbird song learning modes in individuals (see Chapter 6).

If the 'acoustic properties' hypothesis is true, the acoustic properties of a syllable type will determine its likelihood of being sung by both sexes. That is, we should be able to distinguish sex-specific types from types common to both sexes based on some combination of acoustic traits (Figure 5.4). My results support this hypothesis, revealing that male and female syllables occupy largely distinct regions of acoustic feature-space in an LDA ordination, with shared syllables occupying a cluster between in what might be called a 'sex-neutral' range (Figure 5.8). Shared syllables can indeed be discriminated from sex-specific based on acoustic properties—though the 17 features that best separate the clusters are not straightforward to interpret (Table 5.4). In the ordination, the similarities and differences between sites in the LDA ordination may provide an example of interaction between morphology and culture (Whiten et al., 2011). That is, the consistent separation of male-specific and female-specific memes in acoustic space could be due to morphological constraints of body size and syrinx structure, as males are 20% larger than females (Heather and Robertson, 2000) and have different syrinx morphologies (Roper, 2018). On the other hand, the spread and shape of the clusters vary widely between sites, and this geographic variation likely results from site-specific cultures—the outcome of cultural drift, for example (Slater, 1989; Graham et al., 2018a). It will be revealing to see whether other species share the bellbird pattern of repertoire overlap between sexes, and I propose that the honeyeaters (Family Meliphagidae)—of which the bellbird is a member—provide a good study group to begin investigating this question.

5.5.3 The pattern of meme sharing between sites

I reasoned that the pattern of sharing between sites could indicate the stringency of open- or closed-ended learning in bellbirds. (1) If bellbirds are completely closed-ended, there should be a strong pattern of acoustic isolation by distance between sites; because dispersing adults would be unable to change their repertoires after dispersal, the degree of site–site repertoire overlap would directly reflect population connectivity. (2) At the other extreme, I reasoned that if immigrant bellbirds completely abandoned their source population dialect upon arrival, learning local memes instead, then there should be little-to-no sharing between sites and no correlation with distance, because local cultures would be free to diverge independently. (3) Intermediate to these extremes is the possibility that immigrants continue to sing *some* of their source population memes but *augment* their repertoire with local population memes upon arrival. This would predict an isolation-by-distance pattern, but weaker than in a closed-ended learning scenario.

I found that both sexes shared a small percentage of syllable types between sites. However, the only two statistically significant cases of isolation by distance were for males: (i) fine-scale repertoire overlap with Hauturu versus distance from Hauturu (r_S=0.90, df=3, $0.03<p<0.05$); and (ii) broad-scale repertoire overlap between sites versus distance between sites (ρ=-0.799, p=0.003).

Crucially, types shared between sites were almost totally limited to *Pipe* and *Stutter* families, for both sexes (except for Hauturu–Tawharanui, which shared more families as a result of the recent founder event discussed earlier). This implies that immigrants largely abandon source memes upon arrival, except for types from *Pipe* and *Stutter* families. That is, there is indication here that most fine-scale meme types are 'local only', perhaps incurring high aggression at new sites, with *Pipe* and *Stutter* families constituting 'widely acceptable' memes that can be sung freely at multiple sites (see the 'colony password' hypothesis of Feekes, 1977). Therefore, this pattern best supports (3)—an open-ended learning strategy with limited retention of source memes upon immigration.

Given the simplicity of *Pipe* syllables and their continuous pitch variation, could a particular *Pipe* type arise at multiple sites by coincidence, rather than through transmission? If *Pipe* type sharing between sites was spurious, we should not expect the isolation-by-distance pattern observed for males, because there would be no link between number of *Pipe* types shared and geographic distance. Therefore, my results suggest that *Pipe* types (i.e. specific pitches of *Pipe*) are indeed transmitted as categorical memes, at least for males.

Given the large size and geographic centrality of the Hauturu population—which make it a likely source of dispersing bellbirds to other sites in the archipelago—I expected the islands around Hauturu to form a chain of 'stepping stones' for bellbird dispersal, thus leading to progressively more dissimilar repertoires away from the central source population. There are many examples of such chains; for example, in North Island saddleback *Philesturnus carunculatus rufusater*, serial translocations have resulted in a progressive decrease in song sharing when moving from the ancestral population through to serially translocated populations in the Hauraki Gulf (Parker et al., 2012). In chaffinches *Fringilla ceolebs*, sequential colonisation of the Atlantic Islands appears to have led to progressive loss of syntactical structure along the island chain (Lachlan et al., 2013). And greenish warblers *Phylloscopus trochiloides*—a so-called 'ring species'—show intergrading variation in song characteristics in a ring encircling the Tibetan Plateau (Irwin et al., 2001, 2005). In contrast to these examples, I found that for both male and female bellbirds, sites at increasing distances from Hauturu did not share a progressively diminishing subset of syllable types with Hauturu, but different (apparently unrelated) subsets. Therefore, my results suggest direct dispersal to each site from Hauturu, rather than serial dispersal along the island chain.

Lastly, I predicted that female-biased dispersal in bellbirds should result in higher flow of female memes between sites than male memes, and thus higher inter-site sharing for females. Surprisingly, I found the opposite: despite female-biased dispersal, female bellbirds shared *fewer* syllable families between sites than did males. A similar pattern was observed in eastern whipbirds, where female song had much greater differentiation with distance than male song (Mennill and Rogers, 2006). The authors surmise that increased variation in females could result from different selective pressures for the two sexes; male song consistency may be enforced by sexual selection, whereas female song (they suggest) may be free from such a constraint and thus rapidly evolve.

In bellbirds, a faster syllable mutation rate for females is believable, based on the higher turnover of syllable types across consecutive years on Tiri. I suggest that an increased mutation rate for female syllables could plausibly account for greater divergence between female groups in bellbirds, driven by differences in male and female sexual/social selection (Tobias et al., 2012) and song learning modes (Ljubičić et al., 2016). Examining the mechanisms of dispersal, selection and learning modes will require sophisticated field experiments (see Chapter 6). It will be interesting to see whether the observed pattern of male and female sharing between sites exists for other species, as more studies on female song are conducted.

5.6 Conclusion

To my knowledge, this is one of the first studies on female song culture in a multi-site metapopulation (Mennill and Rogers, 2006; Graham et al., 2018a, 2018b). Such an undertaking is logistically challenging; in contrast to a laboratory situation, it is difficult to amass large amounts of data, the identity of individuals is not known, and there is little control over the social context of singing—which limits assessment of individual-level mechanisms. However, there are striking advantages to my study system. By exposing the large-scale patterns of cultural diversification for male and female bellbirds, I have found some of the first evidence for female song dialects in songbirds, and revealed contrasting cultural processes for males and females. My key findings were as follows.

First, with equal syllable diversity and largely separate repertoires, female bellbird song is not a simplified subset of male song, nor can it be construed as an optional accessory to male song. Rather, it appears as much an elaborate signal as male song. This underlines the inadequacy of male-centric sexual selection theory, and calls for comparisons of male and female repertoires in many more species (Riebel et al., 2019)—such a scenario may turn out to be common amongst songbirds.

Second, I found good visual evidence that acoustic properties determine which syllable types are transmitted between sexes. I suggested playback experiments to test the idea that sex-specific types function in intra-sexual communication, and shared types for inter-sexual communication.

Third, I found intriguing similarities and differences in the patterns of sharing between sites for males and females. Remarkably, although both sexes had complex and varied repertoires, sharing between sites was limited to only *Stutter* family syllables and pure-tone whistle notes of various pitches (*Pipe* family syllables)—a far starker population divergence than I expected, given high dispersal between sites (Baillie et al., 2014). To explain this phenomenon, I reasoned that immigrant bellbirds may abandon the memes of their source population and instead adopt the local memes of the new population—retaining only *Pipe* and *Stutter* syllables. Between-site sharing decreased with distance for males, but not females; furthermore, despite females moving between sites more frequently, males had greater site–site repertoire overlap. These contrasting cultural patterns suggest there are sex differences in dispersal, meme mutation rates and/or song-learning modes. Detailed field-based experiments will help to resolve the individual-level mechanisms giving rise to the population-level patterns I have uncovered.

5.7 Author contributions

Dianne Brunton and Wesley Webb conceived of the study. Wesley Webb, Michelle Roper, Dianne Brunton, Barbara Evans, and Aaron Harmer conducted the fieldwork. Wesley Webb, Yukio Fukuzawa, and Michelle Roper created the song database. Wesley Webb, Yukio Fukuzawa and Matthew Pawley analysed the data, with guidance from Dianne Brunton and Aaron Harmer. Wesley Webb wrote the chapter with guidance from Dianne Brunton and Aaron Harmer.

5.8 References

Anderson, S. H., and Craig, J. L. (2003). Breeding biology of bellbirds (*Anthornis melanura*) on Tiritiri Matangi Island. *Notornis* 50, 75–82.

Aplin, L. M. (2019). Culture and cultural evolution in birds: a review of the evidence. *Anim. Behav.* 147, 179–187. doi:10.1016/j.anbehav.2018.05.001.

Arcese, P., Stoddard, P. K., and Hiebert, S. M. (1988). The form and function of song in female song sparrows. *The Condor* 90, 44–50. doi:10.2307/1368431.

Aunger, R. ed. (2001). *Darwinizing culture: the status of memetics as a science*. 1 edition. Oxford ; New York: Oxford University Press.

Austin, V. I., Higgott, C., Viguier, A., Grundy, L., Russell, A. F., and Griffith, S. C. (2019). Song rate and duetting in the chirruping wedgebill (*Psophodes cristatus*): frequency, form and functions. *Emu - Austral Ornithol.* 0, 1–9. doi:10.1080/01584197.2018.1561193.

Baillie, S. M. (2011). Population genetics, biogeography and ecological interactions of the New Zealand bellbird (*Anthornis melanura*) and their avian malaria parasites: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph. D.) in Zoology at Massey University, Albany, New Zealand. Available at: http://muir.massey.ac.nz/handle/10179/2863 [Accessed July 18, 2014].

Baillie, S. M., Ritchie, P. A., and Brunton, D. H. (2014). Population genetic connectivity of an endemic New Zealand passerine after large-scale local extirpations: a model of re-colonization potential. *Ibis* 156, 826–839. doi:10.1111/ibi.12182.

Baker, M. C., and Gammon, D. E. (2008). Vocal memes in natural populations of chickadees: why do some memes persist and others go extinct? *Anim. Behav.* 75, 279–289. doi:10.1016/j.anbehav.2007.05.022.

Baptista, L. F., and Schuchmann, K.-L. (1990). Song learning in the Anna hummingbird (*Calypte anna*). *Ethology* 84, 15–26.

Bartle, J. A., and Sagar, P. M. (1987). Intraspecific variation in the New Zealand bellbird *Anthornis melanura*. *Notornis* 34, 253–306.

Beecher, M. D., and Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* 20, 143–149. doi:10.1016/j.tree.2005.01.004.

Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *Q. Rev. Biol.* 74, 21–45. doi:10.1086/392950.

Boyd, R., and Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago Press.

Bradley, D. W., and Mennill, D. J. (2009). Solos, duets and choruses: vocal behaviour of the rufous-naped wren (*Campylorhynchus rufinucha*), a cooperatively breeding neotropical songbird. *J. Ornithol.* 150, 743–753.

Brakes, P., Dall, S. R. X., Aplin, L. M., Bearhop, S., Carroll, E. L., Ciucci, P., et al. (2019). Animal cultures matter for conservation. *Science* 363, 1032–1034. doi:10.1126/science.aaw3557.

Brown, R. N., and Lemon, R. E. (1979). Structure and evolution of song form in the wrens *Thryothorus sinaloa* and *T. felix*. *Behav. Ecol. Sociobiol.* 5, 111–131. doi:10.1007/BF00293301.

Brunton, D. H., Evans, B. A., and Ji, W. (2008a). Assessing natural dispersal of New Zealand bellbirds using song type and song playbacks. *N. Z. J. Ecol.* 32, 147–154.

Brunton, D. H., Evans, B., Cope, T., and Ji, W. (2008b). A test of the dear enemy hypothesis in female New Zealand bellbirds (*Anthornis melanura*): female neighbors as threats. *Behav. Ecol.* 19, 791–798. doi:10.1093/beheco/arn027.

Brunton, D. H., and Li, X. (2006). The song structure and seasonal patterns of vocal behavior of male and female bellbirds (*Anthornis melanura*). *J. Ethol.* 24, 17–25. doi:10.1007/s10164-005-0155-5.

Buddle, M. G. A. (1941). Birds of the Poor Knights. *Emu - Austral Ornithol.* 41, 56–68. doi:10.1071/MU941056.

Catchpole, C. K. (1987). Bird song, sexual selection and female choice. *Trends Ecol. Evol.* 2, 94–97. doi:10.1016/0169-5347(87)90165-0.

Catchpole, C. K., and Slater, P. J. B. (2008). *Bird song: Biological themes and variations*. 2nd ed. New York, NY, US: Cambridge University Press doi:10.1017/CBO9780511754791.

Cavalli-Sforza, L. L., and Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press.

Claidière, N., and André, J.-B. (2011). The transmission of genes and culture: a questionable analogy. *Evol. Biol.* 39, 12–24. doi:10.1007/s11692-011-9141-8.

Clarke, A. L., Saether, B.-E., and Røskaft, E. (1997). Sex biases in avian dispersal: a reappraisal. *Oikos*, 429–438. doi:10.2307/3546885.

Clarke, K. R., and Gorley, R. N. (2015). *PRIMER v7: User Manual/Tutorial*. PRIMER-EPlymouth.

Colwell, R. (2013). EstimateS*: Statistical estimation of species richness and shared species from samples*. Available at: http://purl.oclc.org/estimates.

Craig, J. L., and Douglas, M. E. (1984). Temporal partitioning of a nectar resource in relation to competitive asymmetries. *Anim. Behav.* 32, 624–625. doi:10.1016/S0003-3472(84)80300-0.

Craig, J. L., and Douglas, M. E. (1986). Resource distribution, aggressive asymmetries and variable access to resources in the nectar feeding bellbird. *Behav. Ecol. Sociobiol.* 18, 231–240. doi:10.1007/BF00290827.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Dixon, W. J., and Mood, A. M. (1946). The Statistical Sign Test. *J. Am. Stat. Assoc.* 41, 557–566. doi:10.1080/01621459.1946.10501898.

Driskell, A., Christidis, L., Gill, B. J., Boles, W. E., Barker, F. K., and Longmore, N. W. (2007). A new endemic family of New Zealand passerine birds: adding heat to a biodiversity hotspot. *Aust. J. Zool.* 55, 73–78.

Evans, C., and Kleindorfer, S. (2016). Superb fairy-wren (*Malurus cyaneus*) sons and daughters acquire song elements of mothers and social fathers. *Front. Ecol. Evol.* 4, 9.

Feekes, F. (1977). Colony-specific song in *Cacicus cela* (Icteridae, Aves) - password hypothesis. *Ardea* 65, 197–202.

Gibbs, H. L. (1990). Cultural evolution of male song types in Darwin’s medium ground finches, Geospiza fortis. *Anim. Behav.* 39, 253–263. doi:10.1016/S0003-3472(05)80869-3.

Graham, B. A., Heath, D. D., and Mennill, D. J. (2017a). Dispersal influences genetic and acoustic spatial structure for both males and females in a tropical songbird. *Ecol. Evol.* 7, 10089–10102. doi:10.1002/ece3.3456.

Graham, B. A., Heath, D. D., Walter, R. P., Mark, M. M., and Mennill, D. J. (2018a). Parallel evolutionary forces influence the evolution of male and female songs in a tropical songbird. *J. Evol. Biol.* 31, 979–994. doi:10.1111/jeb.13279.

Graham, B. A., Heath, D. D., Walter, R. P., and Mennill, D. J. (2018b). Immigrant song: males and females learn songs after dispersal in a tropical bird. *Behav. Ecol.* 29, 711–723. doi:10.1093/beheco/ary029.

Graham, B. A., Sandoval, L., Dabelsteen, T., and Mennill, D. J. (2017b). A test of the Acoustic Adaptation Hypothesis in three types of tropical forest: degradation of male and female Rufous-and-white Wren songs. *Bioacoustics* 26, 37–61. doi:10.1080/09524622.2016.1181574.

Greenwood, P. J. (1980). Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.* 28, 1140–1162. doi:10.1016/S0003-3472(80)80103-5.

Hall, M. L. (2000). The function of duetting in magpie-larks: conflict, cooperation, or commitment? *Anim. Behav.* 60, 667–677. doi:10.1006/anbe.2000.1517.

Hall, M. L., Rittenbach, M. R. D., and Vehrencamp, S. L. (2015). Female song and vocal interactions with males in a neotropical wren. *Behav. Evol. Ecol.* 3, 12. doi:10.3389/fevo.2015.00012.

Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., et al. (1989). Similarity measures in scientometric research: The Jaccard index versus Salton's cosine formula. *Inf. Process. Manag.* 25, 315–318. doi:10.1016/0306-4573(89)90048-4.

Heather, B. D., and Robertson, H. A. (2000). *The field guide to the birds of New Zealand.* Viking\Penguin Books.

Hill, S. D., Brunton, D. H., Anderson, M. G., and Ji, W. (2018). Fighting talk: complex song elicits more aggressive responses in a vocally complex songbird. *Ibis* 160, 257–268.

Illes, A. E. (2014). Context of female bias in song repertoire size, singing effort, and singing independence in a cooperatively breeding songbird. *Behav. Ecol. Sociobiol.*, 1–12. doi:10.1007/s00265-014-1827-3.

Irwin, D. E., Bensch, S., Irwin, J. H., and Price, T. D. (2005). Speciation by distance in a ring species. *Science* 307, 414–416. doi:10.1126/science.1105201.

Irwin, D. E., Bensch, S., and Price, T. D. (2001). Speciation in a ring. *Nature* 409, 333. doi:10.1038/35053059.

Janik, V. M., and Slater, P. J. (1997). Vocal learning in mammals. *Adv. Study Behav.* 26, 59–100.

Jenkins, P. F. (1978). Cultural transmission of song patterns and dialect development in a free-living bird population. *Anim. Behav.* 26, 50–78. doi:10.1016/0003-3472(78)90007-6.

Lachlan, R. F., Verzijden, M. N., Bernard, C. S., Jonker, P.-P., Koese, B., Jaarsma, S., et al. (2013). The progressive loss of syntactical structure in bird song along an island colonization chain. *Curr. Biol.* 23, 1896–1901. doi:10.1016/j.cub.2013.07.057.

Laland, K. N. (2004). Social learning strategies. *Anim. Learn. Behav.* 32, 4–14. doi:10.3758/BF03196002.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292–305. doi:10.2307/2407393.

Langmore, N. E., Davies, N. B., Hatchwell, B. J., and Hartley, I. R. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 141–146.

Levin, R. N. (1996). Song behaviour and reproductive strategies in a duetting wren, *Thryothorus nigricapillus*: I. Removal experiments. *Anim. Behav.* 52, 1093–1106. doi:10.1006/anbe.1996.0257.

Ljubičić, I., Hyland Bruno, J., and Tchernichovski, O. (2016). Social influences on song learning. *Curr. Opin. Behav. Sci.* 7, 101–107. doi:10.1016/j.cobeha.2015.12.006.

Logue, D. M., Droessler, E. E., Roscoe, D. W., Vokey, J. R., Rendall, D., and Kunimoto, R. M. (2007). Sexually antithetical song structure in a duet singing wren. *Behaviour* 144, 331–350. doi:10.1163/156853907780425749.

Lynch, A., Plunkett, G. M., Baker, A. J., and Jenkins, P. F. (1989). A model of cultural evolution of chaffinch song derived with the meme concept. *Am. Nat.* 133, 634–653. doi:10.1086/284942.

Mann, N. I., Dingess, K. A., Barker, F. K., Graves, J. A., and Slater, P. J. B. (2009). A comparative study of song form and duetting in neotropical *Thryothorus* wrens. *Behaviour* 146, 1–43.

McNamara, A. (2011). Can we measure memes? *Front. Evol. Neurosci.* 3, 1. doi:10.3389/fnevo.2011.00001.

Mennill, D. J., Doucet, S. M., Newman, A. E. M., Williams, H., Moran, I. G., Thomas, I. P., et al. (2018). Wild birds learn songs from experimental vocal tutors. *Curr. Biol.* 28, 3273-3278.e4. doi:10.1016/j.cub.2018.08.011.

Mennill, D. J., and Rogers, A. C. (2006). Whip it good! Geographic consistency in male songs and variability in female songs of the duetting eastern whipbird *Psophodes olivaceus*. *J. Avian Biol.* 37, 93–100. doi:10.1111/j.0908-8857.2006.03548.x.

Mennill, D. J., and Vehrencamp, S. L. (2005). Sex differences in singing and duetting behavior of neotropical rufous-and-white wrens (*Thryothorus rufalbus*). *The Auk* 122, 175–186. doi:10.1642/0004-8038(2005)122[0175:SDISAD]2.0.CO;2.

Miskelly, C. (2013). Conservation translocations of New Zealand birds, 1863-2012. *Notornis* 60, 3–28.

Molles, L. E., Hudson, J. D., and Waas, J. R. (2006). The mechanics of duetting in a new zealand endemic, the kokako (*Callaeas cinerea wilsoni*): song at a snail's pace. *Ethology* 112, 424–436. doi:10.1111/j.1439-0310.2005.01168.x.

Mundinger, P. C. (1980). Animal cultures and a general theory of cultural evolution. *Ethol. Sociobiol.* 1, 183–223. doi:10.1016/0162-3095(80)90008-4.

Nordby, J. C., Campbell, S. E., and Beecher, M. D. (2002). Adult song sparrows do not alter their song repertoires. *Ethology* 108, 39–50. doi:10.1046/j.1439-0310.2002.00752.x.

Nottebohm, F. (1969). The "critical period" for song learning. *Ibis* 111, 386–387. doi:10.1111/j.1474-919X.1969.tb02551.x.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., and Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5. doi:10.1038/ncomms4379.

Oren, K., W., F. M., and Nicole, C. (2018). Integrative studies of cultural evolution: crossing disciplinary boundaries to produce new insights. *Philos. Trans. R. Soc. B Biol. Sci.* 373, 20170048. doi:10.1098/rstb.2017.0048.

Paris, D., Nicholls, A. O., Hall, A., Harvey, A., and Massaro, M. (2016). Female-biased dispersal in a spatially restricted endemic island bird. *Behav. Ecol. Sociobiol.* 70, 2061–2069. doi:10.1007/s00265-016-2210-3.

Parker, K. A., Anderson, M. J., Jenkins, P. F., and Brunton, D. H. (2012). The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecol. Lett.* 15, 778–785. doi:10.1111/j.1461-0248.2012.01797.x.

Paton, J. A., Manogue, K. R., and Nottebohm, F. (1981). Bilateral organization of the vocal control pathway in the budgerigar, *Melopsittacus undulatus*. *J. Neurosci.* 1, 1279–1288.

Payne, R. B., Payne, L. L., and Doehlert, S. M. (1988). Biological and cultural success of song memes in indigo buntings. *Ecology* 69, 104–117. doi:10.2307/1943165.

Perreault, C. (2012). The pace of cultural evolution. *PLoS ONE* 7, e45150. doi:10.1371/journal.pone.0045150.

Podos, J., and Warren, P. S. (2007). "The evolution of geographic variation in birdsong," in *Advances in the Study of Behavior*, eds. H. J. Brockmann, T. J. Roper, M. Naguib, K. E. Wynne-Edwards, C. Barnard, and J. Mitani (Academic Press), 403–458. Available at: http://www.sciencedirect.com/science/article/pii/S0065345407370095.

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., and Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature* 434, 455.

Potts, T. H. (1882). *Out in the open: a budget of scraps of natural history, gathered in New Zealand*. Lyttelton Times Company.

Potvin, D. A., and Clegg, S. M. (2015). The relative roles of cultural drift and acoustic adaptation in shaping syllable repertoires of island bird populations change with time since colonization. *Evolution* 69, 368–380. doi:10.1111/evo.12573.

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Behav. Evol. Ecol.* 3, 40. doi:10.3389/fevo.2015.00040.

Rasch, G., and Craig, J. L. (1988). Partitioning of nectar resources by New Zealand honeyeaters. *N. Z. J. Zool.* 15, 185–190. doi:10.1080/03014223.1988.10422613.

Richerson, P. J., Boyd, R., and Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *Proc. Natl. Acad. Sci.* 107, 8985–8992. doi:10.1073/pnas.0914631107.

Riebel, K., Odom, K. J., Langmore, N. E., and Hall, M. L. (2019). New insights from female bird song: towards an integrated approach to studying male and female communication roles. *Biol. Lett.* 15, 20190059.

Robertson, C. J. (1996). Tui (*Prosthemadera novaeseelandiae*) mimic parakeet calls at Raoul Island. *Notornis* 43, 52–53.

Rogers, A. C., Langmore, N. E., and Mulder, R. A. (2007). Function of pair duets in the eastern whipbird: cooperative defense or sexual conflict? *Behav. Ecol.* 18, 182–188. doi:10.1093/beheco/arl070.

Roper, M. M. (2018). Sexual dimorphism of song and life history trade-offs in the New Zealand bellbird: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy (Ph. D.) in Ecology at Massey University, Albany, New Zealand.

Rose, E. M., Mathew, T., Coss, D. A., Lohr, B., and Omland, K. E. (2018). A new statistical method to test equivalence: an application in male and female eastern bluebird song. *Anim. Behav.* 145, 77–85. doi:10.1016/j.anbehav.2018.09.004.

RStudio Team (2015). *RStudio: integrated development environment for R*. Boston, MA: RStudio, Inc. Available at: http://www.rstudio.com/.

Sanvito, S., Galimberti, F., and Miller, E. H. (2007). Observational evidences of vocal learning in southern elephant seals: a longitudinal study. *Ethology* 113, 137–146. doi:10.1111/j.1439-0310.2006.01306.x.

Searcy, W. A., and Yasukawa, K. (1995). *Polygyny and sexual selection in red-winged blackbirds*. Princeton University Press.

Shennan, S. (2011). Descent with modification and the archaeological record. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 1070–1079. doi:10.1098/rstb.2010.0380.

Slater, P. J. (1989). Bird song learning: causes and consequences. *Ethol. Ecol. Evol.* 1, 19–46.

Spurr (2008). Recapture of bellbird likely 8 years old. *Notornis* 55.

Templeton, C. N., Rivera-Cáceres, K. D., Mann, N. I., and Slater, P. J. B. (2011). Song duets function primarily as cooperative displays in pairs of happy wrens. *Anim. Behav.* 82, 1399–1407. doi:10.1016/j.anbehav.2011.09.024.

Tobias, J. A., Montgomerie, R., and Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2274–2293. doi:10.1098/rstb.2011.0280.

Todt, D., and Geberzahn, N. (2003). Age-dependent effects of song exposure: song crystallization sets a boundary between fast and delayed vocal imitation. *Anim. Behav.* 65, 971–979. doi:10.1006/anbe.2003.2127.

Vitt, D. H. (1979). The moss flora of the Auckland Islands, New Zealand, with a consideration of habitats, origins, and adaptations. *Can. J. Bot.* 57, 2226–2263. doi:10.1139/b79-272.

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., and Dale, J. (2016). Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. *Front. Ecol. Evol.* 4, 22.

Whitehead, H., and Rendell, L. (2014). *The cultural lives of whales and dolphins*. University of Chicago Press.

Whiten, A., Hinde, R. A., Laland, K. N., and Stringer, C. B. (2011). Culture evolves. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 938–948. doi:10.1098/rstb.2010.0372.

Wickler, W., and Seibt, U. (1980). Vocal duetting and the pair bond. *Z. Für Tierpsychol.* 52, 217–226. doi:10.1111/j.1439-0310.1980.tb00713.x.

Wickler, W., and Seibt, U. (1988). "Gender dialects" — a new phenomenon from birdsong epigenesis. *Naturwissenschaften* 75, 51–52. doi:10.1007/BF00367445.

Wilkins, M. R., Seddon, N., and Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. *Trends Ecol. Evol.* 28, 156–166. doi:10.1016/j.tree.2012.10.002.

Yamaguchi, A. (1998). A Sexually Dimorphic Learned Birdsong in the Northern Cardinal. *Condor Ornithol. Appl.* 100, 504–511. doi:10.2307/1369716.

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Mach. Learn.* 42, 31–60.

Zwartz, J. (2018). Woof Woof the talking tui. *Radio N. Z.* Available at: https://www.radionz.co.nz/national/programmes/ourchangingworld/audio/2018673079/woof-woof-the-talking-tui [Accessed March 3, 2019].

6 Conclusions and future directions

6.1 Overview

As a result of male-centric, northern-hemisphere-biased sexual selection theory, elaborate female traits in songbirds have been largely overlooked as unusual or non-functional by-products of male evolution (Darwin, 1871; Lande, 1980; Tobias et al., 2012). However, recent research has revealed that female song is present in *most* surveyed songbirds (Webb et al., 2016) and was in fact the ancestral condition to the clade (Odom et al., 2014). Additionally, a high proportion of songbird species have colourful females, and both song and showy colours have demonstrated female-specific functions in a growing number of species (see references in Webb et al., 2016). We have much to learn about the evolution and functions of elaborate female traits in general, and female song in particular. My thesis has aimed to extend the horizons of female birdsong in three ways (p. 15). First, it elucidates the evolutionary relationship between female song and plumage elaboration across the songbirds, demonstrating that these two traits have evolved in concert and suggesting overlapping, reinforcing functions. Second, the thesis introduces new accessible software tools for the measurement and analysis of song complexity, facilitating the analysis and comparison of female song. Third, I apply those software tools by showing—through a detailed field study on a large natural metapopulation—how vocal culture operates differentially in males and females. Here I highlight the significance (and limitations) of my findings for understanding elaborate female traits, and discuss future directions for building on the platform of this new knowledge.

6.2 Evolution of female song and plumage colouration

Significance and limitations

Understanding the evolutionary drivers of elaborate female traits is a key component of filling in the female song knowledge gap. This includes examining the evolutionary relationships *between* elaborate traits—for example, between female song and female plumage elaboration. In Chapter 2 (Webb et al., 2016) my co-authors and I used phylogenetically-informed analysis to reveal the relationship between female song presence and plumage elaboration across the songbirds. Our study showed that (1) female song is more prevalent amongst species with elaborated (classically "male-like") female plumage, (2) the co-occurrence of female song and female plumage elaboration is widely dispersed across the songbird phylogeny, suggesting many independent origins of this evolutionary association, (3) there is clear support for an evolutionary correlation between female song and female plumage colour elaboration, and (4) the presence of female song accounts for 16.2% of the variation in female plumage elaboration (and vice versa). These results suggest that contrary to the trade-off hypothesis (Darwin, 1871), female plumage colouration and female song likely evolved together under similar selection pressures and that their respective functions are reinforcing. Current selection theory suggests the primary context for this multimodal signalling is social competition for resources (Tobias et al., 2012), and I posited that the different signalling ranges of plumage and song may have favoured the evolutionary maintenance of both traits. Overall, the pattern reported in this study is consistent with female song and plumage colouration being functional signals (Webb et al., 2016), which I hope will encourage research into the functions and drivers of elaborate female traits.

As with all research, there are some limitations that warrant discussion. First, to score female song presence/absence, our study relied on descriptions of bird vocalisations contributed by hundreds of ornithologists in the Handbook of the Birds of the World (del Hoyo et al., 2004–2011). Remarkably, there is no standard definition for bird song. Although most bird enthusiasts would feel they recognise when a bird is singing, an adequate definition of bird song remains elusive, with over 80 criteria proposed (Spector, 1994). These criteria relate to function, structure, complexity, taxonomy, musicality, and volume, amongst other things. Some definitions even restrict song to male birds, evidencing the entrenched male bias in song research. Hundreds of contributors, each bringing implicit song definitions to their vocalisation descriptions, will certainly have introduced noise in our female song dataset. However, while this may have reduced precision, it is unlikely to have biased our results. That we found a strong evolutionary correlation between traits despite this noise demonstrates the robustness of the underlying biological pattern.

Second, at the time I undertook the research there was little quantitative information on female song complexity, and so analysis was limited to scoring female song as a binary presence/absence characteristic. As I highlighted in the paper, both traits are considered costly (Oberweger and Goller, 2001; Berg et al., 2005; Schmidt et al., 2013; McGraw et al., 2002; Walther and Clayton, 2005; Kleindorfer et al., 2016) and so there may be a trade-off in investment between the *degree* of song elaboration and plumage colour elaboration not apparent at the presence/absence level. As expounded below, it is now possible to build on my existing research to look at the evolution of female song *complexity* and whether it trades off with plumage elaboration.

Future directions: evolution of female song elaboration

Although we uncovered a strongly significant positive correlation between female song *presence* and female plumage colourfulness across the songbirds (Webb et al., 2016), there may be a trade-off in investment between the *degree* of song elaboration and plumage colour elaboration not apparent at the presence/absence level. Now that elaborated female traits have gained recognition as a major knowledge gap, recent high-profile requests for the documentation of female song (Odom and Benedict, 2018) have motivated widespread gathering of female song recordings into online repositories, such as the Cornell Female Birdsong Project (http://femalebirdsong.org/). There are steadily increasing numbers of female song recordings available from around the world. Our own bioacoustics software *Koe* (Fukuzawa, Webb et al., in review, Methods in Ecology and Evolution; koe.io.ac.nz) allows unprecedentedly streamlined and large-scale classification and analysis of song elaboration. With growing female song data and *Koe*'s novel, efficient tools for quantifying song elaboration, the resources are now available to explore the evolution of female song complexity and plumage elaboration in detail.

The songbirds (oscine passerines) are a recent branch of the evolutionary tree, originating in Australia/New Zealand then diversifying explosively across the globe (today accounting for ~50% of all bird species). Female song was likely present in the ancestral songbird (Odom et al., 2014), but nothing is yet known about the evolution of female song *complexity*—for example, whether female song complexity shows an evolutionary trade-off with other female elaboration, whether it is driven by environmental or life history factors (e.g. latitude), or what evolutionary pathways female song complexity has taken from the ancestral songbird to present-day species. Here I propose three important questions for future research on the evolution of elaborate female traits across the songbirds:

(1) What is the evolutionary relationship between female song *complexity* and female plumage colour elaboration?

Do these traits evolve in a concerted fashion, an antagonistic fashion, or in ways unrelated to each other? The answer will shed light on the costs and possible functions of the two traits. A negative association would suggest that the costs of complex song and elaborate plumage impose an evolutionary trade-off between traits (Darwin, 1871). That is, given limited resources, it is better to do one thing well rather than two things poorly. Alternatively, a positive evolutionary association would indicate that both traits are driven by similar selection pressures and that their functions are overlapping and reinforcing. Finally, a lack of relationship would suggest that the two traits are driven by separate pressures and that their functions evolve independently.

(2) Is female song more complex in the Southern Hemisphere—and if so, why?

The Southern Hemisphere, including Australasia, has many examples of complex female song (Molles et al., 2006; Bradley and Mennill, 2009; Brunton et al., 2016; Dalziell and Welbergen, 2016), whereas in the Northern Hemisphere, where most song research has been conducted, complex female song appears less common (Odom et al., 2014). I propose that future research (i) test for a quantitative difference in female song complexity between hemispheres, and (ii) resolve the phylogenetic, environmental and life-history factors associated with female song complexity. I predict that there are systemic differences in drivers between Northern and Southern Hemisphere species that can account for the apparent difference in female song complexity.

(3) What are the evolutionary trajectories of female song complexity?

By quantifying song complexity within and between songbird lineages, it will be possible to trace the pathways of female song evolution. Is complexity tightly linked to phylogeny, or is it more dependent on non-genetic factors (e.g. ecological and social environments)? Overall, has female song complexity tended to increase or decrease over evolutionary time?

6.3 *Koe* bioacoustics software

Significance and limitations

Classifying acoustic units is a key step in studying repertoires and sequence structure in animal communication (Kershenbaum et al., 2016). Manual classification by eye and ear remains the most widely used method for most species, but as increasingly large and diverse datasets are collected (e.g. facilitated by recent advances in automated field audio-recording), there is a pressing need for new tools and techniques to expedite the classification process. Existing bioacoustics software offers visualisation and labelling functionality only within one or few sound files at a time, not on the large comparative scale required—forcing researchers to resort to cumbersome and non-integrated workflows that severely limit the scope of bioacoustics research. *Koe* bioacoustics software (Chapter 4; Fukuzawa et al., 2020) was developed in response to this need in my own field study. I required a means of classifying my dataset of 21,500 bellbird song syllables in order to calculate repertoire sizes and assess repertoire overlap between sites, sexes and years. Through ongoing collaboration with computer scientist Yukio Fukuzawa and colleagues at Massey University, *Koe* has developed into a web-based, open-source bioacoustics tool (koe.io.ac.nz) for large-scale, high-resolution classification and analysis of animal vocalisations. It features tools for segmenting, comparing and classifying acoustic units and analysing diversity and sequence structure. *Koe*'s interactive ordination and unit tables provide a major advance in classification speed and robustness over existing methods, and a variety of measurements can be bulk-extracted from syllable datasets to analyse vocalisation complexity.

The potential of *Koe* for advancing female song research is substantial. First, by enabling unprecedentedly streamlined analysis of song complexity, female song elaboration can now be readily quantified within and between individuals, populations and species, and compared with conspecific male song. This will help to answer important questions of how female singing behaviour relates to social status, reproductive success, and other social/life-history factors (Brunton et al., 2016).

Second, being web-based and freely accessible, *Koe* provides a platform for online collaboration. This could be of immense benefit for democratising the widespread gathering and analysis of female song data urged by Odom and Benedict (2018). If researchers collaborate to build a centralised female song database in *Koe*, it will be possible to answer increasingly detailed questions on the function and evolution of female song. A further advantage of online collaboration is the potential for easy inter-observer reliability studies. For example, I have run third-year undergraduate sound analysis labs using *Koe* to collaboratively classify large datasets of bellbird song syllables. Syllable classification data for all collaborators is automatically saved to the *Koe* server and can be retrieved to collate and compare. By making it easy for many judges to label a dataset, *Koe* can facilitate robust classification validation compared to the traditional approach of classification by 1–2 researchers.

As a newly developed program, *Koe* has some limitations. While deemed user-friendly by fellow ornithologists and late-stage undergraduates, *Koe* has yet to be trialled on a non-academic general userbase. To encourage maximal uptake of our software (and its use in citizen science), accessibility is key. To this end, a *Koe* focus group to test program usability on non-specialists will be beneficial. Secondly, while *Koe* has proved a robust tool for bellbird song research, it has yet to be thoroughly trialled in other species. In theory the program is suitable for analysing any animal vocalisations with discrete acoustic units, but it is difficult to anticipate *a priori* the specific requirements of diverse taxa and diverse proximate and ultimate questions that interest bio-acousticians. Early adopters of *Koe* are currently trialling the program to study the vocalisations of common blackbird *Turdus merula*, bottlenosed dolphin *Tursiops truncatus*, North Island saddleback *Philesturnus rufusater*, short-tailed bat *Mystacina tuberculata*, superb fairy wren *Malurus cyaneus*, and more; as these projects progress, it will become apparent which aspects of the program require development to make *Koe* a robust and flexible platform.

Future development and applications of *Koe*

Recent advances in machine learning open possibilities for automating tasks in bioacoustics (Bianco et al., 2019). For example, segmenting audio recordings into constituent acoustic units has traditionally been a manual process; current automation based on amplitude thresholds, such as the *band-limited energy detector* in *Raven Pro* (The Cornell Lab of Ornithology, 2014), may work for simple, stereotyped vocalisations in restricted pitch ranges, but are inaccurate for complex song, especially in noisy environments. By contrast, machine learning trained on a dataset of manual segmentation can provide highly accurate, context-dependent segmentation robust to background noise (Figure 6.1). Implementing machine-learning-based automated segmentation in *Koe* could dramatically reduce user effort and further streamline the bioacoustics workflow.

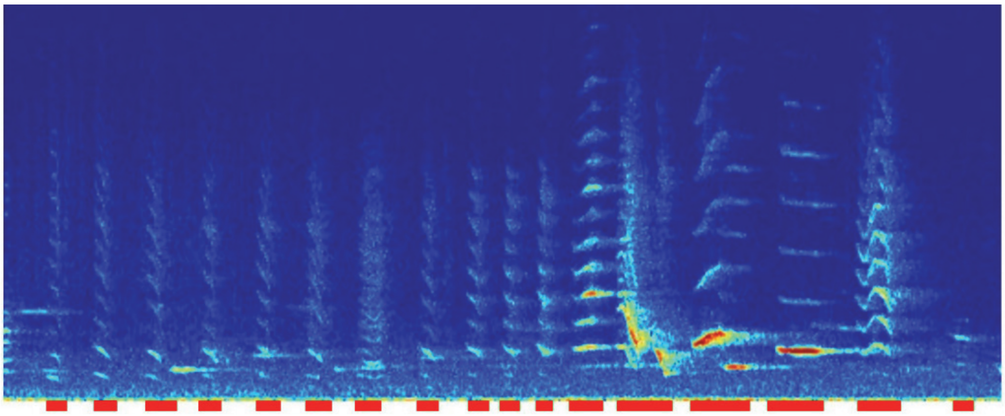


Figure 6.1—Development of auto-segmentation using machine learning in *Koe*.

Yukio Fukuzawa is developing machine-learning-based automatic segmentation. After training the algorithm on a small set of user-segmented songs, the software accurately delineates syllables (red bars), even in the presence of some background bird noise. Used with permission from Yukio Fukuzawa.

Another area of *Koe* with great development potential is syntax analysis. Currently this includes sequence mining algorithms (cSPADE) and network visualisations, which are useful for seeing which pairs of units tend to occur consecutively in acoustic sequences (see Appendix 7.2). An interesting, previously unexplored use of syntax analysis is to assist with refining syllable classification. To illustrate, if two user-defined syllable classes are both acoustically similar *and* sung in similar syntactic context, then these two classes are likely one class from the point of view of the vocalising animal. Thus, the software could recommend merging the two classes. If the user accepts the merge suggestion, acoustic and syntactic similarities between classes could be recalculated and the next pair of similar classes presented to the user for merging. Combining acoustic and syntactic information in this way would provide a novel, rapid and robust means of refining classification (in development by Yukio Fukuzawa).

Biologists habitually venture into places with poor internet connectivity, and sometimes need to process audio recordings in the field. Therefore we are working towards an offline version of *Koe* that can be deployed to a local machine. When an internet connection is re-established, the local version could sync data to the *Koe* server for the benefits of online collaboration and secure, centralised data storage.

Finally, it is worth emphasising that *Koe* is open source, with code freely available on GitHub (https://github.com/fzyukio/koe). As any developer can branch the code, the future potential of *Koe* is not limited to our (currently) small team. Our dream is that *Koe* continues to incorporate new technologies and techniques as they arise, to remain on the cutting edge of bioacoustics research.

6.4 Cultural evolution of female song

Significance and limitations

My study provides some of the first evidence of female song dialects in songbirds. Comparing between the male and female sectors of each population, I found comparable levels of syllable diversity, similar-shaped distributions of syllable type prevalence, and largely distinct syllable repertoires. These results are insightful because they demonstrate that female bellbird song is not a simplified subset of male song, as in some northern hemisphere species (Arcese et al., 1988); nor can it be construed as an accessory to male song, contributing token gestures to a largely male signal. Rather, with equal syllable diversity and largely separate repertoires, female bellbird song appears as much an elaborate signal as male song. If this is true of bellbirds, it is parsimonious to expect this to be true of other songbirds—which I hope will encourage song researchers to compare male and female song in many more species (Riebel et al., 2019).

Regarding patterns of sharing between sites, for both sexes most syllable types were site-specific; despite a large and varied repertoire within sites, generally the only types shared between sites were *Stutter* family syllables and pure-tone whistle notes of various pitches (*Pipe* family syllables). Between-site sharing of these syllable types decreased with distance for males, but there was no relationship between sharing and distance for females. Furthermore, despite females moving between sites more frequently, males had greater site–site repertoire overlap. These contrasting patterns of sharing across the archipelago indicate different cultural processes at play for the two sexes, perhaps stemming from sex differences in dispersal, meme mutation rates and song-learning modes (Graham et al., 2018a, 2018b)—all largely unknown for bellbirds. By taking a population-level approach, my study was limited in its ability to distinguish individual-level mechanisms of culture; to do so will require sophisticated field-based experiments as described below.

Future directions: mechanisms of male and female song culture

My field study has uncovered the *pattern* of male and female cultural diversification across an archipelago, focusing on population-level repertoires. However, to understand the *mechanisms* of cultural change it is necessary to combine these observations with an experimental approach (Mennill et al., 2018). In a recent research proposal, Professor Dianne Brunton presents an innovative strategy for determining the roles of individual-level processes in shaping population-level culture. At her chosen study site—Tawharanui—she proposes fitting bellbirds with passive integrated transponder (PIT) tags and backpack microphones (for individual identification and targeted song recording, respectively). By experimentally introducing new syllable types at Tawharanui via playback, it will be possible to monitor which birds learn the memes, and how the memes spread through the population. This could reveal the roles of song learning biases and individual movement—the two key processes hypothesised to regulate population-level culture (Aplin et al., 2014; Aplin, 2016; Tchernichovski et al., 2017; Somveille et al., 2018). For example, is meme transmission random, or are there learning biases towards prevalent or rare memes? Are dominant individuals preferred as song tutors? Are the learning modes of males and females different? The answers to these questions will illuminate the mechanisms underlying the patterns of meme sharing within and between sites I have observed.

6.5 Closing remarks

If you are reading this, you are almost certainly a PhD student looking for examples of structure and formatting for your own thesis. I did the same thing! Let me offer some words of camaraderie. First, the way my thesis turned out is very different to what I planned. Some goals that seemed straightforward were not. Research permits were slow, bellbirds played 'hard to get', expansion fronts failed to expand, databases got corrupted, and bizarre results boggled my brain. So take heart—setbacks are normal (Figure 6.2). Second, embrace collaboration! Through working with bird colouration expert Professor James Dale, and talented computer scientist Yukio Fukuzawa, opportunities arose that were far more rewarding than my original ideas. Not only is collaboration fun, but working across disciplines will broaden your mind and train you to communicate with those outside your field. Finally, let me close by imparting my most valuable nugget of wisdom: *Cambria* for body text and *Fira Sans* for headings. You're welcome.

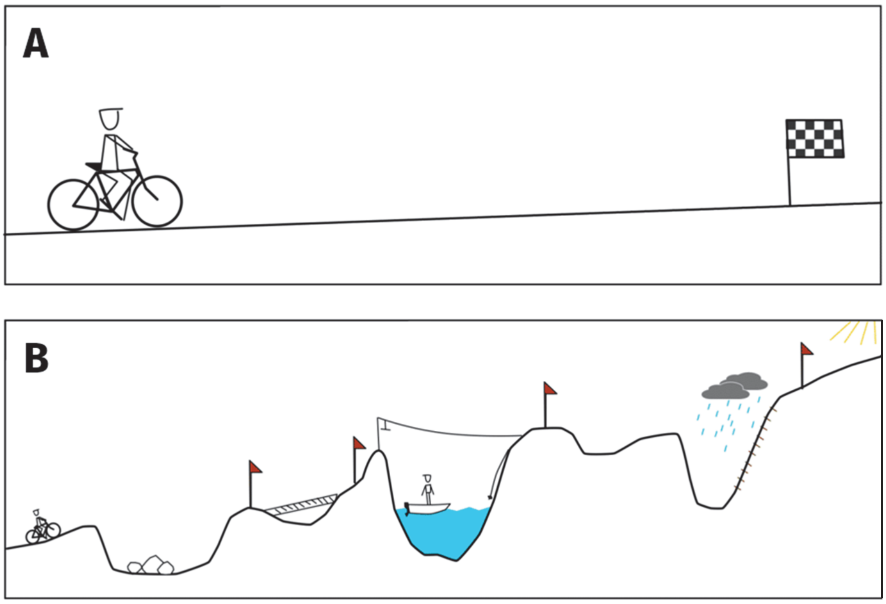


Figure 6.2—Expectations of thesis versus reality.

(**A**) Expectations. (**B**) Reality. Adapted from http://thedoghousediaries.com/5468.

6.6 References

Aplin, L. (2016). Understanding the multiple factors governing social learning and the diffusion of innovations. *Curr. Opin. Behav. Sci.* 12, 59–65. doi:10.1016/j.cobeha.2016.09.003.

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., and Sheldon, B. C. (2014). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature* 518, 538–541. doi:10.1038/nature13998.

Arcese, P., Stoddard, P. K., and Hiebert, S. M. (1988). The form and function of song in female song sparrows. *The Condor* 90, 44–50. doi:10.2307/1368431.

Berg, M. L., Beintema, N. H., Welbergen, J. A., and Komdeur, J. (2005). Singing as a handicap: the effects of food availability and weather on song output in the Australian reed warbler *Acrocephalus australis*. *J. Avian Biol.* 36, 102–109.

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., et al. (2019). Machine learning in acoustics: theory and applications. *ArXiv190504418 Phys.* Available at: http://arxiv.org/abs/1905.04418 [Accessed August 7, 2019].

Bradley, D. W., and Mennill, D. J. (2009). Solos, duets and choruses: vocal behaviour of the rufous-naped wren (*Campylorhynchus rufinucha*), a cooperatively breeding neotropical songbird. *J. Ornithol.* 150, 743–753.

Brunton, D. H., Roper, M. M., and Harmer, A. M. (2016). Female song rate and structure predict reproductive success in a socially monogamous bird. *Front. Ecol. Evol.* 4, 13.

Dalziell, A. H., and Welbergen, J. A. (2016). Elaborate mimetic vocal displays by female superb lyrebirds. *Front. Ecol. Evol.* 4. doi:10.3389/fevo.2016.00034.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Fukuzawa, Y., Webb, W. H., Pawley, M. D. M., Roper, M. M., Marsland, S., Brunton, D. H., et al. (2020). Koe: Web-based software to classify acoustic units and analyse sequence structure in animal vocalizations. *Methods Ecol. Evol.* doi:10.1111/2041-210X.13336.

Graham, B. A., Heath, D. D., Walter, R. P., Mark, M. M., and Mennill, D. J. (2018a). Parallel evolutionary forces influence the evolution of male and female songs in a tropical songbird. *J. Evol. Biol.* 31, 979–994. doi:10.1111/jeb.13279.

Graham, B. A., Heath, D. D., Walter, R. P., and Mennill, D. J. (2018b). Immigrant song: males and females learn songs after dispersal in a tropical bird. *Behav. Ecol.* 29, 711–723. doi:10.1093/beheco/ary029.

del Hoyo, J., Elliott, A., and Christie, D. (2004). *Handbook of the birds of the world Vol. 9–16*. Lynx Edicions.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev.* 91, 13–52. doi:10.1111/brv.12160.

Kleindorfer, S., Evans, C., and Mahr, K. (2016). Female in-nest chatter song increases predation. *Biol. Lett.* 12, 20150513.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292–305. doi:10.2307/2407393.

McGraw, K. J., Mackillop, E. A., Dale, J., and Hauber, M. E. (2002). Different colors reveal different information: how nutritional stress affects the expression of melanin- and structurally based ornamental plumage. *J. Exp. Biol.* 205, 3747–3755.

Mennill, D. J., Doucet, S. M., Newman, A. E. M., Williams, H., Moran, I. G., Thomas, I. P., et al. (2018). Wild birds learn songs from experimental vocal tutors. *Curr. Biol.* 28, 3273-3278.e4. doi:10.1016/j.cub.2018.08.011.

Molles, L. E., Hudson, J. D., and Waas, J. R. (2006). The mechanics of duetting in a new zealand endemic, the kokako (*Callaeas cinerea wilsoni*): song at a snail's pace. *Ethology* 112, 424–436. doi:10.1111/j.1439-0310.2005.01168.x.

Oberweger, K., and Goller, F. (2001). The metabolic cost of birdsong production. *J. Exp. Biol.* 204, 3379–3388.

Odom, K. J., and Benedict, L. (2018). A call to document female bird songs: Applications for diverse fields. *The Auk* 135, 314–325. doi:10.1642/AUK-17-183.1.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., and Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5. doi:10.1038/ncomms4379.

Riebel, K., Odom, K. J., Langmore, N. E., and Hall, M. L. (2019). New insights from female bird song: towards an integrated approach to studying male and female communication roles. *Biol. Lett.* 15, 20190059.

Schmidt, K. L., Moore, S. D., MacDougall-Shackleton, E. A., and MacDougall-Shackleton, S. A. (2013). Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Anim. Behav.* 86, 25–35.

Somveille, M., Firth, J. A., Aplin, L. M., Farine, D. R., Sheldon, B. C., and Thompson, R. N. (2018). Movement and conformity interact to establish local behavioural traditions in animal populations. *PLOS Comput. Biol.* 14, e1006647. doi:10.1371/journal.pcbi.1006647.

Spector, D. A. (1994). Definition in biology: the case of “bird song.” *J. Theor. Biol.* 168, 373–381. doi:10.1006/jtbi.1994.1117.

Tchernichovski, O., Feher, O., Fimiarz, D., and Conley, D. (2017). How social learning adds up to a culture: from birdsong to human public opinion. *J. Exp. Biol.* 220, 124–132. doi:10.1242/jeb.142786.

The Cornell Lab of Ornithology (2014). *Raven Pro: Interactive Sound Analysis Software*. Ithaca, NY Available at: http://www.birds.cornell.edu/raven.

Tobias, J. A., Montgomerie, R., and Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2274–2293. doi:10.1098/rstb.2011.0280.

Walther, B. A., and Clayton, D. H. (2005). Elaborate ornaments are costly to maintain: evidence for high maintenance handicaps. *Behav. Ecol.* 16, 89–95.

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., and Dale, J. (2016). Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. *Front. Ecol. Evol.* 4, 22.

7 Appendices

7.1 Appendix for Chapter 2—Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism

7.1.1 Reprint of published paper

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., & Dale, J. (2016). Female Song Occurs in Songbirds with More Elaborate Female Coloration and Reduced Sexual Dichromatism. *Frontiers in Ecology and Evolution*, 22. https://doi.org/10.3389/fevo.2016.00022

PDF follows. Further publicity of this research includes:

'Female Singers are Colourful'—Australasian Evolution Society article

I was invited to write an article summarising the research for the Australasian Evolution Society: http://ausevo.com/2016-02-10-female_song_dichromatism/

British Ornithological Union 2017 Twitter Conference (BOU17TC) presentation

I presented the research at the British Ornithological Union 2017 Twitter conference, in six tweets. Thumbnails and tweet links are provided below.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 15 March 2016
doi: 10.3389/fevo.2016.00022

Female Song Occurs in Songbirds with More Elaborate Female Coloration and Reduced Sexual Dichromatism

Wesley H. Webb[1]*, *Dianne H. Brunton*[1], *J. David Aguirre*[1], *Daniel B. Thomas*[1], *Mihai Valcu*[2] *and James Dale*[1]

[1] *Institute of Natural and Mathematical Sciences, Massey University, Auckland, New Zealand,* [2] *Department of Behavioral Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany*

OPEN ACCESS

Edited by:
Peter Schausberger,
University of Natural Resources and Life Sciences Vienna, Austria

Reviewed by:
Ken Kraaijeveld,
Vrije Universiteit Amsterdam, Netherlands
Karan J. Odom,
University of Maryland, Baltimore County, USA

***Correspondence:**
Wesley H. Webb
w.webb@massey.ac.nz

Specialty section:
This article was submitted to Behavioral and Evolutionary Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *06 November 2015*
Accepted: *29 February 2016*
Published: *15 March 2016*

Citation:
Webb WH, Brunton DH, Aguirre JD, Thomas DB, Valcu M and Dale J (2016) Female Song Occurs in Songbirds with More Elaborate Female Coloration and Reduced Sexual Dichromatism. Front. Ecol. Evol. 4:22. doi: 10.3389/fevo.2016.00022

Elaborate plumages and songs in male birds provide classic evidence for Darwinian sexual selection. However, trait elaboration in birds is not gender-restricted: female song has recently been revealed as a taxonomically-widespread trait within the songbirds (oscine Passerines), prompting increased research into likely functions and social/ecological correlates. Here we use phylogenetically-informed comparative analysis to test for an evolutionary association between female song and plumage color elaboration in songbirds. If there is an evolutionary trade-off between signaling modes, we predict a negative correlation between acoustic and visual elaboration. This trade-off hypothesis has been commonly proposed in males but has mixed empirical support. Alternatively, if song and plumage have similar or overlapping functions and evolve under similar selection pressures, we predict a positive correlation between female song and female plumage elaboration. We use published data on female song for 1023 species of songbirds and a novel approach that allows for the reliable and objective comparison of color elaboration between species and genders. Our results reveal a significant positive correlation between female colorfulness and female song presence. In species where females sing, females (but not males) are on average more colorful—with concomitantly reduced average sexual dichromatism. These results suggest that female plumage and female song likely evolved together under similar selection pressures and that their respective functions are reinforcing. We discuss the potential roles of sexual vs. social selection in driving this relationship, and the implications for future research on female signals.

Keywords: female song, multimodal signaling, oscine, Passeriformes, plumage coloration, trade-off, trait correlation

INTRODUCTION

Elaborate male traits often provide evidence for classic sexual selection (Darwin, 1871), but the possession of elaborate traits by females is less well understood. A traditional perspective holds that female trait elaboration is non-adaptive, perhaps detrimental, and results from "shared genetic architecture" with males (Darwin, 1871; Lande, 1980; see Tobias et al., 2012 for review). While current evidence does suggest that female ornamentation is correlated with conspecific male ornamentation to some extent (Bonduriansky and Chenoweth, 2009; Dale et al., 2015a) the view that elaborate female traits are purely non-adaptive pleiotropic effects has been strongly refuted by recent research.

Bird song and elaborate plumage are often considered costly traits (Song: Oberweger and Goller, 2001; Nowicki et al., 2002; Berg et al., 2005; Schmidt et al., 2013; but see Gil et al., 2006. Plumage: McGraw et al., 2002; Walther and Clayton, 2005; Simpson et al., 2015). In females, both traits can be evolutionarily labile with respect to the conspecific male phenotype. First, female plumage changes have played a greater role than male changes in the evolution of dichromatism in New World blackbirds (Icteridae; Irwin, 1994; Hofmann et al., 2008; Price and Eaton, 2014), tanagers (Thraupidae, Burns, 1998) and fairy-wrens (Maluridae, Johnson et al., 2013) and this pattern was shown to hold true for the order Passeriformes as a whole (Dale et al., 2015a).

Second, female song has been recovered as the ancestral condition of songbirds (Odom et al., 2014), indicating that multiple losses of female song have occurred over evolutionary time despite conspecific males retaining the trait (Price, 2015). This shows that female traits are not necessarily tightly constrained by the male phenotype, but are able to evolve rapidly and independently.

Furthermore, female-specific functions of song and plumage ornamentation have been resolved for a growing number of species, revealing these traits to be potentially adaptive for females. For instance, female song has been shown to function in territory and resource defense, mate attraction, mate defense, and pair-bonding (Searcy and Yasukawa, 1995; Langmore et al., 1996; Rogers et al., 2007; Brunton et al., 2008; Templeton et al., 2011; Hall et al., 2015). Female coloration is attractive to mates in some species (Amundsen et al., 1997; Smiseth and Amundsen, 2000; Murphy et al., 2009a,b) and frequently has roles in female-female competition for non-sexual breeding resources (Tobias et al., 2012; Morales et al., 2014).

If elaborated female traits are adaptive, this raises the question: what is the evolutionary relationship between female song and plumage color elaboration? We outline three hypotheses regarding the evolution of multiple elaborate traits: the traits may evolve "in a concerted fashion, in an antagonistic fashion, or in ways unrelated to each other" (Shutler and Weatherhead, 1990). In our context, these options are detailed as follows:

(i) The "trade-off" hypothesis (Darwin, 1871) argues that if two modes of signaling are both costly to produce or maintain, then selection might favor doing one thing well rather than two things badly. This predicts an evolutionary trade-off (an inverse correlation) between signaling modes (Shutler and Weatherhead, 1990). Thus, species with female song are predicted to have less colorful females on average than species lacking female song.

(ii) The "reinforcing signals" hypothesis (inspired by the "redundant signal" hypothesis of Møller and Pomiankowski, 1993) proposes that the two ornamental traits have overlapping, reinforcing functions, acting in concert to convey the condition or status of the signaler. Whereas either trait on its own provides a partial indication of signaler condition or status, in combination the multiple ornaments enable a more accurate assessment by rivals or mates. This predicts a co-evolution of the two traits; that is, a positive correlation between song and plumage color elaboration in females. Thus, species with female song are predicted to have more colorful females on average than species lacking female song.

(iii) The "multiple messages" hypothesis (adapted from Møller and Pomiankowski, 1993) posits that the two traits reveal different information about their bearers. This implies that the characters are driven by different selective pressures, and therefore elaborate vocal and visual sexual signals should evolve independently. Thus, there should be no difference in female color elaboration between species with female song, and those without female song.

Current evidence is inconclusive regarding the three hypotheses. In line with the trade-off hypothesis, an inverse relationship has been found between male plumage brightness and song complexity in cardueline finches (Badyaev et al., 2002). But in support of the reinforcing signals hypothesis, a positive correlation has been found between song length and number of colored patches among Asian barbets (Gonzalez-Voyer et al., 2013), and a positive correlation between the degree of dichromatism and time spent singing among wood warblers (Shutler and Weatherhead, 1990). Finally, Ornelas et al. (2009) found no relationship between dichromatism and song complexity among trogons, and Mason et al. (2014) found no correlation between song and plumage complexity among the tanagers. These conflicting results likely reflect biological and evolutionary differences among focal taxa, as well as methodological differences in how song and plumage were quantified (Mason et al., 2014). Furthermore, all these studies focus on males. The possibility of a generalized macroevolutionary association between multiple ornamental traits in females has not yet been investigated.

In this study we test for an evolutionary correlation between female song and plumage elaboration (male, female and dichromatism) across the songbirds (i.e., Oscines; order Passeriformes, suborder Passeri). We perform phylogenetically-informed comparative analysis using song and plumage data from repositories supplemented with additional data gleaned from the literature.

METHODS

Scoring Female Song

We compiled data describing the presence or absence of song in male and female songbirds (Oscines) from Odom et al. (2014) and del Hoyo et al. (2004–2011). We gave each species one of four scores according to the criteria of Odom et al. (2014, full details therein). Scores included: "present," both males and females of the species sing; "absent," only the male sings; "songless," neither sex sings; or "not enough information" if we could not reliably make a designation. (Note that no species where only females sing has been described.) Out of all 4814 songbird species, 1314 had sufficient information on song to reliably score the species. Because the lack of female song in songless species might be the product of different selection pressures than in species with male song, we omitted songless species (291 species) from our analysis. Our final species pool included 1023 singing species comprised of 656 species where both sexes sing (64%), and 367 species where only males sing (36%).

Scoring Plumage Elaboration

Plumage color scores for the 1023 songbird species were obtained from Dale et al. (2015b). Briefly, for both sexes of each species of passerine (Order: Passeriformes), the mean red, green and blue (RGB) values on 3 dorsal and 3 ventral patches were measured using digital image processing software (Valcu and Dale, 2013) on scanned images from handbook plates. For each patch of each sex of each species, it was determined how "male-like" that patch is by scoring the proportion of males in the nearest 1% of similarly colored patches in other species. The method results in scores where low values correspond to males or females with drab, classically-"female-like" plumage, and high scores correspond to males or females with elaborate, classically-"male-like" plumage. This approach is transferable to other color quantification methods. There is a high correlation between scores determined with handbook plates vs. analogous scores determined with UV-VIS (ultraviolet to visible) reflectance spectra from museum specimens (Dale et al., 2015a). This result provides critical validation of the method because although human and avian vision have considerable overlap (Badyaev and Hill, 2003; Seddon et al., 2010), birds can also see UV light not visible to humans (Cuthill, 2006). See Dale et al. (2015a) for detailed methodology. Sexual dichromatism was calculated for each species as the male plumage color score minus the female plumage score.

Trait Correlation Test

To test for an evolutionary correlation between female song and female plumage elaboration, we first performed Pagel's correlation test (Pagel, 1994), in R 3.1.2 (R Development Core Team, 2014) using the "geiger" and "phytools" packages (Harmon et al., 2008; Revell, 2012). The Pagel test controls for phylogenetic relatedness and requires no designation of independent and response variables. We assigned a song character state and a plumage character state to each tip of a phylogeny, and tested the null hypothesis that the two traits had evolved independently. As the Pagel test requires both traits to be binary, plumage scores were binned into binary characters according to an arbitrary cut-off, which was moved in integer increments from 35 (1022 of 1023 species with female plumage elaboration present) to 71 (1 of 1023 species), to study the sensitivity of the correlation test to changes in plumage cutoff value. The phylogenies used for this analysis were obtained from the Hackett backbone (Hackett et al., 2008) supertrees at http://birdtree.org (Jetz et al., 2012).

MCMCglmm Analysis

To estimate the strength of the correlation between female song presence and female plumage elaboration, we fit a multivariate generalized linear mixed model using the "MCMCglmm" package (Monte Carlo Markov Chain generalized linear mixed model; Hadfield, 2010) in R (version 3.1.2). MCMCglmm allowed us to fit a model which had a response vector that contained a mixture of Gaussian and non-Gaussian distributed variables. Female plumage elaboration is a continuous measure (see above); hence we assumed a Gaussian error distribution. For female song, the response vector contained binary presence scores (0 = female song absent, and 1 = female song present) and accordingly we assumed a Bernouli error distribution and used a logit link function. Fixed effects in our model included female plumage elaboration, the presence of female song, and male plumage elaboration as a covariate. Phylogeny was fit as a random effect using the methods described in Hadfield and Nakagawa (2010) to calculate the inverse numerator relationship for phylogenetic effects. For the phylogenetic effects we also allowed separate random intercepts for female song and female plumage elaboration and a non-zero covariance between these two traits by assuming an unstructured variance-covariance structure.

Priors for the location effects were diffuse about zero and had a large variance (10^8). For the variance components we used priors conforming to a scaled non-central F-distribution (Gelman, 2006) with the location parameter equal to zero. The scale parameter for female color elaboration was equal to half of the phenotypic variation in female color elaboration, and for female song the scale parameter was equal to $p(1 - p)$, where p is the mean probability of female song across the dataset. For the residual covariance matrix we assumed an inverse-Wishart distributed prior for female coloration. For female song (which is a binary trait), it is not possible to estimate a residual variance, so we fixed the prior at a value of 1 (Hadfield, 2014). The MCMC chain had 20,600,000 iterations, with a burn-in of 600,000 and a thinning interval of 20,000, resulting in ~1000 samples of the posterior distribution of the parameters. Model fit was confirmed by ensuring that autocorrelation was low and the trait means lay within the 95% highest posterior density (HPD) intervals of the posterior predictive distribution of each trait. To incorporate some of the uncertainty in the phylogenetic relationships among bird species, we applied the statistical model described above to 10 different phylogenetic trees randomly selected from http://birdtree.org (Jetz et al., 2012). Finally, we examined the convergence of the phylogenetic variances and covariances estimated from the 10 models (each using different trees, and therefore with different numerator relationship matrices for the phylogenetic effects), with the Gelman and Rubin (1992) diagnostic, R. For these 10 trees, the point estimate was $R = 1.2$ indicating moderate convergence. This phylogenetic uncertainly is incorporated in all the estimates of the posterior means and the HPD intervals we present.

To assess the significance of the phylogenetic correlation between female song (FS) and female plumage elaboration (FP), we first calculated the posterior distribution of the correlation using $cor_{(FS,\ FP)} = cov_{(FS,\ FP)} \div (var_{FS} \times var_{FP})^{0.5}$

If the 95% HPD intervals of the posterior distribution of the correlation did not overlap zero, we interpreted this as evidence for a significant phylogenetic correlation between female plumage elaboration and female song. Similarly, to assess the significance of the association between male color elaboration and our two female traits, we tested whether the 95% HPD intervals of the fixed interactions between male plumage and female plumage elaboration, as well as between male plumage and female song, overlapped zero.

RESULTS

Female Song in Relation to Plumage Color Elaboration

For many species in our sample, males and females have similar plumage elaboration scores (**Figure 1A**, points lying along the diagonal). However, there are also many sexually dichromatic species where the male is more colorful and the female is more drab (**Figure 1A**, points in the upper left of the distribution). Overlaying the density of female song presence on this plumage elaboration scatter (**Figure 1B**) reveals that female song presence is most concentrated amongst species with high female plumage scores and reduced average sexual dichromatism.

Phylogenetic Distribution

The phylogenetic distribution of the co-occurrence of female elaboration traits is visualized in **Figure 2**, where color-coded species tips correspond to the presence/absence of female song and female plumage elaboration. Instances where both traits co-occur (257 species, 25% of the total sample) are not concentrated within one or a few clades but are dispersed throughout the tree, suggesting many separate origins of the co-evolution of these traits (**Figure 2**). Wide phylogenetic dispersion persists when higher plumage elaboration cutoff values or different trees from Jetz et al. (2012) are used. Also numerous and widely dispersed around the tree are species with only one trait or the other: female song but not female plumage elaboration (399 species, 39% of the total sample); female plumage elaboration but not female song (72 species, 7% of the total sample), and neither trait present (295 species, 29% of the total sample).

Pagel Correlation Test

The Pagel test provides strong support for correlated evolution of female song presence and female plumage elaboration for all plumage cut-off values between 41 and 54 (**Figure 3**, $P < 0.0001$). This is a large range by comparison to the entire range of female plumage scores (35–71), and even at a cut-off as high as 60, the correlation is still statistically significant at the 5% level. The presence of an evolutionary association between female song and female color elaboration is therefore highly robust to alternative cut-off values used to categorize female coloration as elaborated.

MCMCglmm

Despite the phylogenetic uncertainty introduced by using 10 randomly selected phylogenetic trees from Jetz et al. (2012), we found a strongly significant evolutionary correlation between female plumage elaboration and the presence of female song (**Figure 4**, mean $cor_{(FS,\ FP)} = 0.402$, 95% HPD = 0.220–0.583). Accounting for evolutionary relationships among bird species we found that male and female color elaboration were positively associated (posterior mean = 0.472, 95% HPD = 0.422–0.527), confirming the apparent pattern seen in **Figure 1**. In addition there was a positive, but not statistically significant, relationship between the degree of male color elaboration and the presence of female song (posterior mean = 0.044, 95% HPD = $^{-}0.014$–0.102).

DISCUSSION

We used phylogenetically-informed analysis to reveal the relationship between female song presence and plumage elaboration across the songbirds. Our study shows that (1) female song is more concentrated amongst species with elaborated (classically "male-like") female plumage, (2) the co-occurrence of female song and female plumage elaboration is widely dispersed across the songbird phylogeny, suggesting many independent origins of this evolutionary association, (3) Pagel correlation tests demonstrate clear support for an evolutionary correlation between female song and female plumage color elaboration, and

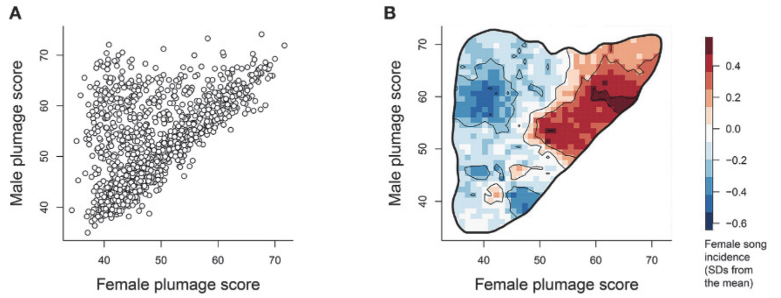


FIGURE 1 | (A) Male plumage score vs. female plumage score for each species ($N = 1023$). **(B)** Contour map depicting the average of the (scaled) female song values overlaid within the 99.8% volume contour of the male vs. female plumage score distribution. Values within the plot were calculated by superimposing a 50 × 50 grid over the scatter occurring between 30 and 80 and then calculating at each grid point the mean female song scores of the closest 3% of species. The units represent standard deviations away from the mean incidence of song, with redder values indicating increasingly higher occurrence of female song.

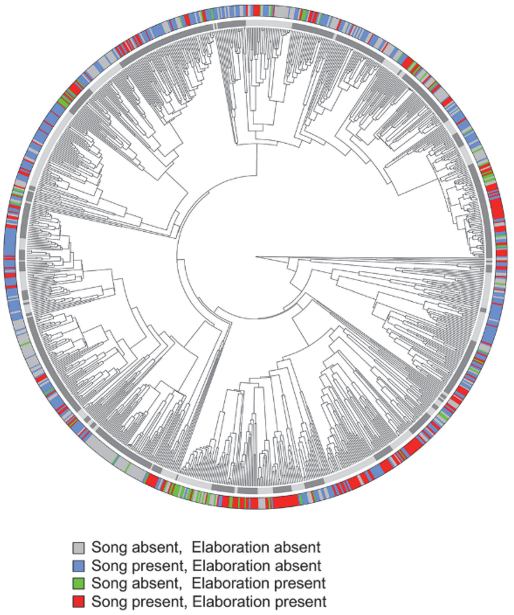


FIGURE 2 | Phylogenetic distribution of female song and female plumage elaboration. Plumage elaboration is here treated as a binary character with cutoff = plumage score of 50. Red labels represent species where female song and female plumage elaboration (plumage score >50) co-occur. The inner ring of alternating gray tones depicts boundaries of the 86 oscine families in our sample.

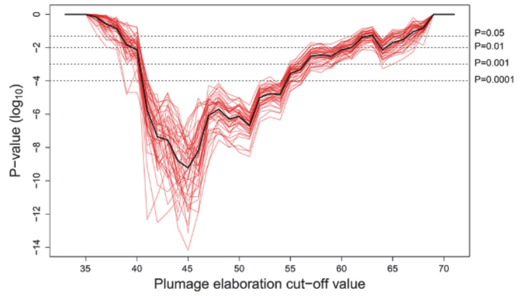


FIGURE 3 | Pagel correlation test *P*-values (Pagel, 1994) at all possible plumage elaboration cutoff values, for each of 50 phylogenetic trees from http://birdtree.org (Jetz et al., 2012). The black line indicates the mean of the 50 runs (red lines). Evidence of correlated evolution is strong (<0.0001) for all plumage cut-off values between 41 and 54.

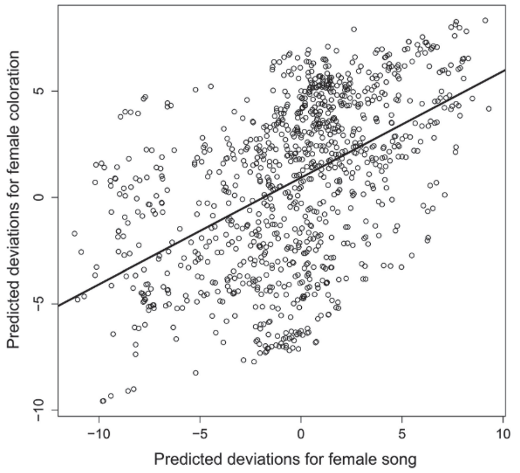


FIGURE 4 | Plot of best linear unbiased predictions of the phylogenetic association between female song and female plumage color elaboration. Points indicate the predicted phylogenetic deviations of each species (i.e., the 1023 species tips within the pruned phylogenetic tree) for the two female display traits, and solid line denotes the posterior mean phylogenetic correlation between female song with female plumage.

(4) estimates from MCMCglmm models suggest that the presence of female song accounts for 16.2% ($R = 0.402$) of the variation in female plumage elaboration (and vice versa).

Mason et al. (2014) provided the largest comparative study of multimodal sexual signaling to date, finding no relationship between plumage elaboration and song elaboration across the males of 301 tanager species (Thraupidae). In contrast, our results show a clear positive relationship between female song and female plumage color elaboration across the songbirds. The difference between our results and those of Mason et al. could stem from a number of factors. First, Mason et al. consider only male elaboration. The selective pressures acting on males and females are not necessarily equivalent, and so female traits may exhibit a different relationship to each other than do male traits (Tobias et al., 2012). Second, we use different methods of scoring song and plumage elaboration; notably, the song data of Mason et al. are quantitative, whereas our data are presence/absence. Finally, there are differences in scale between our studies. The presence of negative or non-significant relationships within some families such as the Thraupidae does not preclude a positive correlation at the broader taxonomic level of the songbirds as a whole. These considerations apply equally to other studies of multimodal signaling which find a negative relationship or no relationship between song and plumage (Badyaev et al., 2002; Ornelas et al., 2009).

Our results are most consistent with the "reinforcing signals" hypothesis; a positive correlation between female song and female plumage elaboration is expected if the two signaling modes tend to have reinforcing, overlapping functions, and thus have co-evolved together in response to similar selective pressures. Females may use song and plumage as a multimodal signal to reinforce the reliability and/or potency of the message to receivers.

A central tenet of sexual selection theory is that males and females often differ in their routes to reproductive success; for males, competition for mates is paramount, whereas for females, access to resources that affect fecundity is thought to be of greater importance (Rosvall, 2011). Therefore, the primary selection operating on females is likely to be non-sexual "social selection" for ecological or social resources, such as foraging territories, nest sites and paternal investment (West-Eberhard, 1979, 1983; Tobias et al., 2012). Such a view would suggest that a key function of both female song and plumage color elaboration may be signaling the female status (or resource-holding potential: Searcy and Nowicki, 2005) to competitive rivals, a prediction that is borne out in a number of studies (Tobias et al., 2012). This is not to deny that sexual selection for elaborate female traits may also be important (Clutton-Brock, 2007; Rosvall, 2011; Clutton-Brock and Huchard, 2013); indeed, the same ornaments can perform both sexual and (non-sexual) social functions (Kraaijeveld et al., 2004).

The co-occurrence of both visual and acoustic indicators of status (or resource-holding potential) likely reinforces the overall message and facilitates effective communication under different signaling scenarios. Song can be communicated over longer distances, without a clear line of sight; and as sound is propagated radially (Fahy and Gardonio, 2007), vocal communication does not depend on precise directionality of the signaler in relation to the receiver. Song is amenable to rapid temporal changes, conveying the short-term intentions of the individual and encoding complex information about signaler identity. Plumage coloration, by contrast, is a more permanent feature (though birds may be able to mediate plumage display behaviorally). In general, plumage coloration is a more direction-sensitive close-range signal than song. Thus, to signal quality and ward off rivals, selection could favor song for long-range broadcasting and plumage elaboration for close encounters, even if the message and intended receiver are the same.

If differing signaling scenarios promote female song or plumage elaboration differentially, this might explain the many species in our sample with only one trait or the other. Given that 16.2% of variation in female plumage elaboration is explained by female song presence (and vice versa), a remaining 83.8% of variation in each trait is thus attributable to other factors. That is, the relationship between female song and female plumage elaboration is complex and likely mediated by additional effects of, for example, habitat type, predation risk, territory size, and social structure. Closed habitat might strongly favor female song for effective communication if the range and efficacy of visual signaling is impaired. Or, for species experiencing high predation, elaborate female plumage may impose too great a risk around the nest (Martin and Badyaev, 1996) and be selected against. Or, in colony-living species where individuals remain in close proximity, close-range visual signaling may be sufficient for female signaling needs, making female song redundant. In short, there is much work left to do in identifying and quantifying the factors that contribute to visual and acoustic ornamentation in female songbirds.

If song and elaborate plumage are both costly (e.g., Oberweger and Goller, 2001; McGraw et al., 2002), as expected if the signals are "honest" indicators of aspects of female quality, then why have we not observed a trade-off between signaling modes? There are a number of potential explanations. First, while our results demonstrate that *presence* of female song does not trade off with female plumage elaboration, we lack information on the degree of female song *elaboration* (e.g., in terms of repertoire size, vocal agility, and time spent singing) and thus we cannot evaluate whether there is some degree of trade-off of resource allocation between modes. As more studies focus on quantifying the vocal performance of female birds, addressing whether singing females with bright plumage have "cheaper" songs than those with drab plumage will be practicable. It is also conceivable that when comparing across species, trade-offs might not manifest if different species are selected to invest different levels of resources into the overall message.

Another possibility is that resources involved in song and plumage development are not limiting for species where status signaling is strongly favored by selection. Rather than functioning as condition-dependent indices or handicaps, the honesty of these signals may instead be socially enforced, with cheaters (i.e., individuals who signal having greater status than they actually have) being punished through increased aggression by conspecifics (Tibbetts and Dale, 2004; Tibbetts and Izzo, 2010). The social cost in such cases will depend on how accurately the signal reflects true quality of the individual, rather than the number of ornaments involved in the signal *per se*. That is, under social costs, female song and plumage elaboration may be no more costly to produce than either trait alone, in which case we would not expect a trade-off between traits.

CONCLUSION AND FUTURE DIRECTIONS

We have demonstrated strong evidence for a positive co-evolutionary relationship between plumage elaboration and song in female songbirds, a result which supports an overlapping function of the two traits (i.e., the reinforcing signals hypothesis). We have suggested, in light of current selection theory (Tobias et al., 2012), that the primary context for this multimodal signaling is non-sexual social competition for ecological or social resources, and that the different signaling ranges of plumage and song may have favored the evolutionary maintenance of both traits. Our finding raises several questions for future research, including: (1) Does this pattern hold for the sub-oscines? Though generally poorer singers than the songbirds, and lacking vocal learning (but see Kroodsma et al., 2013), many sub-oscines vocalize for mate attraction and territorial defense (Chelén et al., 2005) and thus may be subject to similar evolutionary pressures. (2) Are female ornaments gained and lost more frequently than male ornaments (Kraaijeveld, 2014) and is there a consistent order of female trait evolution? That is, do gains (or losses) of song follow gains (or losses) in plumage elaboration, or vice versa? And (3), what are the social and ecological drivers of the evolutionary association of female song and female plumage elaboration? It is our hope that the patterns reported in this study contribute to future research on the functions of both visual and acoustic ornamentation in females.

FUNDING

WW and DB were supported by a Marsden Fund Grant (13-MAU-004) from the Royal Society of New Zealand.

AUTHOR CONTRIBUTIONS

Conceived of the study: WW, JD, and DB; collected the data: JD, MV, WW; analyzed the data: WW, JD, JA, DT; wrote the paper: WW, JD, and JA with input from the other authors.

ACKNOWLEDGMENTS

We thank Guest Associate Editors Michelle Hall and Naomi Langmore for initiating this special issue on female song, and the many ornithologists and scientists who have contributed data which made this study possible. Thanks to Aaron Harmer for his helpful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fevo.2016.00022

REFERENCES

Amundsen, T., Forsgren, E., and Hansen, L. T. T. (1997). On the function of female ornaments: male bluethroats prefer colourful females. *Proc. Biol. Sci.* 264, 1579–1586. doi: 10.1098/rspb.1997.0220

Badyaev, A. V., and Hill, G. E. (2003). Avian sexual dichromatism in relation to phylogeny and ecology. *Annu. Rev. Ecol. Evol. Syst.* 34, 27–49. doi: 10.1146/annurev.ecolsys.34.011802.132441

Badyaev, A. V., Hill, G. E., and Weckworth, B. V. (2002). Species divergence in sexually selected traits: increase in song elaboration is related to decrease in plumage ornamentation in finches. *Evolution* 56, 412–419. doi: 10.1111/j.0014-3820.2002.tb01350.x

Berg, M. L., Beintema, N. H., Welbergen, J. A., and Komdeur, J. (2005). Singing as a handicap: the effects of food availability and weather on song output in the Australian reed warbler *Acrocephalus australis*. *J. Avian Biol.* 36, 102–109. doi: 10.1111/j.0908-8857.2005.03285.x

Bonduriansky, R., and Chenoweth, S. F. (2009). Intralocus sexual conflict. *Trends Ecol. Evol.* 24, 280–288. doi: 10.1016/j.tree.2008.12.005

Brunton, D. H., Evans, B., Cope, T., and Ji, W. (2008). A test of the dear enemy hypothesis in female New Zealand bellbirds (*Anthornis melanura*): female neighbors as threats. *Behav. Ecol.* 19, 791–798. doi: 10.1093/beheco/arn027

Burns, K. J. (1998). A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. *Evolution* 52, 1219–1224. doi: 10.2307/2411252

Chelén, A. A. R., Garcia, C. M., and Riebel, K. (2005). Variation in the song of a sub-oscine, the vermilion flycatcher. *Behaviour* 142, 1115–1132. doi: 10.1163/156853905774405326

Clutton-Brock, T. (2007). Sexual selection in males and females. *Science* 318, 1882–1885. doi: 10.1126/science.1133311

Clutton-Brock, T. H., and Huchard, E. (2013). Social competition and selection in males and females. *Philos. Trans. R. Soc. B Biol. Sci.* 368, 20130074. doi: 10.1098/rstb.2013.0074

Cuthill, I. C. (2006). "Mechanisms and Measurements, Vol. 1," in *Bird Coloration*, eds G. E. Hill and K. J. McGraw (Cambridge, MA: Harvard University Press), 3–40.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015a). The effects of life-history and sexual selection on male and female plumage coloration. *Nature*. 527, 367–370. doi: 10.1038/nature15509

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., and Valcu, M. (2015b). Data from: the effects of life history and sexual selection on male and female plumage colouration. *Dryad Digit. Repository*. doi: 10.5061/dryad.1rp0s

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London: John Murray.

del Hoyo, J., Elliott, A., and Christie, D. (2004–2011). *Handbook of the Birds of the World*, Vol. 9–16. Barcelona: Lynx Edicions.

Fahy, F. J., and Gardonio, P. (2007). *Sound and Structural Vibration: Radiation, Transmission and Response*. Oxford, UK; Burlington, MA: Academic Press.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1, 515–534. doi: 10.1214/06-BA117A

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136

Gil, D., Naguib, M., Riebel, K., Rutstein, A., and Gahr, M. (2006). Early condition, song learning, and the volume of song brain nuclei in the zebra finch (*Taeniopygia guttata*). *J. Neurobiol.* 66, 1602–1612. doi: 10.1002/neu.20312

Gonzalez-Voyer, A., den Tex, R.-J., Castelló, A., and Leonard, J. A. (2013). Evolution of acoustic and visual signals in Asian barbets. *J. Evol. Biol.* 26, 647–659. doi: 10.1111/jeb.12084

Hackett, S. J., Kimball, R. T., Reddy, S., Bowie, R. C., Braun, E. L., Braun, M. J., et al. (2008). A phylogenomic study of birds reveals their evolutionary history. *Science* 320, 1763–1768. doi: 10.1126/science.1157704

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* 33, 1–22. doi: 10.18637/jss.v033.i02

Hadfield, J. D. (2014). *MCMCglmm Course Notes*. Available online at: https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf

Hadfield, J. D., and Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* 23, 494–508. doi: 10.1111/j.1420-9101.2009.01915.x

Hall, M. L., Rittenbach, M. R. D., and Vehrencamp, S. L. (2015). Female song and vocal interactions with males in a neotropical wren. *Behav. Evol. Ecol.* 3, 12. doi: 10.3389/fevo.2015.00012

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). GEIGER: investigating evolutionary radiations. *Bioinformatics* 24, 129–131. doi: 10.1093/bioinformatics/btm538

Hofmann, C. M., Cronin, T. W., and Omland, K. E. (2008). Evolution of sexual dichromatism. 1. Convergent losses of elaborate female coloration in New World orioles (*Icterus* spp.). *Auk* 125, 778–789. doi: 10.1525/auk.2008.07112

Irwin, R. E. (1994). The evolution of plumage dichromatism in the New World blackbirds: social selection on female brightness. *Am. Nat.* 144, 890–907. doi: 10.1086/285717

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., and Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature* 491, 444–448. doi: 10.1038/nature11631

Johnson, A. E., Jordan Price, J., and Pruett-Jones, S. (2013). Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecol. Evol.* 3, 3030–3046. doi: 10.1002/ece3.686

Kraaijeveld, K. (2014). Reversible trait loss: the genetic architecture of female ornaments. *Annu. Rev. Ecol. Evol. Syst.* 45, 159–177. doi: 10.1146/annurev-ecolsys-120213-091550

Kraaijeveld, K., Gregurke, J., Hall, C., Komdeur, J., Mulder, R. A. (2004). Mutual ornamentation, sexual selection, and social dominance in the black swan. *Behav. Ecol.* 15, 380–389. doi: 10.1093/beheco/arh023

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., et al. (2013). Behavioral evidence for song learning in the suboscine bellbirds (Procnias spp.; Cotingidae). *Wilson J. Ornithol.* 125, 1–14. doi: 10.1676/12-033.1

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292–305. doi: 10.2307/2407393

Langmore, N. E., Davies, N. B., Hatchwell, B. J., and Hartley, I. R. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proc. R. Soc. Lond. B. Biol. Sci.* 263, 141–146. doi: 10.1098/rspb.1996.0022

Martin, T. E., and Badyaev, A. V. (1996). Sexual dichromatism in birds: importance of nest predation and nest location for females versus males. *Evolution* 50, 2454–2460. doi: 10.2307/2410712

Mason, N. A., Shultz, A. J., and Burns, K. J. (2014). Elaborate visual and acoustic signals evolve independently in a large, phenotypically diverse radiation of songbirds. *Proc. R. Soc. Lond. B Biol. Sci.* 281, 20140967. doi: 10.1098/rspb.2014.0967

McGraw, K. J., Mackillop, E. A., Dale, J., and Hauber, M. E. (2002). Different colors reveal different information: how nutritional stress affects the expression of melanin- and structurally based ornamental plumage. *J. Exp. Biol.* 205, 3747–3755.

Møller, A. P., and Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behav. Ecol. Sociobiol.* 32, 167–176. doi: 10.1007/BF00173774

Morales, J., Gordo, O., Lobato, E., Ippi, S., Martínez-de la Puente, J., Tomás, G., et al. (2014). Female-female competition is influenced by forehead patch expression in pied flycatcher females. *Behav. Ecol. Sociobiol.* 68, 1195–1204. doi: 10.1007/s00265-014-1730-y

Murphy, T. G., Hernández-Muciño, D., Osorio-Beristain, M., Montgomerie, R., and Omland, K. E. (2009a). Carotenoid-based status signaling by females in the tropical streak-backed oriole. *Behav. Ecol.* 20, 1000–1006. doi:10.1093/beheco/arp089

Murphy, T. G., Rosenthal, M. F., Montgomerie, R., and Tarvin, K. A. (2009b). Female American goldfinches use carotenoid-based bill coloration to signal status. *Behav. Ecol.* 20, 1348–1355. doi: 10.1093/beheco/arp140

Nowicki, S., Searcy, W., and Peters, S. (2002). Brain development, song learning and mate choice in birds: a review and experimental test of the "nutritional stress hypothesis." *J. Comp. Physiol. A* 188, 1003–1014. doi: 10.1007/s00359-002-0361-3

Oberweger, K., and Goller, F. (2001). The metabolic cost of birdsong production. *J. Exp. Biol.* 204, 3379–3388.

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., and Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5:3379. doi: 10.1038/ncomms4379

Ornelas, J. F., González, C., and Espinosa De Los Monteros, A. (2009). Uncorrelated evolution between vocal and plumage coloration traits in the trogons: a comparative study. *J. Evol. Biol.* 22, 471–484. doi: 10.1111/j.1420-9101.2008.01679.x

Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. Biol. Sci.* 255, 37–45. doi: 10.1098/rspb.1994.0006

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Behav. Evol. Ecol.* 3:40. doi: 10.3389/fevo.2015.00040

Price, J. J., and Eaton, M. D. (2014). Reconstructing the evolution of sexual dichromatism: current color diversity does not reflect past rates of male and female change. *Evolution* 68, 2026–2037. doi: 10.1111/evo.12417

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things): phytools: R package. *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Rogers, A. C., Langmore, N. E., and Mulder, R. A. (2007). Function of pair duets in the eastern whipbird: cooperative defense or sexual conflict? *Behav. Ecol.* 18, 182–188. doi: 10.1093/beheco/arl070

Rosvall, K. A. (2011). Intrasexual competition in females: evidence for sexual selection? *Behav. Ecol.* 22, 1131–1140. doi: 10.1093/beheco/arr106

Schmidt, K. L., Moore, S. D., MacDougall-Shackleton, E. A., and MacDougall-Shackleton, S. A. (2013). Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Anim. Behav.* 86, 25–35. doi: 10.1016/j.anbehav.2013.03.036

Searcy, W. A., and Nowicki, S. (2005). The *Evolution of Animal Communication: Reliability and Deception in Signaling Systems*. Princeton, NJ: Princeton University Press.

Searcy, W. A., and Yasukawa, K. (1995). *Polygyny and Sexual Selection in Red-Winged Blackbirds*. Princeton, NJ: Princeton University Press.

Seddon, N., Tobias, J. A., Eaton, M., and Ödeen, A. (2010). Human vision can provide a valid proxy for avian perception of sexual dichromatism. *Auk* 127, 283–292. doi: 10.1525/auk.2009.09070

Shutler, D., and Weatherhead, P. J. (1990). Targets of sexual selection: song and plumage of wood warblers. *Evolution* 44, 1967–1977. doi: 10.2307/2409607

Simpson, R. K., Johnson, M. A., and Murphy, T. G. (2015). Migration and the evolution of sexual dichromatism: evolutionary loss of female coloration with migration among wood-warblers. *Proc. R. Soc. Lond. B Biol. Sci.* 282, 20150375. doi: 10.1098/rspb.2015.0375

Smiseth, P. T., and Amundsen, T. (2000). Does female plumage coloration signal parental quality? A male removal experiment with the bluethroat (Luscinia s. svecica). *Behav. Ecol. Sociobiol.* 47, 205–212. doi: 10.1007/s002650050657

Templeton, C. N., Rivera-Cáceres, K. D., Mann, N. I., and Slater, P. J. B. (2011). Song duets function primarily as cooperative displays in pairs of happy wrens. *Anim. Behav.* 82, 1399–1407. doi: 10.1016/j.anbehav.2011.09.024

Tibbetts, E. A., and Dale, J. (2004). A socially enforced signal of quality in a paper wasp. *Nature* 432, 218–222. doi: 10.1038/nature02949

Tibbetts, E. A., and Izzo, A. (2010). Social punishment of dishonest signalers caused by mismatch between signal and behavior. *Curr. Biol.* 20, 1637–1640. doi: 10.1016/j.cub.2010.07.042

Tobias, J. A., Montgomerie, R., and Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2274–2293. doi: 10.1098/rstb.2011.0280

Valcu, M., and Dale, J. (2013). *colorZapper: A Platform for Scoring Biological Coloration*. R package version 1.0. Available online at: https://github.com/valcu/colorZapper

Walther, B. A., and Clayton, D. H. (2005). Elaborate ornaments are costly to maintain: evidence for high maintenance handicaps. *Behav. Ecol.* 16, 89–95. doi: 10.1093/beheco/arh135

West-Eberhard, M. J. (1979). Sexual selection, social competition, and evolution. *Proc. Am. Philos. Soc.* 123, 222–234.

West-Eberhard, M. J. (1983). Sexual selection, social competition, and speciation. *Q. Rev. Biol.* 58, 155–183. doi: 10.1086/413215

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

https://twitter.com/wesleythewebb/status/935712425235578886

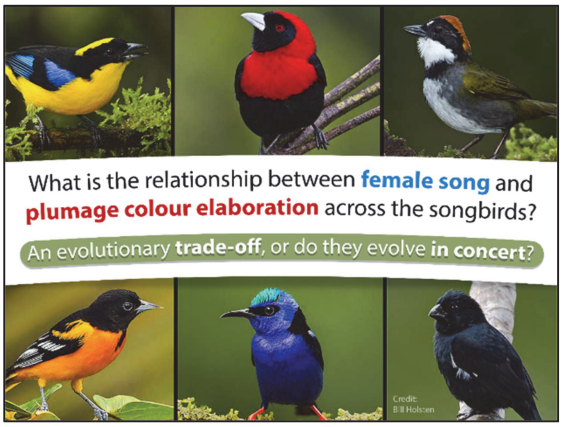
https://twitter.com/wesleythewebb/status/935712926626754560

https://twitter.com/wesleythewebb/status/935713445063794688

https://twitter.com/wesleythewebb/status/935713932005711872

https://twitter.com/wesleythewebb/status/935714434735071232

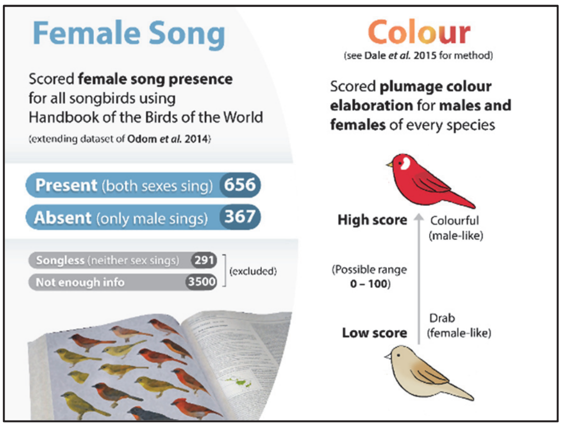
https://twitter.com/wesleythewebb/status/935714939095748608



What is the relationship between **female song** and **plumage colour elaboration** across the songbirds?

An evolutionary **trade-off**, or do they evolve **in concert**?

Credit: Bill Holsten



Female Song

Scored **female song presence** for all songbirds using Handbook of the Birds of the World

(extending dataset of **Odom *et al.* 2014**)

Present (both sexes sing) **656**

Absent (only male sings) **367**

Songless (neither sex sings) **291**

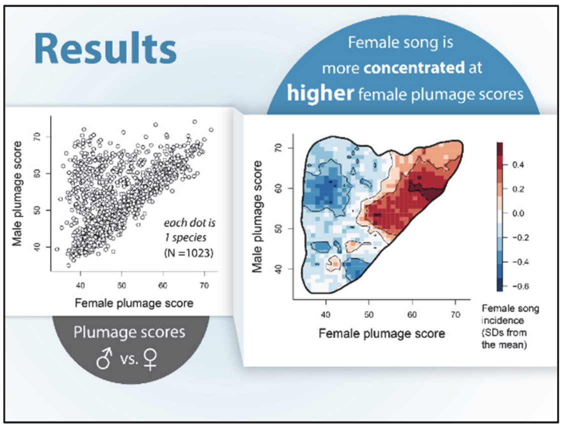
Not enough info **3500**

(excluded)

Colour

(see **Dale *et al.* 2015** for method)

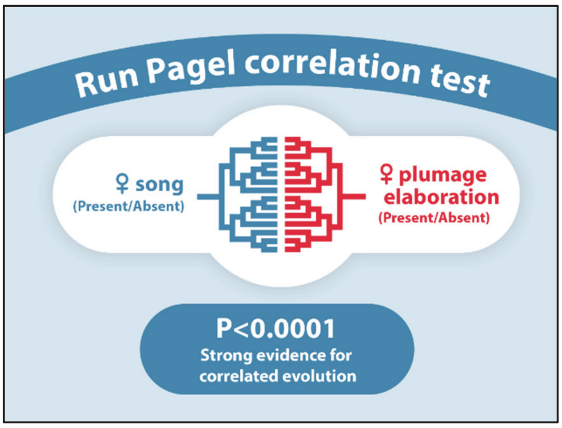
Scored **plumage colour elaboration** for **males and females** of every species



Results

Female song is more **concentrated** at **higher** female plumage scores

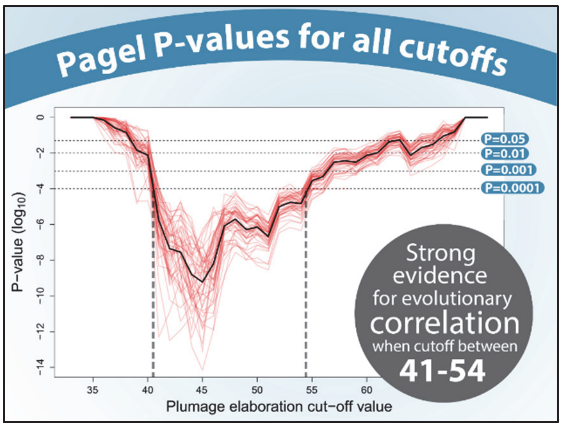
Plumage scores ♂ vs. ♀



Run Pagel correlation test

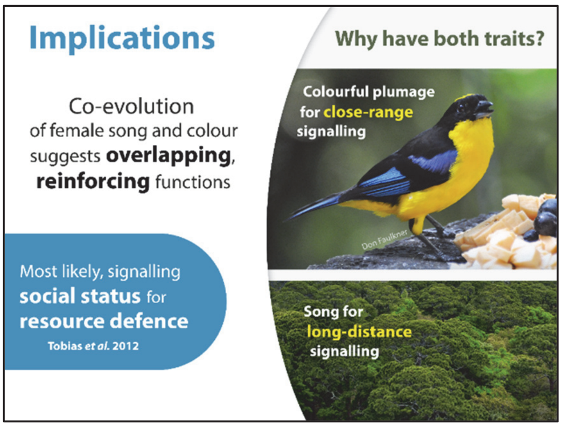
P<0.0001

Strong evidence for correlated evolution



Pagel P-values for all cutoffs

Strong evidence for evolutionary correlation when cutoff between **41-54**



Implications

Co-evolution of female song and colour suggests **overlapping**, **reinforcing** functions

Most likely, signalling **social status** for **resource defence**

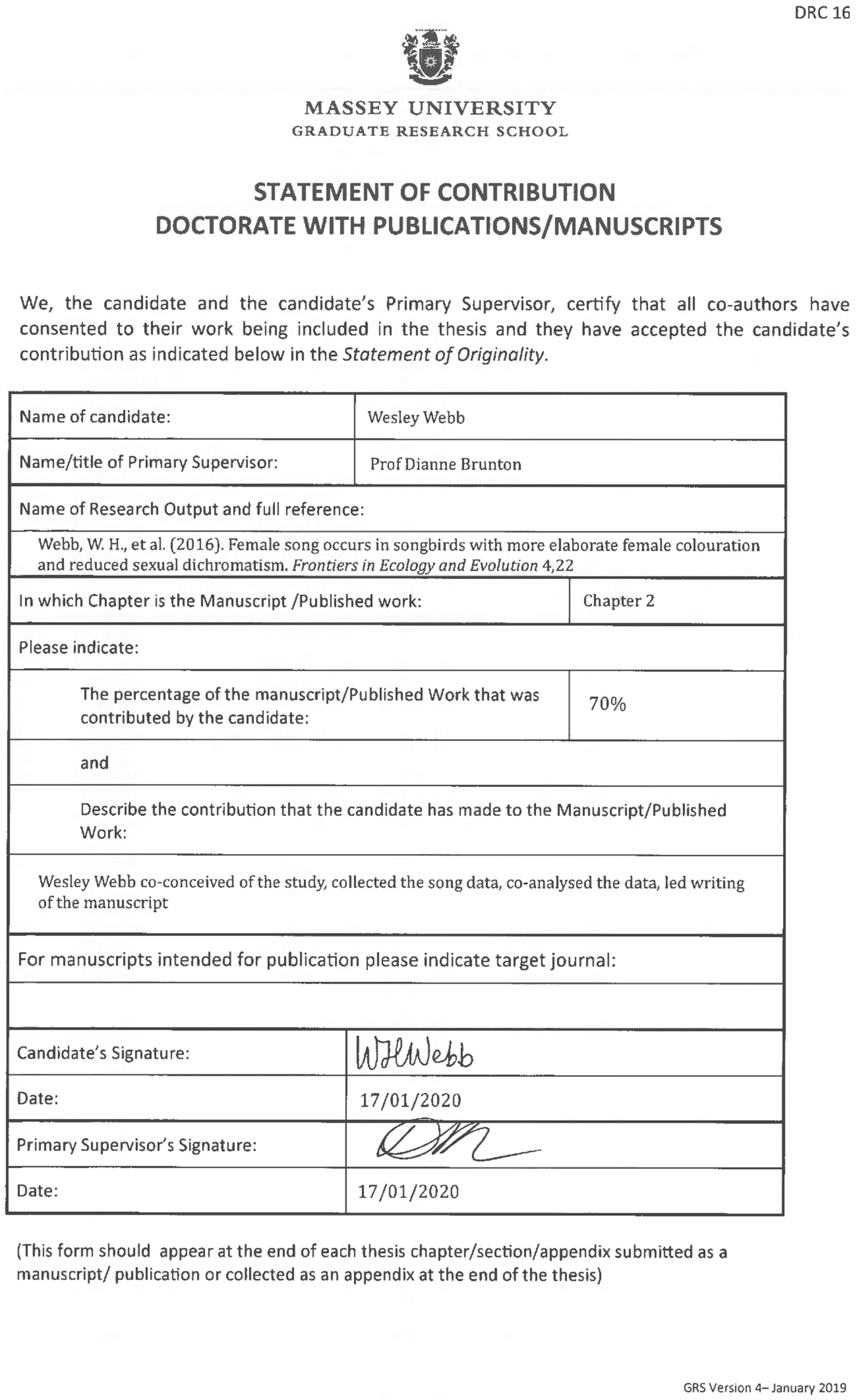
Tobias *et al.* 2012

Why have both traits?

Colourful plumage for **close-range** signalling

Song for **long-distance** signalling

Don Faulkner



DRC 16

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Wesley Webb
Name/title of Primary Supervisor:	Prof Dianne Brunton
Name of Research Output and full reference:	
Webb, W. H., et al. (2016). Female song occurs in songbirds with more elaborate female colouration and reduced sexual dichromatism. *Frontiers in Ecology and Evolution* 4,22	
In which Chapter is the Manuscript /Published work:	Chapter 2
Please indicate:	
The percentage of the manuscript/Published Work that was contributed by the candidate:	70%
and	
Describe the contribution that the candidate has made to the Manuscript/Published Work:	
Wesley Webb co-conceived of the study, collected the song data, co-analysed the data, led writing of the manuscript	
For manuscripts intended for publication please indicate target journal:	
Candidate's Signature:	WHWebb
Date:	17/01/2020
Primary Supervisor's Signature:	
Date:	17/01/2020

(This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis)

GRS Version 4– January 2019

7.2 Appendix for Chapter 4—*Koe*: web-based software for classifying acoustic units and analysing sequence structure in animal vocalisations

7.2.1 Reprint of published paper

Yukio Fukuzawa*, Wesley Webb*, Matthew Pawley, Michelle Roper, Stephen Marsland, Dianne Brunton, & Andrew Gilman. (2020). *Koe*: Web-based software to classify acoustic units and analyse sequence structure in animal vocalisations. *Methods in Ecology and Evolution*. Online Early, doi:10.1111/2041-210X.13336.

*These two authors contributed equally as co-first authors.

Received: 1 May 2019 | Accepted: 28 October 2019
DOI: 10.1111/2041-210X.13336

APPLICATION

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Koe: Web-based software to classify acoustic units and analyse sequence structure in animal vocalizations

Yukio Fukuzawa[1] | Wesley H. Webb[1] | Matthew D.M. Pawley[1] | Michelle M. Roper[1] | Stephen Marsland[2] | Dianne H. Brunton[1] | Andrew Gilman[1]

[1]School of Natural and Computational Sciences, Massey University, Auckland, New Zealand

[2]School of Mathematics and Statistics, Victoria University, Wellington, New Zealand

Correspondence
Wesley H. Webb
Email: w.webb@massey.ac.nz

Funding information
Massey University Doctoral Scholarship; Royal Society of New Zealand Marsden Fund, Grant/Award Number: 13-MAU-004, 17-MAU-154 and 17-UOA-295

Handling Editor: Luca Börger

Abstract

1. Classifying acoustic units is often a key step in studying repertoires and sequence structure in animal communication. Manual classification by eye and ear remains the primary method, but new tools and techniques are urgently needed to expedite the process for large, diverse datasets.
2. Here we introduce *Koe*, an application for classifying and analysing animal vocalizations. *Koe* offers bulk-labelling of units via interactive ordination plots and unit tables, as well as visualization and playback, segmentation, measurement, data filtering/exporting and new tools for analysing repertoire and sequence structure – in an integrated environment.
3. We demonstrate *Koe* with a real-world case study of New Zealand bellbird *Anthornis melanura* songs from an archipelago metapopulation. Having classified 21,500 units in *Koe*, we compare repertoires and sequence structure between sites and sexes.
4. *Koe* is web-based (koe.io.ac.nz.) and easy to use, making it ideal for collaboration, education and citizen science. By enabling large-scale, high-resolution classification and analysis of animal vocalizations, *Koe* expands the possibilities for bioacoustics research.

KEYWORDS
cataloguing, clustering, feature extraction, labelling, sequence, song, syllable, syntax

1 | INTRODUCTION

Acoustic communication is important for many species across the animal kingdom, used in contexts such as territorial defence, mate attraction, group coordination, and recognition of individuals and species (Catchpole & Slater, 2008; Janik & Sayigh, 2013; Peters, Searcy, & Marler, 1980; Soltis, 2009; Sperber, Werner, Kappeler, & Fichtel, 2017; Starnberger, Preininger, & Hödl, 2014). Often, acoustic communication is structured as a sequence of distinct acoustic units, where information is encoded in the types of units and sometimes their sequence structure (i.e. temporal arrangement) (Kershenbaum et al., 2016). In animals with simple, stereotyped songs, it may be possible to derive useful measures of similarity by comparing entire songs. However, many species have complex, flexible vocalizations, where unit-level (e.g. syllable) analysis is essential, for example, to assess unit sharing or evaluate sequence structure. A common workflow for studying such cases is summarized in Figure 1 (see Kershenbaum et al., 2016 for a review).

Classifying units involves identifying unit classes (i.e. producing a 'catalogue' of class types) and then assigning units to those classes;

Yukio Fukuzawa and Wesley H. Webb contributed equally.

Dianne H. Brunton and Andrew Gilman are co-seniors.

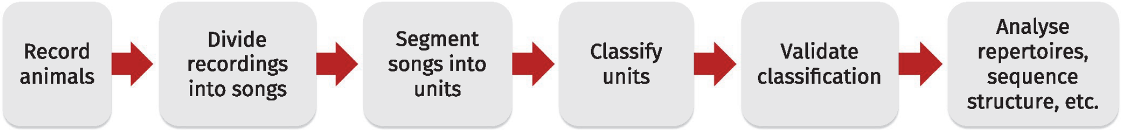


FIGURE 1 A typical process for studies in bioacoustics

only once units have been classified can repertoires and sequence structures be analysed and compared. Classification is a challenge, however; thus far, attempts at automation have not proven generalizable (Priyadarshani, Marsland, & Castro, 2018) and the primary approach for most species remains manual classification based on human visual and auditory perception (Kershenbaum et al., 2016).

Although human perception-based methods excel at judging unit similarity (Duda, Hart, & Stork, 2012; Sayigh, Esch, Wells, & Janik, 2007), they can nonetheless be very time consuming. The researcher needs to view spectrograms and hear playback of units, comparing similarity of many units at a time, in order to discern and assign classes. We know of no tools designed for this purpose; existing acoustics software offers visualization and labelling functionality only within one or few files at a time, not on the large comparative scale required (Table S1). Researchers currently resort to cumbersome and non-integrated workflows that severely limit the scope of bioacoustics research.

Here we introduce *Koe*, new web-based software designed to expedite large-scale classification and analysis of acoustic units. *Koe* offers bulk-labelling of units via novel interactive ordination plots, as well as visualization and playback, segmentation, measurement, data filtering/exporting and new tools for analysing sequence structure, in an integrated environment.

Koe was developed in *Python 3* (Rossum, 1995), which provides flexible server-side computation (Jones, Oliphant, & Peterson, 2001; Oliphant, 2007; Pedregosa et al., 2012). For the web framework, we chose *Django* (Django Software Foundation, 2013). Being web-based gives *Koe* numerous benefits: it is accessible from any device, requires no setup, supports multiple simultaneous users, and can run in multiple browser windows at once (useful for viewing different program views concurrently). This makes it ideal for collaboration, teaching and citizen science. The official application, maintained by the authors, is available at koe.io.ac.nz. Source code is available at http://github.com/fzyukio/koe.

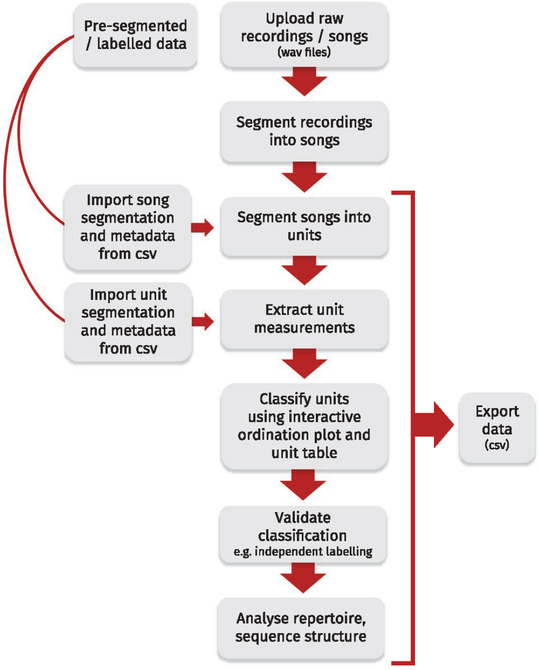


FIGURE 2 Suggested *Koe* workflow. *Koe* data is integrated; any step in the workflow can be revisited, and modifications will dynamically update throughout the program. Raw recordings/songs (in wav format) are uploaded. Recordings are divided into songs, which are segmented into units. *Koe* extracts unit features to produce similarity indices and interactive ordination plots, which help a user to rapidly and accurately classify units. Once complete, classification can be validated by independent labelling. *Koe* provides tools for analysing repertoires and sequence structure. Data can also be exported for external use

2 | KOE WORKFLOW

We designed *Koe* to provide end-to-end functionality in an intuitive and flexible workflow (Figure 2). Each step in the workflow has a program view tailored to that task. Here we highlight significant features of each program view; for more detail, see *Koe* documentation (http://github.com/fzyukio/koe/wiki).

2.1 | Upload and segment recordings

In *Koe*, the user can partition raw recordings into individual vocalization bouts, termed 'songs' (*Upload & split raw recordings* view), and segment songs into their constituent acoustic units (*Segment songs into units* view). The tasks are similar; they involve the user selecting start/endpoints of songs or units, respectively, on a spectrogram, then committing the selections to the database (stored securely on the *Koe* server). It is up to the user to decide on segmentation criteria appropriate to their species/question. The user can adjust playback speed, spectrogram contrast and time-axis zoom. Saved selections become available in other program views. Acoustic units are not stored as audio segments, but as start/endpoint information referencing the source song; the user can freely readjust unit segmentation and program views will update dynamically. For any songs already segmented into units in other software, start/endpoints can be imported as a csv file. Pre-partitioned song files can also be uploaded directly to the database in *View all songs*. Users can grant database access to other users, with custom permission levels.

2.2 | Extract acoustic features from units

In quantitative analysis, raw signals are often replaced by their compact representations using a specific set of features, extracted from the signals. Analyses performed on compact representations can be more effective and computationally efficient. In *Extract unit features* view, *Koe* can extract a wide array of spectral-, tempo- and chroma-related features for bioacoustic analyses (see http://github.com/fzyukio/koe/wiki). Extracted features are utilized by *Koe* to construct ordinations and similarity indices (sections 2.3.1 and 2.3.2), and can be exported as csv files for analysis in other software.

2.3 | Classify units

Manual classification requires the ability to visually/acoustically compare and label large numbers of units quickly. *Koe* offers interactive ordination plots, unit tables and class exemplars as complementary tools for this purpose.

2.3.1 | Interactive ordination plots

A previously unexplored potential of ordination is to expedite the manual classification of units. *Koe*'s interactive ordination plots (Figure 3) incorporate audio playback, spectrograms and classification functionality, so that a user can simultaneously use their audio-visual perception of unit similarity and the structure of the data to rapidly and robustly classify units.

Koe implements three ordination techniques: Principal Component Analysis (PCA; Pearson, 1901), Independent Component Analysis (ICA; Comon, 1994), and t-distributed Stochastic Neighbour Embedding (t-SNE; Van Der Maaten & Hinton, 2008). t-SNE aims to preserve local structure in the data and is particularly effective for defining and discriminating between clusters.

The user encircles groups of points on the plot to see spectrograms and hear playback of units; if a selection appears acoustically consistent, it is classified in bulk. The user can toggle visibility of each class independently and can zoom to examine structural detail.

2.3.2 | Unit tables

As a complement to the ordination plot, units can be viewed as an interactive table (Figure 4). Each unit is represented by a row containing spectrogram, audio and associated information (class label, unit

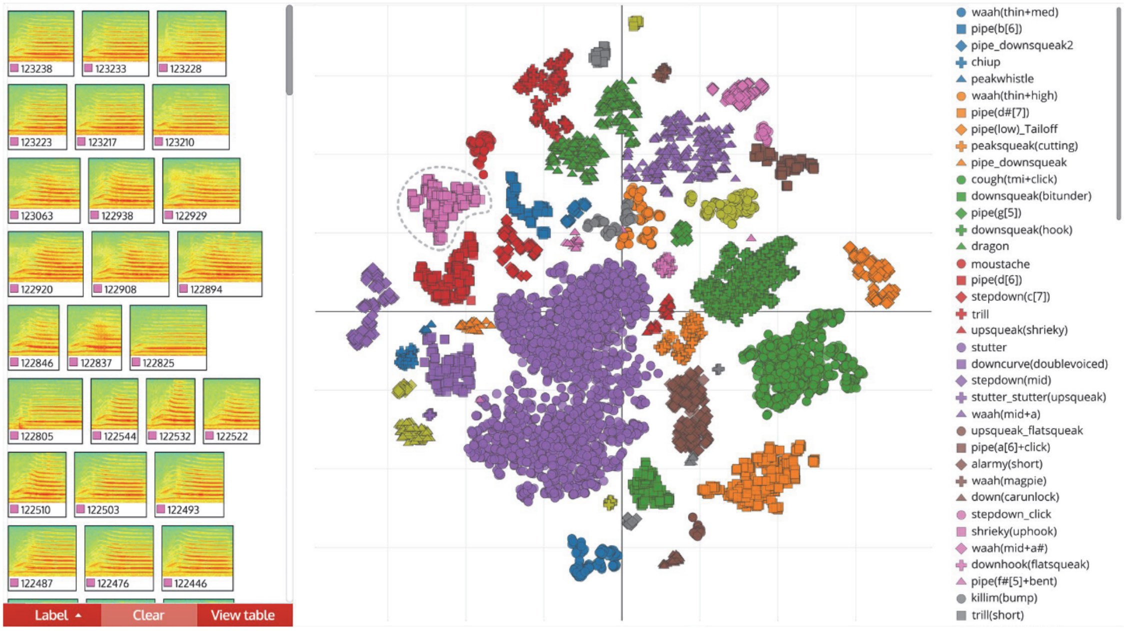
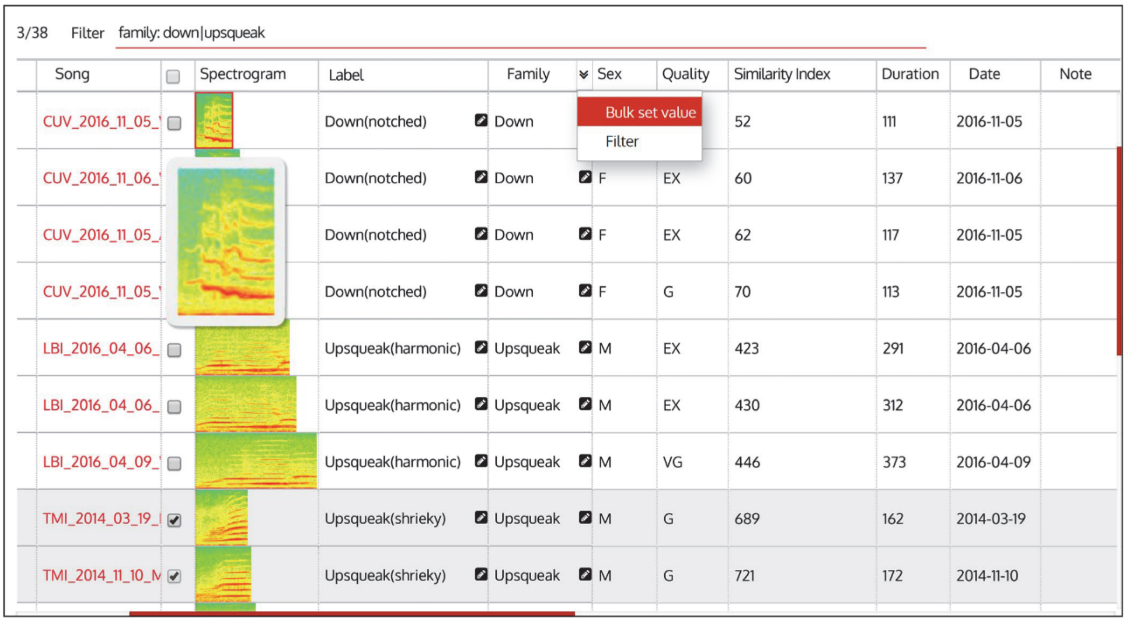


FIGURE 3 *Koe's* interactive *Ordination* view. Interactive ordination view allows the user to encircle groups of points on the plot with the lasso tool, to view their spectrograms and hear their audio. Mousing over a point in a selection highlights the corresponding spectrogram in the left-hand panel. Selections can be labelled in bulk directly on the plot or opened as a unit table to view detailed unit information. The user can zoom, toggle the visibility of classes and export the plot as a vector graphic. This example shows a t-SNE ordination of 7,189 syllables of male and female bellbird song



3/38 Filter family: down|upsqueak

Song	☐	Spectrogram	Label	Family	Sex	Quality	Similarity Index	Duration	Date	Note
CUV_2016_11_05_'	☐		Down(notched)	Down			52	111	2016-11-05	
CUV_2016_11_06_'			Down(notched)	Down	F	EX	60	137	2016-11-06	
CUV_2016_11_05_			Down(notched)	Down	F	EX	62	117	2016-11-05	
CUV_2016_11_05_'			Down(notched)	Down	F	G	70	113	2016-11-05	
LBI_2016_04_06_	☐		Upsqueak(harmonic)	Upsqueak	M	EX	423	291	2016-04-06	
LBI_2016_04_06_	☐		Upsqueak(harmonic)	Upsqueak	M	EX	430	312	2016-04-06	
LBI_2016_04_09_'	☐		Upsqueak(harmonic)	Upsqueak	M	VG	446	373	2016-04-09	
TMI_2014_03_19_	☑		Upsqueak(shrieky)	Upsqueak	M	G	689	162	2014-03-19	
TMI_2014_11_10_M	☑		Upsqueak(shrieky)	Upsqueak	M	G	721	172	2014-11-10	

FIGURE 4 *Unit table* view in *Koe*. *Koe's* unit table is designed for classifying, annotating and filtering units. Each unit row contains a spectrogram which becomes enlarged during mouse-over. Unit audio plays when a spectrogram is clicked. The table can be sorted/filtered by any columns. Sorting by the *similarity index* column arranges units by spectral similarity for expedited labelling

duration, song ID, individual ID, date, etc.). The table can be sorted/filtered by any column. A notable feature is the *similarity index*, which ranks units based on acoustic similarity. The index is produced as follows: from the raw feature measurements or from the ordination, *Koe* calculates pairwise Euclidean distance between each pair of units, then constructs a ladderized dendrogram using agglomerative hierarchical clustering (UPGMA) (Sokal, 1958). The order of the dendrogram leaf nodes becomes the similarity index. Sorting by the *similarity index* column orders the table so that similar units arrange together, allowing them to be selected and labelled in large batches.

2.3.3 | Class exemplars

A class catalogue is a useful reference during classification. *Koe* produces one automatically in *Exemplars* view, with exemplar spectrograms and playback for every class, making visual/acoustic comparison easy (Figure S2). The catalogue updates dynamically as classification progresses, displaying 10 randomly chosen exemplars per class to reflect within-class variability.

2.3.4 | Classification granularity

Without a priori knowledge of the animal's perception of units, a researcher must identify classes and assign units based on their own perception. Classifying units at multiple hierarchical levels of granularity increases robustness by enabling analyses at different scales. *Koe* offers up to three granularity levels (fine, medium, and broad-scale) for labelling in the *Ordination* and *Unit table* views. For example, the units in Figure 4 are labelled at two granularity levels: the broad-scale *Upsqueak* family is subdivided at the fine-scale into *Upsqueak(harmonic)* and *Upsqueak(shrieky)*.

2.3.5 | Validate classification through independent labelling

To ensure robustness of manual classification, a common validation method is to have several judges independently label the dataset and calculate their degree of agreement; high agreement lends credibility to the classification (Nelson, Hallberg, & Soha, 2004; Parker, Anderson, Jenkins, & Brunton, 2012). Typically, 2–5 judges are used, but robustness can improve substantially with more judges (Jones, Bijleveld, & Bijleveld, 2001). *Koe* facilitates independent labelling experiments. A subset of songs/units can be selected and copied to a new database. The database owner grants labelling access to participating judges, who label the dataset online. The labels of participants are automatically saved to the server and are compiled to evaluate concordance (see Case Study below for a real-world example using 74 judges).

2.4 | Analyse sequence structure

For many animals, songs are comprised of acoustic units ordered into a sequence (Kershenbaum et al., 2016). The aim of sequence analysis is to reveal patterns in the sequence structure of songs, which *Koe* can do in three ways, outlined below.

2.4.1 | Filter songs by subsequence

In *View all songs* a user can filter for songs that contain a certain subsequence of unit labels (Figure 5). This could be used to identify all instances of a certain song type, for example.

2.4.2 | Discover patterns using sequence rule mining (cSPADE)

As one way of exploring rules that may govern song structure (i.e. syntax; Lachlan et al., 2013), *Koe* uses the cSPADE (constrained Sequential Pattern Discovery using Equivalence classes) algorithm (Zaki, 2001) to discover commonly occurring sequences in a set of songs.

We consider a sequence to be an ordered list of *x* acoustic units denoted as A → B → C → … → x.

A sequence rule has two parts: a left and right side. The rule states that when the left side occurs, the right side follows: [Left side] ⇒ Right side. For example, the rule [A → B] ⇒ C states that when the sequence [A → B] occurs, C comes next. The left side can be a sequence of any length, but in our implementation the right side is a single unit (see https://github.com/fzyukio/koe/wiki, *Mine Sequence Structure*).

The cSPADE algorithm calculates the credibility of sequence rules. Credible rules have a large confidence factor, a large level of support and a value of lift greater than one, as defined below using the example rule, [A → B] ⇒ C.

Support

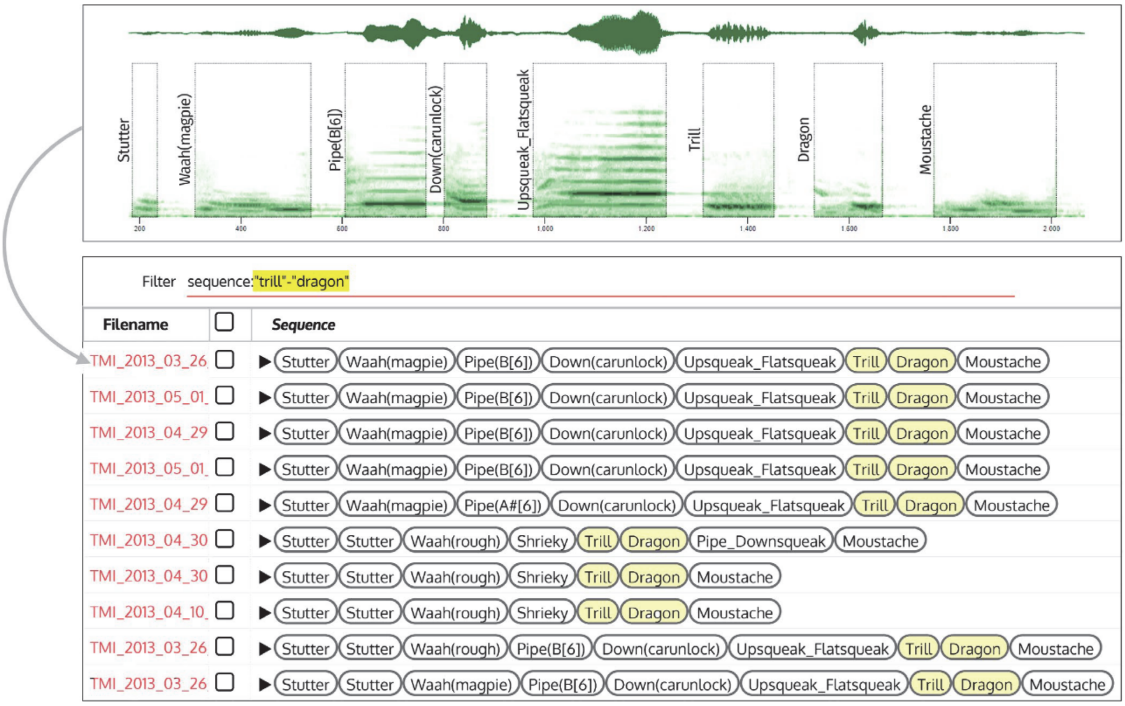
The proportion of songs in the database that contain the entire sequence [A → B → C] at least once.

Confidence

The proportion of those songs containing [A → B] that also contain [A → B → C].

Lift

A measure of the strength of the association relative to chance. Lift is equal to the confidence of the rule, divided by the proportion of songs containing the right side. Thus it gives the ratio of (a) the proportion of songs with [A → B] that transition to C, versus (b) the



Filter sequence:"trill"-"dragon"

Filename	☐	Sequence
TMI_2013_03_26	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_05_01_	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_29	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_05_01_	☐	▶ Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_29	☐	▶ Stutter Waah(magpie) Pipe(A#[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_04_30	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Pipe_Downsqueak Moustache
TMI_2013_04_30	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Moustache
TMI_2013_04_10_	☐	▶ Stutter Stutter Waah(rough) Shrieky Trill Dragon Moustache
TMI_2013_03_26	☐	▶ Stutter Stutter Waah(rough) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache
TMI_2013_03_26	☐	▶ Stutter Stutter Waah(magpie) Pipe(B[6]) Down(carunlock) Upsqueak_Flatsqueak Trill Dragon Moustache

FIGURE 5 *Koe's* songs list view. *View all songs* displays songs in an interactive table, with one song per row. Segmented songs (like the example in the top panel) are represented as a sequence of unit labels. Entire songs can be played, or individual labels clicked for unit playback and spectrograms. Here the filter has returned all songs containing *Trill* followed by *Dragon* (highlighted)

proportion of songs expected to contain [A → B → C] by chance association of [A → B] and C.

To demonstrate cSPADE, consider the ten songs shown in Figure 5 to be a population of songs. The rule [*Stutter* → *Waah(magpie)*] ⇒ *Pipe(B6)* has a **support** of 0.4 since the entire sequence occurs in four of the 10 songs. The rule has a **confidence** of 0.8, because in four of the five songs that contain *Stutter* → *Waah(magpie)*, the transition to *Pipe(B6)* occurs. The proportion of songs with *Pipe(B6)* is 0.6, so the **lift** of this rule is 0.8/0.6 = 1.33. That is, the association [*Stutter* → *Waah(magpie)*] ⇒ *Pipe(B6)* occurs in 1.33 times as many songs as expected by chance association of [*Stutter* → *Waah(magpie)*] and *Pipe(B6).*

2.4.3 | Visualize sequence structure as a network

Two-unit associations from cSPADE can be visualized using a *directed network*. The network models the direction and strength of association between pairs of units, across a population of songs. Units are represented by nodes which are joined by lines (edges) if the units occur consecutively. The order of units is indicated by arrow direction, and strength of association between units (lift) is represented by edge thickness. Visually cluttered networks can be simplified using the filter, e.g. to show only associations with high lift.

Thus, *Koe* offers pattern recognition and sequence visualization. If a user desires to make formal inferences, such as assessing the influence of experimental factors (fixed effects) and random effects of individuals on sequence structure, sophisticated Markovian frameworks exist (e.g. Sarkar, Chabout, Macopson, Jarvis, & Dunson, 2018). A user can export sequence data from *Koe* for external analysis.

3 | CASE STUDY: NEW ZEALAND BELLBIRD SONG

Here we showcase *Koe*'s analysis features with a large (21,500-unit) dataset of New Zealand bellbird *Anthornis melanura* (hereafter 'bellbird') songs, manually classified in *Koe* using the methods described above. The songs were recorded at six sites in an archipelago metapopulation (Figure 6). Bellbird songs are complex and diverse, both in units (i.e. syllables) and sequence structure (Brunton & Li, 2006; Heather & Robertson, 2000). Furthermore, both sexes of bellbird sing prolifically with sexually dimorphic songs and site-specific dialects. This presents a challenging test case for classification and for analysing repertoires and sequence structure. For a step-by-step case study tutorial see the *Koe* documentation (https://github.com/fzyukio/koe/wiki/Analysis-Tutorial%3A-Case-Study-of-NZ-bellbird-song).

3.1 | Validating our classification with independent labelling

We harnessed the citizen science potential of *Koe* to evaluate inter-observer reliability using 74 judges (to our knowledge the largest number of judges yet used). From our labelled dataset of 21,500 bellbird syllables (units), we constructed a subset database of 200 syllables in *Koe*. The 200 syllables consisted of 18 label classes from two populations – Tawharanui and Tiritiri Matangi (Tiri) – with 3–20 of each class, including 4 'other' syllables that did not match the exemplars (the classes and numbers of each class were chosen randomly). The 74 judges (all naïve to *Koe*, spectrograms and bellbird song) created *Koe* accounts and were granted online access to the database. Judges were asked to independently label the syllables using the unit table by comparing against labelled class exemplars. To reduce bias, the judges were not given the exact number of each class, but rather a range that included the true value. Hence they were told there were 3–20 of each class, and 3–20 'other' syllables that did not match the exemplars. For each syllable, the percentage of judges whose labels matched our own label was calculated: average 89.6%; median 95.6%. This is a high degree of agreement for inter-observer reliability studies (Nelson et al., 2004; Parker et al., 2012), lending validity to our manual classification.

3.2 | Analysing syllable diversity

Data can be explored quickly using *Koe*'s filter, which combines different criteria specified by a regular expression. For example, `duration:<100; sex:F; label:Downsqueak` returns only units with duration shorter than 100 ms, vocalized by female birds, and label containing *Downsqueak*. Below, we use the filter to compare male and female unit diversity across the whole metapopulation, then between males at two sites.

3.2.1 | What is the overall unit diversity for male and female bellbirds across all six sites?

Exemplars view displays the catalogue and count of unit classes, revealing 754 fine-scale classes in all. In *Unit table* view we made subset collections by sex. Viewing the subsets in *Exemplars* view revealed 517 fine-scale classes sung by males, 415 sung by females and 178 (24%) sung by both.

3.2.2 | How many male unit types are now shared between Tawharanui and source population Hauturu?

In 2005, bellbirds from Hauturu naturally recolonized Tawharanui Peninsula following pest control. At that point, repertoires of source and founder populations were indistinguishable (Brunton, Evans, & Ji, 2008). After a decade of potential cultural divergence, what is the current degree of male repertoire overlap between the sites?

In *Unit table* view we made subsets containing units sung by males from Tawharanui (`sex:M; song:TAW`) and Hauturu (`sex:M;`

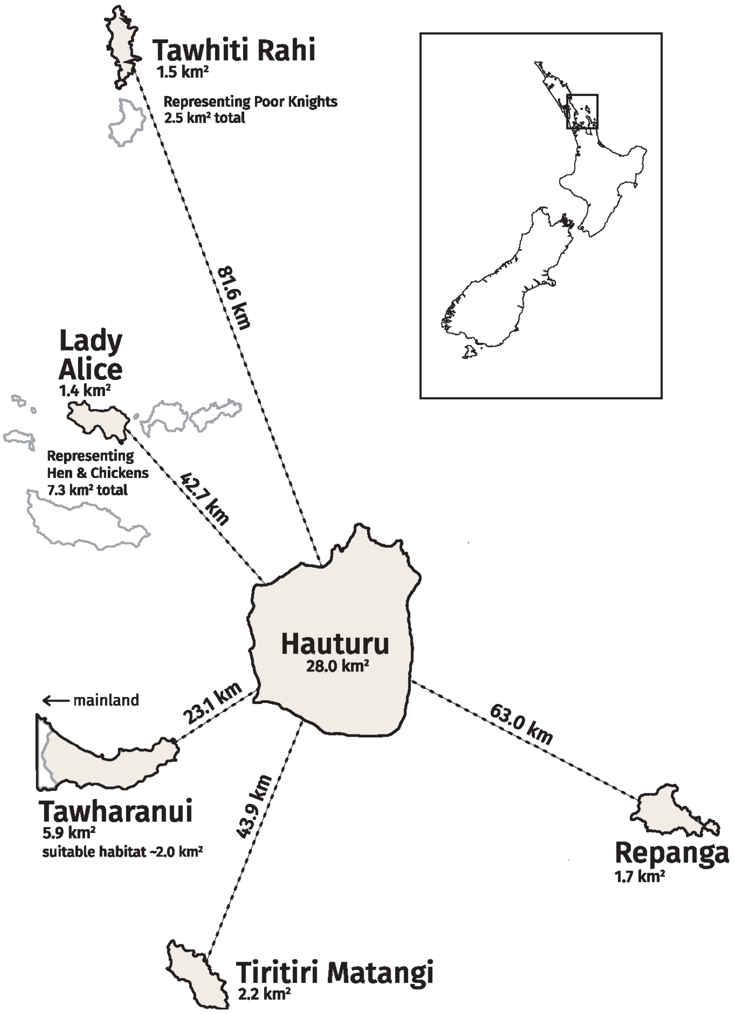


FIGURE 6 The six sites where NZ bellbird *Anthornis melanura* populations were recorded in the Hauraki Gulf, New Zealand. Between 2013 and 2017, songs were recorded from each site using portable recorders and shotgun microphones. For simplicity, only distances from Hauturu are shown (distances not to scale)

`song:LBI`). These subsets showed 141 male fine-scale unit classes on Hauturu, and 104 at Tawharanui. By examining the intersect of the two class lists in *R Studio* (RStudio Team, 2015), we found 35 Tawharanui unit classes shared with Hauturu (34%) – since colonization 66% of the unit classes sung by males at Tawharanui appear to have been innovated.

3.3 | Evaluating song structure in bellbirds

To analyse sequence structure in detail, *Mine sequence structure* shows the strength of association between unit classes using the cSPADE algorithm (Zaki, 2001), and produces network visualizations. This can be used to visually compare sequence structure between any data subsets, e.g. individuals, sexes, sites or seasons. Here we compare the sequence structure of male and female bellbird song on Tiri.

3.3.1 | Male and female song structure as revealed by cSPADE

One-unit 'associations' show prevalence of syllable classes in the population. The three most prevalent classes for males are *Dragon*, *Cough(TMI+click)* and *Downsqueak(hook)*, occurring in 52%, 40% and 35% of male songs, respectively (Table 1). The three most prevalent female classes are *Stutter*, *Chiup* and *Downcurve(doublevoiced)*, occurring in 93%, 59% and 36% of female songs, respectively.

The most prevalent male two-unit association, *Cough(TMI+click) ⇒ Downsqueak(hook)*, occurs in 32% of songs. It has a confidence of 0.81; i.e. when *Cough(TMI+click)* occurs, *Downsqueak(hook)* comes next in 81% of songs. It has a lift of 2.33; i.e. *Downsqueak(hook)* follows *Cough(TMI+click)* in 2.33 times the number of songs as expected by chance. Male song most often starts with *Stutter* (29% of songs)

TABLE 1 Key features of male and female song structure as extracted by the cSPADE algorithm. Results for NZ bellbird males (389 songs comprised of 5,305 syllables) and females (138 songs comprised of 1,884 syllables) recorded on Tiri between February 2013 and December 2015. Rule length is how many units constitute the association. Song count is how many songs the association occurs in. See text for definitions of support, confidence and lift

Rule length	Association rule	Song count	Support	Confidence	Lift
Male population					
1	Dragon	201	0.52	---	---
1	Cough(TMI+click)	156	0.40	---	---
1	Downsqueak(hook)	135	0.35	---	---
…	…	…	…	…	…
2	Cough(TMI+click) ⇒ Downsqueak(hook)	126	0.32	0.81	2.33
2	START ⇒ Stutter	113	0.29	0.29	0.95
2	START ⇒ Cough(TMI+click)	90	0.23	0.23	0.57
2	Dragon ⇒ END	115	0.30	0.57	0.57
2	Moustache ⇒ END	46	0.12	0.81	0.81
2	Cough(TMI+click) ⇒ END	45	0.12	0.29	0.29
…	…	…	…	…	…
3	[Trill → Dragon] ⇒ Moustache	32	0.08	0.86	5.9
…	…	…	…	…	…
Female population					
1	Stutter	129	0.93	---	---
1	Chiup	82	0.59	---	---
1	Downcurve(doublevoiced)	50	0.36	---	---
…	…	…	…	…	…
2	START ⇒ Stutter	125	0.91	0.91	0.97
2	Stutter ⇒ Stutter	117	0.85	0.91	0.97
2	Stutter ⇒ END	52	0.38	0.40	0.40
2	Downcurve(doublevoiced) ⇒ Chiup	49	0.36	0.98	1.65
…	…	…	…	…	…
9	[Stutter → Stutter → Stutter → Stutter →Stutter → Stutter → Stutter → Stutter] ⇒ Stutter	21	0.15	0.75	0.80
…	…	…	…	…	…
11	[Stutter → Stutter → Stutter → Stutter →Stutter → Stutter → Stutter → Stutter → Stutter → Stutter] ⇒ Stutter	10	0.07	0.77	0.82

or *Cough(TMI+click)* (23%). It most often ends with *Dragon* (30%), *Moustache* (12%) or *Cough(TMI+click)* (12%). By contrast, for females, the importance of *Stutter* is apparent. Not only does *Stutter* start in 91% of female songs, it also repeats (i.e. *Stutter* ⇒ *Stutter*) in 85% and ends in 38%. Furthermore, *Stutter* is commonly repeated many (≤11) times.

3.3.2 | Contrasting male and female song structure with network diagrams

Sexual dimorphism in two-unit sequences was examined at the population level with network diagrams with a threshold of support >0.05 (Figure 7). The songs of both sexes contain some nodes with many connections; these unit classes (e.g. *Dragon* in males, *Stutter* in females) can be preceded/followed by many different classes. Both sexes also contain nodes which are always preceded/followed by a certain class; for example, in males *Waah(mid+A)* is always followed by *Dragon*, and in females *Peakwhistle(TMI)* is always preceded by *Stutter*. Both sexes sing repeated units, represented by looped lines. For males: *Dragon*, *Stutter*, and *Cough(TMI+click)*. For females: *Stutter*, *Stutter(wavey)*, *Pipe(D#[7]+)*, and *Pipe(E[7])*. A notable difference in the female network (Figure 7b) is the chain of strongly associated nodes: *Killem* → *Pipe(C#[7]+)* → *Upsqueak_Flatsqueak(1)* → *Downsqueak(twopart)* → *Down_Click_Stepdown* → *Stepdown(C[7]-C[6])*.

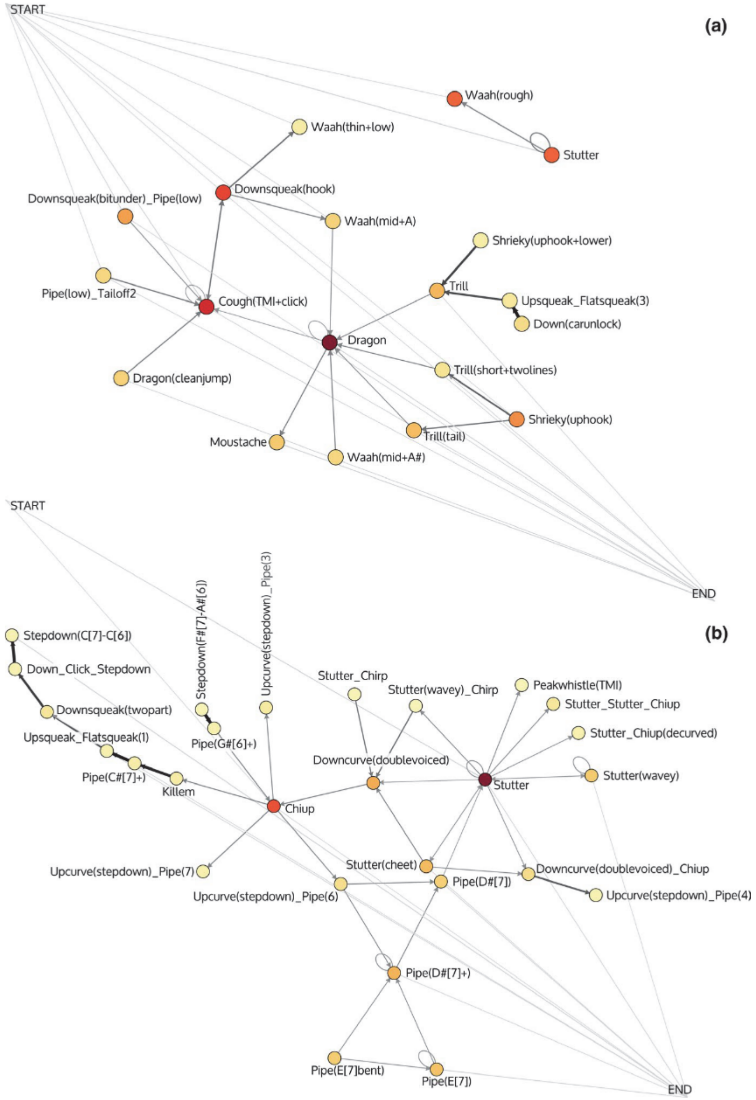


FIGURE 7 Network visualizations of male and female song sequence structure. The networks show two-unit associations for (A) male and (B) female NZ bellbird song on Tiri. The networks are produced from the cSPADE data in Table 1. Key features of male and female song structure as extracted by the cSPADE algorithm. Each node represents a unit class. Colours of the nodes represent frequency of occurrence, with darker, redder nodes being more frequently recorded classes. The arrows show directions of the association. Thickness and darkness of the arrows represent the strength of the association. The networks have been filtered to show only those associations that occur in >5% of songs. Classes that start or end in at least one instance are tethered to the START/END nodes, respectively, with thin grey lines

4 | CONCLUSION

Koe is an end-to-end solution for large-scale, high-resolution classification and analysis of animal vocalizations. It features tools for segmenting, comparing and classifying acoustic units and analysing sequence structure. Interactive ordination and unit tables provide a major advance in classification speed and robustness over existing methods. We have demonstrated the power of *Koe* for large-scale analyses with a case study of 21,500 syllables of bellbird song, classified in *Koe* into over 700 classes and analysed in terms of repertoires and sequence structure. *Koe* was designed primarily for taxa/questions requiring acoustic classification; however, *Koe* is not limited to such cases. Where classification is not the aim, e.g. in animals where vocalizations show graded (non-discrete) variation, *Koe* is still a powerful tool for efficient extraction of measurements, which can be exported for analysis. Being web-based and accessible from any device, *Koe* is ideal for collaboration, teaching and citizen science.

ACKNOWLEDGEMENTS

This manuscript was substantially improved by comments from two anonymous reviewers. We thank Christine Evans and Aaron Harmer for helpful comments during *Koe* development. We thank Ngati Wai and Ngati Manuhiri iwi, the New Zealand Department of Conservation (47948-FAU, 34833–FAU, 41756-FAU, 48000-FAU) and Massey University Animal Ethics Committee (15/21) for permission to record bellbird song at our six sites. We thank Barbara Evans and Aaron Harmer for assistance with song recording. This research was funded by a Massey University Doctoral Scholarship and Royal Society of New Zealand Marsden Fund grants (13-MAU-004, 17-MAU-154, 17-UOA-295).

AUTHORS' CONTRIBUTIONS

Y.F. and W.H.W. conceived of the software and developed the project with guidance from the other authors; Y.F. coded the software; A.G. arranged the hosting server. M.D.M.P., A.G. S.M. gave valuable insights into the algorithms *Koe* utilizes. W.H.W. led the writing of the manuscript and user manual with contributions from the other authors; W.H.W., M.M.R. and D.H.B. contributed song data for the development of *Koe* and for the case study. All authors gave final approval for submission.

DATA AVAILABILITY STATEMENT

Koe source code is available on GitHub (http://github.com/fzyukio/koe) and is archived on Zenodo (https://doi.org/10.5281/zenodo.3538887) (Fukuzawa, agnz, & jren2019, 2019). *Koe* users have access to the *Bellbirds_Case_Study* database analysed in this paper; instructions are available in the user manual at koe.io.ac.nz., which contains a step-by-step tutorial to reproduce our analyses. Label data for reproducing our analyses, and a set of bellbird song recordings, are also available on Dryad Digital Repository (https://doi.org/10.5061/dryad.h44j0zpfg).

ORCID

Yukio Fukuzawa https://orcid.org/0000-0002-5069-8203
Wesley H. Webb https://orcid.org/0000-0002-6676-6194
Matthew D.M. Pawley https://orcid.org/0000-0003-4502-3989
Michelle M. Roper https://orcid.org/0000-0002-4878-2203
Stephen Marsland https://orcid.org/0000-0002-9568-848X
Dianne H. Brunton https://orcid.org/0000-0002-1741-1960
Andrew Gilman https://orcid.org/0000-0002-0050-1927

REFERENCES

Brunton, D. H., Evans, B. A., & Ji, W. (2008). Assessing natural dispersal of New Zealand bellbirds using song type and song playbacks. *New Zealand Journal of Ecology*, *32*, 147–154.

Brunton, D. H., & Li, X. (2006). The song structure and seasonal patterns of vocal behavior of male and female bellbirds (*Anthornis melanura*). *Journal of Ethology*, *24*, 17–25. https://doi.org/10.1007/s10164-005-0155-5

Catchpole, C. K., & Slater, P. J. B. (2008). *Bird song: Biological themes and variations* (2nd ed.). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511754791

Comon, P. (1994). Independent component analysis: A new concept? *Signal Processing*, *36*, 287–314. https://doi.org/10.1016/0165-1684(94)90029-9

Django Software Foundation. (2013). Django: The web framework for perfectionists with deadlines. Djangoproject.Com.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. New York, NY: John Wiley & Sons.

Fukuzawa, Y., agnz, & jren2019. (2019). fzyukio/koe 5.0.0 (Version 5.0.0). Zenodo. http://doi.org/10.5281/zenodo.3538887

Heather, B. D., & Robertson, H. A. (2000). *The field guide to the birds of New Zealand*. Auckland, New Zealand: Viking/Penguin Books.

Janik, V. M., & Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *Journal of Comparative Physiology A*, *199*, 479–489. https://doi.org/10.1007/s00359-013-0817-7

Jones, A., Bijleveld, C. C. J. H., & Bijleveld, C. C. J. H. (2001). The interobserver reliability of scoring sonagrams by eye: A study on methods, illustrated on zebra finch songs. *Animal Behaviour*, *62*, 791–801. https://doi.org/10.1006/anbe.2001.1810

Jones, E., Oliphant, T., & Peterson, P. (2001). SciPy: Open source scientific tools for Python. Retrieved from http://www.scipy.org/

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., … Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews*, *91*, 13–52. https://doi.org/10.1111/brv.12160

Lachlan, R. F., Verzijden, M. N., Bernard, C. S., Jonker, P.-P., Koese, B., Jaarsma, S., … ten Cate, C. (2013). The progressive loss of syntactical structure in bird song along an island colonization chain. *Current Biology*, *23*, 1896–1901. https://doi.org/10.1016/j.cub.2013.07.057

Nelson, D. A., Hallberg, K. I., & Soha, J. A. (2004). Cultural evolution of Puget sound white-crowned sparrow song dialects. *Ethology*, *110*, 879–908. https://doi.org/10.1111/j.1439-0310.2004.01025.x

Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science and Engineering*, *9*(3), 10–20. https://doi.org/10.1109/MCSE.2007.58

Parker, K. A., Anderson, M. J., Jenkins, P. F., & Brunton, D. H. (2012). The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecology Letters*, *15*, 778–785. https://doi.org/10.1111/j.1461-0248.2012.01797.x

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*, 559–572. https://doi.org/10.1080/14786440109462720

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., …, Duchesnay, É. (2012). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peters, S. S., Searcy, W. A., & Marler, P. (1980). Species song discrimination in choice experiments with territorial male swamp and song sparrows. *Animal Behaviour*, *28*, 393–404. https://doi.org/10.1016/S0003-3472(80)80048-0

Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: A review. *Journal of Avian Biology*, *49*, jav-01447. https://doi.org/10.1111/jav.01447

Rossum, G. (1995). *Python tutorial*. Amsterdam, The Netherlands: CWI (Centre for Mathematics and Computer Science).

RStudio Team. (2015). *RStudio: Integrated development environment for R*. Boston, MA: RStudio Inc.

Sarkar, A., Chabout, J., Macopson, J. J., Jarvis, E. D., & Dunson, D. B. (2018). Bayesian semiparametric mixed effects markov models with application to vocalization syntax. *Journal of the American Statistical Association*, *113*, 1515–1527. https://doi.org/10.1080/01621459.2018.1423986

Sayigh, L. S., Esch, H. C., Wells, R. S., & Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour*, *74*, 1631–1642. https://doi.org/10.1016/j.anbehav.2007.02.018

Sokal, R. R. (1958). A statistical method for evaluating systematic relationship. *University of Kansas Science Bulletin*, *28*, 1409–1438.

Soltis, J. (2009). Vocal production, affect expression, and communicative function in African elephant rumble vocalizations. *The Journal of the Acoustical Society of America*, *126*, 2210–2210. https://doi.org/10.1121/1.3248716

Sperber, A. L., Werner, L. M., Kappeler, P. M., & Fichtel, C. (2017). Grunt to go – Vocal coordination of group movements in redfronted lemurs. *Ethology*, *123*, 894–905. https://doi.org/10.1111/eth.12663

Starnberger, I., Preininger, D., & Hödl, W. (2014). From uni- to multimodality: Towards an integrative view on anuran communication. *Journal of Comparative Physiology A*, *200*, 777–787. https://doi.org/10.1007/s00359-014-0923-1

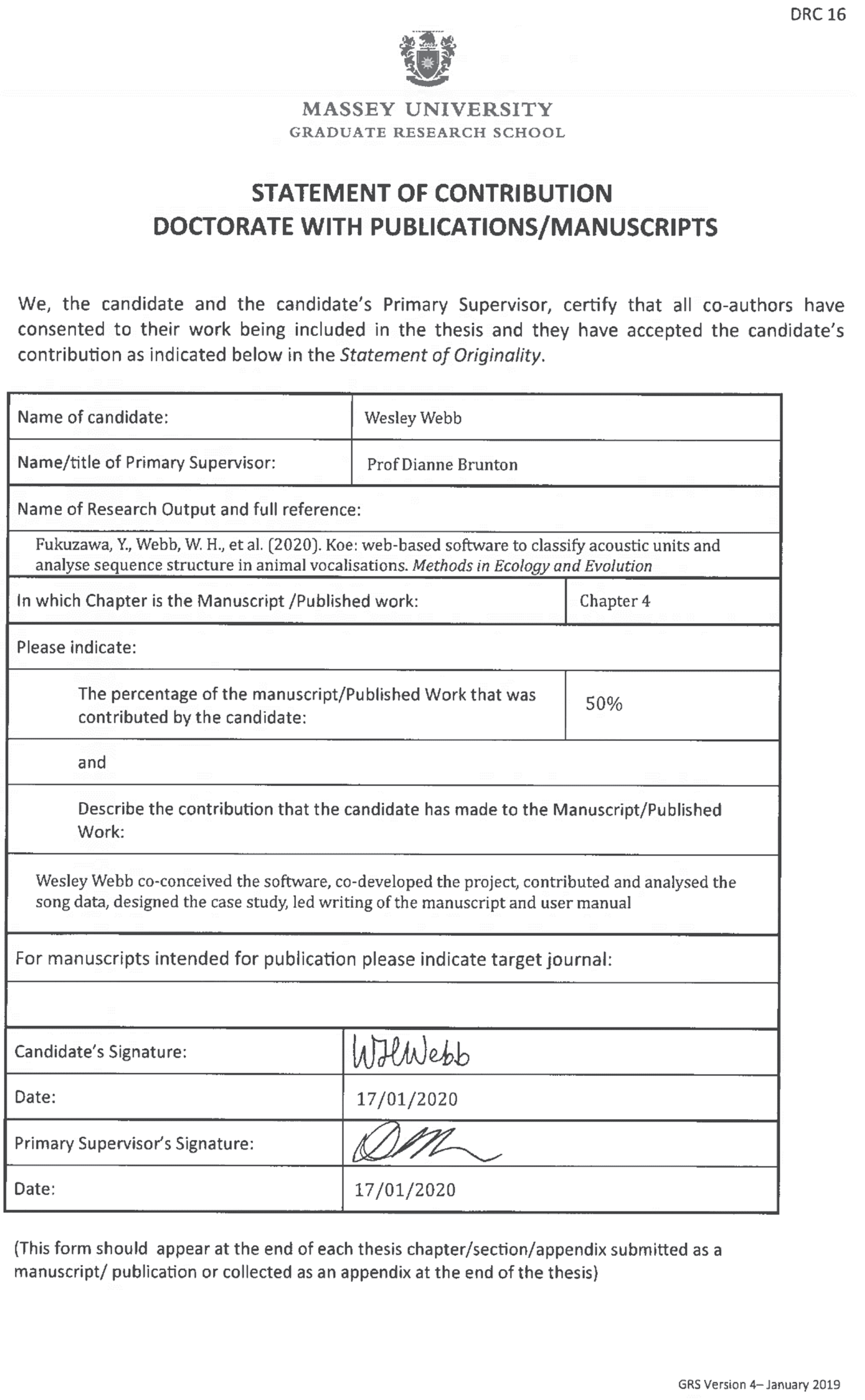
Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605.

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Machine Learning*, *42*, 31–60.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Fukuzawa Y, Webb WH, Pawley MDM, et al. *Koe:* Web-based software to classify acoustic units and analyse sequence structure in animal vocalizations. *Methods Ecol Evol*. 2020;00:1–11. https://doi.org/10.1111/2041-210X.13336



DRC 16

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Wesley Webb
Name/title of Primary Supervisor:	Prof Dianne Brunton
Name of Research Output and full reference:	
Fukuzawa, Y., Webb, W. H., et al. (2020). Koe: web-based software to classify acoustic units and analyse sequence structure in animal vocalisations. *Methods in Ecology and Evolution*	
In which Chapter is the Manuscript /Published work:	Chapter 4
Please indicate:	
The percentage of the manuscript/Published Work that was contributed by the candidate:	50%
and	
Describe the contribution that the candidate has made to the Manuscript/Published Work:	
Wesley Webb co-conceived the software, co-developed the project, contributed and analysed the song data, designed the case study, led writing of the manuscript and user manual	
For manuscripts intended for publication please indicate target journal:	
Candidate's Signature:	WHWebb
Date:	17/01/2020
Primary Supervisor's Signature:	
Date:	17/01/2020

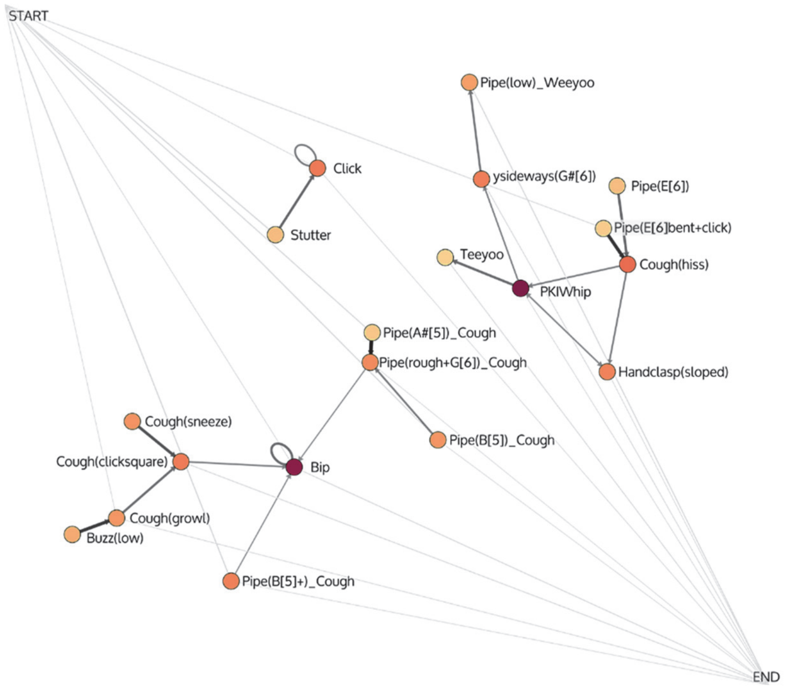
(This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis)

GRS Version 4– January 2019

7.2.2 Network visualisations of male and female NZ bellbird song sequence structure

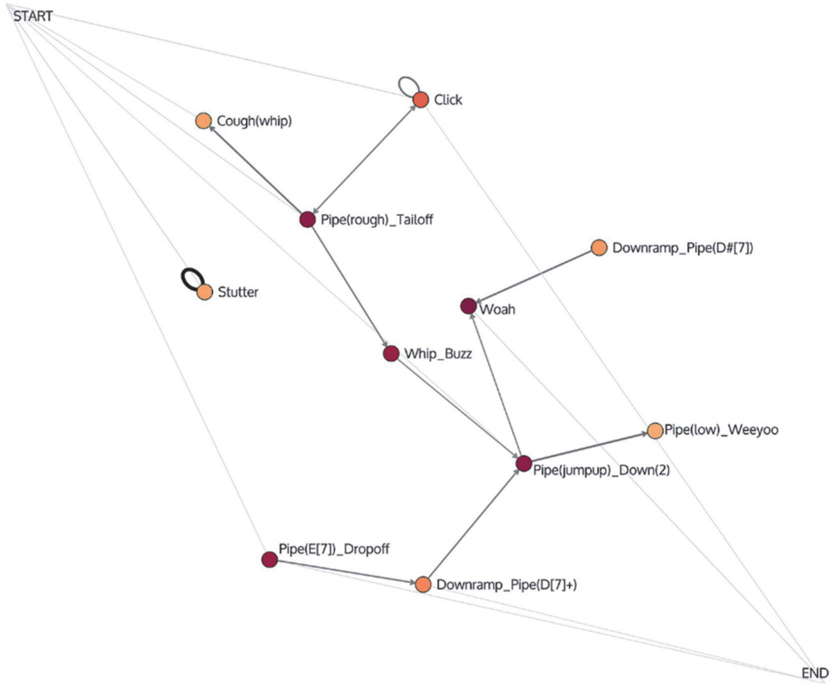
A complete set of network visualisations for all six sites in the Hauraki Gulf archipelago is provided at this link: http://tiny.cc/SyntaxNetworks. An example set of four networks is shown in

Figure 7.1, followed by a comparison of network statistics (Table 7.1) between males and females using nMDS (Figure 7.2).



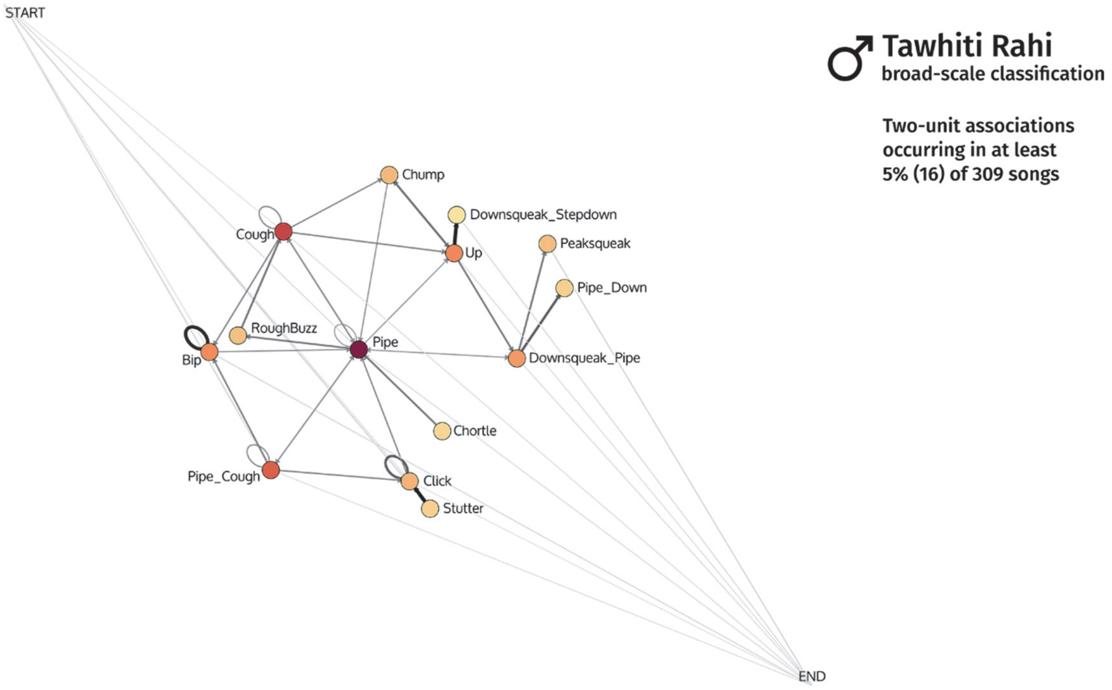
♂ Tawhiti Rahi
fine-scale classification

Two-unit associations occurring in at least 5% (16) of 309 songs



♀ Tawhiti Rahi
fine-scale classification

Two-unit associations occurring in at least 5 of 61 songs


♂ Tawhiti Rahi
broad-scale classification
Two-unit associations
occurring in at least
5% (16) of 309 songs
START
Chump
Downsqueak_Stepdown
Cough
Up
Peaksqueak
Pipe_Down
RoughBuzz
Pipe
Bip
Downsqueak_Pipe
Chortle
Pipe_Cough
Click
Stutter
END

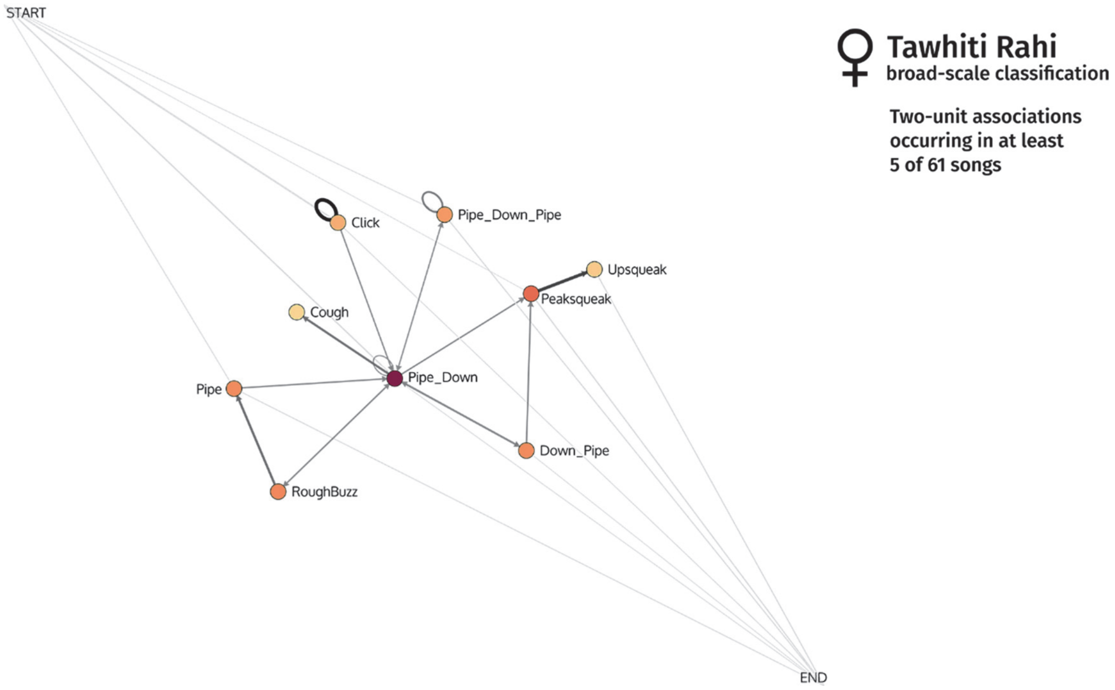

♀ Tawhiti Rahi
broad-scale classification
Two-unit associations
occurring in at least
5 of 61 songs
START
Click
Pipe_Down_Pipe
Upsqueak
Peaksqueak
Cough
Pipe
Pipe_Down
Down_Pipe
RoughBuzz
END

Figure 7.1—Network visualisations of male and female song sequence structure.

(***Above***) These example networks show two-unit associations for male and female NZ bellbird song on Tawhiti Rahi Island, at fine-scale and broad-scale classification levels. Each node represents a unit class. Colours of the nodes represent frequency of occurrence, with darker, redder nodes being more frequently recorded classes. The arrows show directions of the association. Thickness and darkness of the arrows represent the strength of the association. The networks have been filtered to show only those associations that occur in ≥5% of songs and ≥5 songs (i.e. whichever is bigger). Classes that start or end in at least one instance are tethered to the START/END nodes, respectively, with thin grey lines.

7.2.3 Comparison of network features

Network-level statistics, manually derived from the complete set of networks at http://tiny.cc/SyntaxNetworks are presented in Table 7.1. A non-metric MDS on this data (Figure 7.2) found little separation between male and female networks at fine-scale classification, indicating that these statistics do not capture sex differences in syntax. Separation was clearer at broad-scale classification.

Table 7.1—Network-level statistics for all populations.

Columns are explained as follows. Classes: number of different classes in the network. Selfing classes: number of classes that loop back on themselves. Max 'in' connections: the maximum number of classes directly preceding a node. Max 'out' connections: the maximum number of classes coming directly after a node. Max connections (either direction): the maximum number of nodes coming directly before or after a node. Classes that only occur at start: number of classes that are only found at the start of songs, at this support threshold. Classes that only occur at end: number of classes that are only found at the end of songs, at this support threshold.

Network name (site sex scale)	Classes	Selfing classes	Max 'in' connections	Max 'out' connections	Max connections (either direction)	Classes that only occur at start	Classes that only occur at end
PKI M Fine	19	2	3	2	4	5	3
PKI M Broad	14	5	7	5	10	1	2
PKI F Fine	11	2	2	2	4	1	1
PKI F Broad	9	3	5	5	9	0	1
LAI M Fine	8	1	2	2	3	2	2
LAI M Broad	7	2	3	4	5	1	2
LAI F Fine	12	1	3	5	8	1	4
LAI F Broad	4	2	3	3	4	0	1
LBI M Fine	9	3	3	3	5	0	0
LBI M Broad	11	5	5	5	11	0	3
LBI F Fine	25	1	3	5	7	2	1
LBI F Broad	14	3	9	5	14	0	2
REP M Fine	21	3	5	5	10	2	0
REP M Broad	8	5	8	5	12	0	0

REP F Fine	16	3	5	6	10	0	4
REP F Broad	6	5	4	4	7	0	0
TAW M Fine	16	1	3	4	6	1	3
TAW M Broad	15	5	4	5	8	1	1
TAW F Fine	6	1	2	2	3	0	0
TAW F Broad	7	1	3	3	5	0	0
TMI 13 M Fine	18	3	5	3	6	0	1
TMI 13 M Broad	13	6	5	5	9	1	1
TMI 13 F Fine	10	2	3	4	5	0	0
TMI 13 F Broad	8	2	5	4	6	0	0
TMI 14 M Fine	19	3	6	4	8	1	1
TMI 14 M Broad	13	5	5	6	8	0	0
TMI 14 F Fine	17	1	2	3	5	0	3
TMI 14 F Broad	8	3	4	3	6	0	0
TMI 15 M Fine	25	4	9	3	11	1	1
TMI 15 M Broad	16	5	8	5	13	0	1
TMI 15 F Fine	16	4	3	5	7	0	2
TMI 15 F Broad	7	2	3	3	5	0	0
TMI 13-14-15 M Fine	19	3	6	3	8	3	1
TMI 13-14-15 M Broad	12	6	7	5	13	1	0
TMI, 13-14-15 F Fine	28	4	4	9	11	0	1
TMI, 13-14-15 F Broad	10	3	6	5	8	0	0

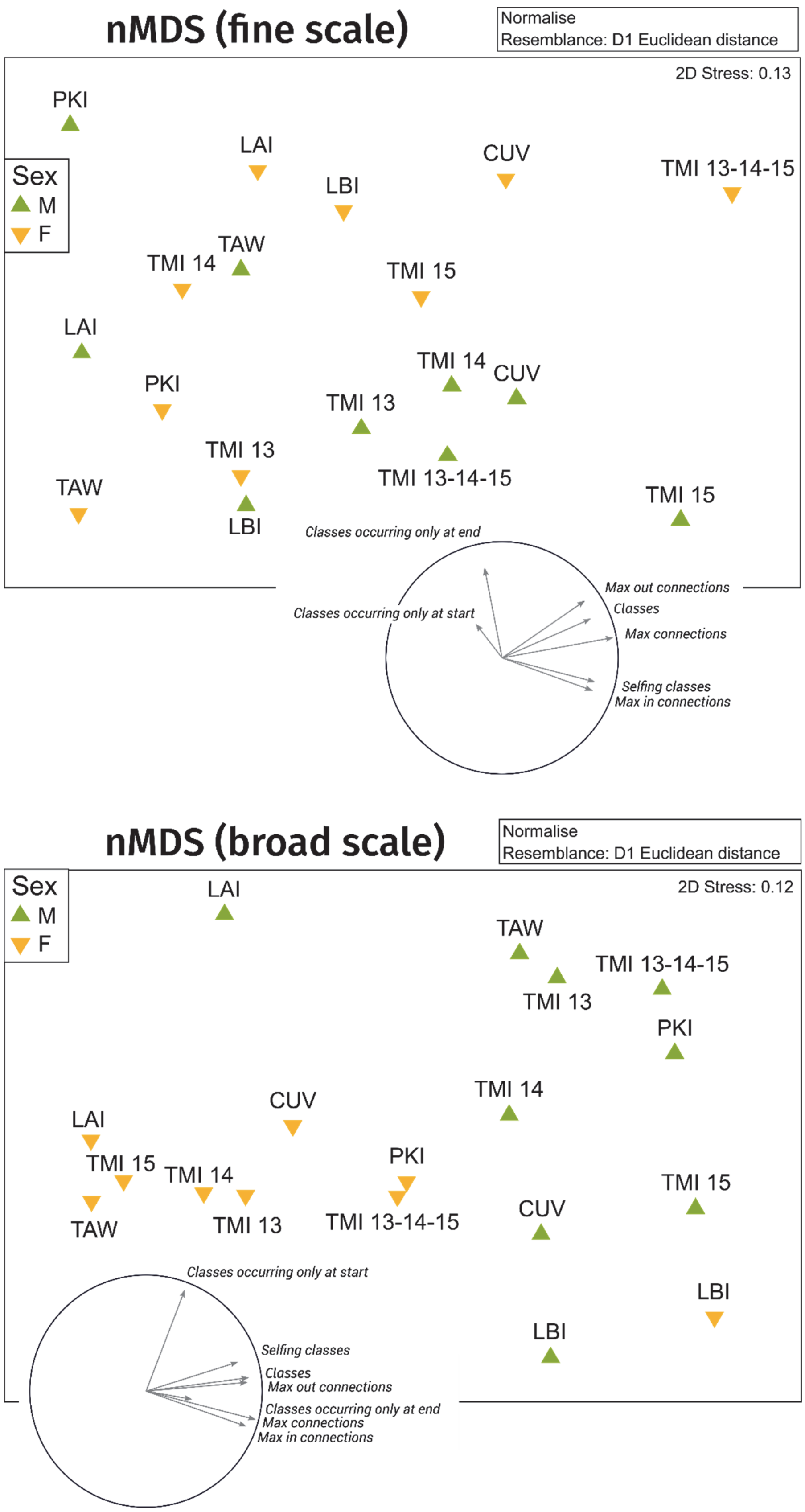


Figure 7.2—Non-metric MDS of all population networks.

Based on network statistics given in Table 7.1

7.3 Appendix for Chapter 5—Cultural evolution of male and female song diversity in an archipelago metapopulation

7.3.1 Complete catalogue of bellbird syllable classes

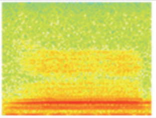
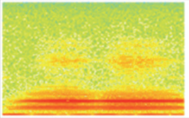
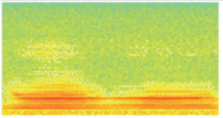
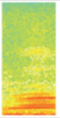
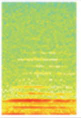
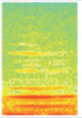
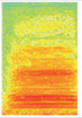
The interactive database is freely accessible on *Koe* (koe.io.ac.nz). Log in with username **Dr_Webb** and password **webb_2019**. Explore the data using the program views accessed by the panel on the left of the screen; *Classify and manage units* views and *View all songs*. For help, see the *Koe* user manual at koe.io.ac.nz.

A PDF catalogue of all observed syllable classes and their counts from across the Hauraki Gulf archipelago metapopulation is provided at this link: http://tiny.cc/BellbirdsCatalogue. Starting pages of the PDF catalogue are shown in Table 7.2 as a sample.

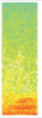
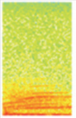
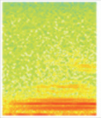
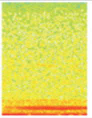
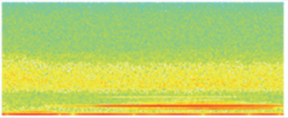
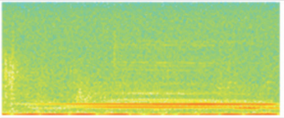
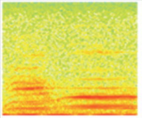
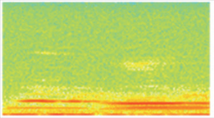
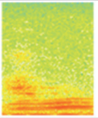
Table 7.2—Catalogue of all syllable classes from across the Hauraki Gulf metapopulation.

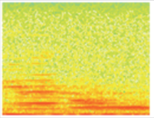
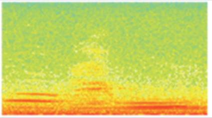
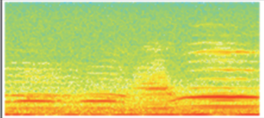
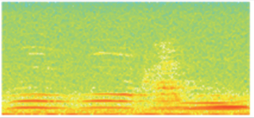
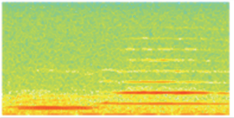
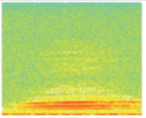
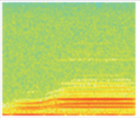
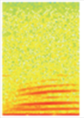
(***Below***) Each row of the table represents one fine-scale syllable class. The broad-scale 'family' to which the class belongs is given in the leftmost column. Classes with an asterisk (*) indicate classes that are sung by both sexes. Counts of the fine-scale class at each site are given in the right-hand columns. For robustness, the fine-scale classes displayed have at least three occurrences within the sex. Site abbreviations are as follows: HAU, Hauturu; TAW, Tawharanui; TMI, Tiri; LAI, Lady Alice; REP, Repanga; PKI, Poor Knights. The spectrogram associated with each class in the table was hand-chosen in *Koe* based on clarity and class representativeness. Time and frequency scalebars are provided at the end of the male and female sub-tables.

Catalogue of male syllable classes

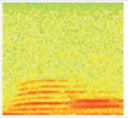
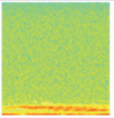
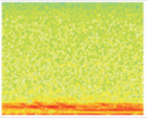
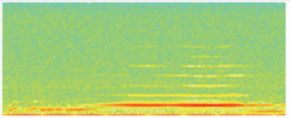
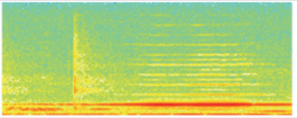
Broad-scale syllable class	Fine-scale syllable class	Spectrogram	HAU	TAW	TMI	LAI	REP	PKI	Total
Alarmy*	Alarmy(cheese2)		3	0	0	0	0	0	**3**
Alarmy*	Alarmy(double)		0	31	0	0	0	0	**31**
Alarmy*	Alarmy(double+high)		4	0	0	0	0	0	**4**
Alarmy*	Alarmy(downflickstart)		0	0	0	11	0	0	**11**
Alarmy*	Alarmy(short)		0	0	286	0	0	1	**287**
Alarmy*	Alarmy(short+highbit)*		0	0	0	0	12	0	**12**
Alarmy*	Alarmy(short+highbit1)		0	0	0	0	10	0	**10**
Alarmy*	Alarmy(short+sharp)		0	0	5	0	0	0	**5**

Sample only. Complete catalogue at http://tiny.cc/BellbirdsCatalogue.

Broad-scale syllable class	Fine-scale syllable class	Spectrogram	HAU	TAW	TMI	LAI	REP	PKI	Total
Alarmy*	Alarmy(short+up)*		0	56	0	0	338	0	**394**
Alarmy*	Alarmy(short+up+hah)		0	25	0	0	0	0	**25**
Alarmy*	Alarmy(trainwhistle+aah)		0	0	0	0	45	0	**45**
Alarmy*	Alarmy(trainwhistle+ooh)		0	23	0	0	15	0	**38**
Alarmy*	Alarmy(tritone1)		0	0	0	0	0	5	**5**
Alarmy*	Alarmy(tritone2)		0	0	0	0	0	3	**3**
Alarmy*	Alarmy(twobitCUV+beware)*		0	0	0	0	38	0	**38**
Alarmy*	Alarmy(twobitCUV+hello)		0	0	0	0	32	0	**32**
Alarmy*	Alarmy(twobitCUV+io)		0	0	0	0	63	0	**63**

Broad-scale syllable class	Fine-scale syllable class	Spectrogram	HAU	TAW	TMI	LAI	REP	PKI	Total
Alarmy*	Alarmy(twobitCUV+nihao)		0	0	0	0	21	0	**21**
Alarmy*	Alarmy(twobitCUV+rightio)		0	0	0	0	29	0	**29**
Alarmy*	Alarmy(twobitCUV+stutter+beware)		0	0	0	0	32	0	**32**
Alarmy*	Alarmy(twobitCUV+stutter+rightio)		0	0	0	0	123	0	**123**
Alarmy*	Alarmy_Pipe(high)		0	0	0	0	26	0	**26**
Alarmy*	SimUpDown_Upsqueak(firsthalf)		0	0	0	0	0	12	**12**
Alarmy*	Stutter(alarmy)*		0	0	0	0	3	0	**3**
Alarmy*	Upsqueak(alarmyLAI)		0	0	0	3	0	0	**3**
Alarmy*	Upsqueak(kitten)*		0	88	0	0	0	0	**88**

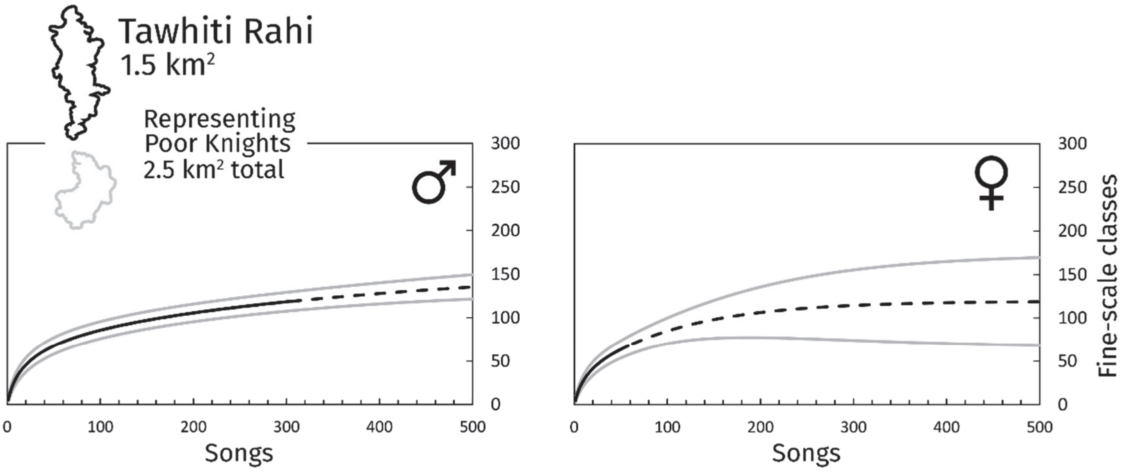
Sample only. Complete catalogue at http://tiny.cc/BellbirdsCatalogue.

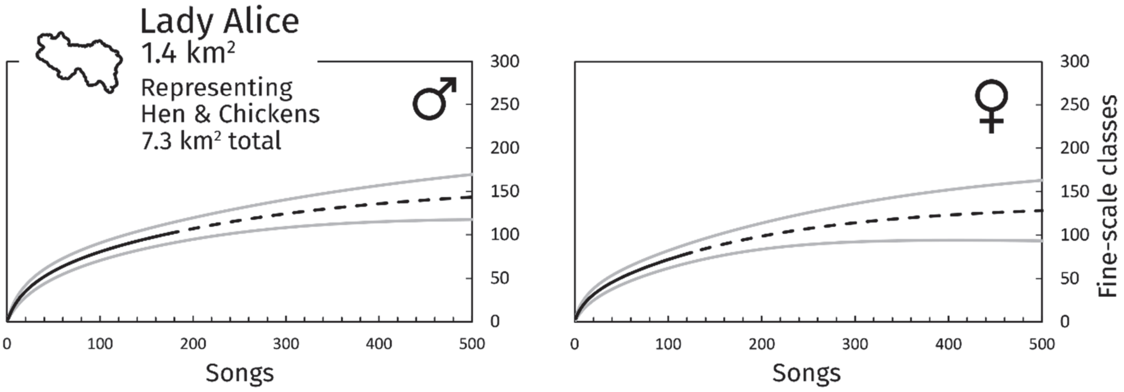
Broad-scale syllable class	Fine-scale syllable class	Spectrogram	HAU	TAW	TMI	LAI	REP	PKI	Total
Alarmy*	Upsqueak(kitten)_Alarmy		0	24	0	0	0	0	**24**
Bip	Bip		0	0	0	0	0	626	**626**
Chiggle*	Chiggle*		0	116	0	0	7	6	**129**
Chiggle*	Chiggle(straight)		0	0	0	5	0	0	**5**
Chortle*	Chortle(click)		0	0	0	0	0	17	**17**
Chortle*	Chortle(LAI)		0	0	0	4	0	0	**4**
Chortle*	Chortle(TAW)		0	10	0	0	0	0	**10**
Chortle*	Chortle(whipsong1)		0	0	0	0	0	4	**4**
Chortle*	Chortle(whipsong1+click)		0	0	0	0	0	6	**6**

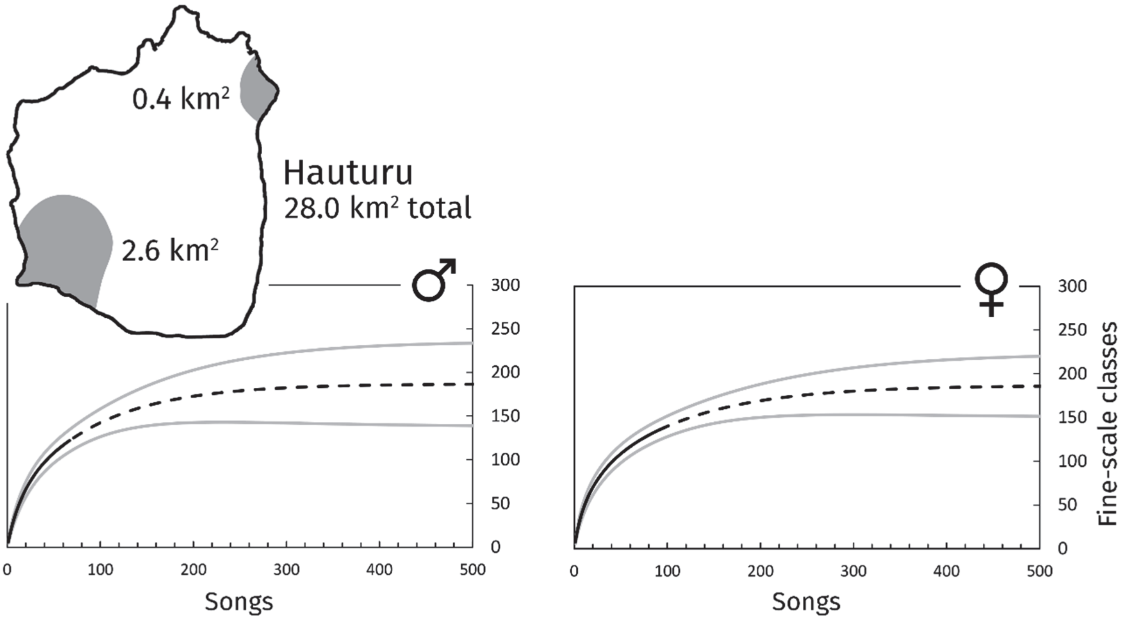
Sample only. Complete catalogue at http://tiny.cc/BellbirdsCatalogue.

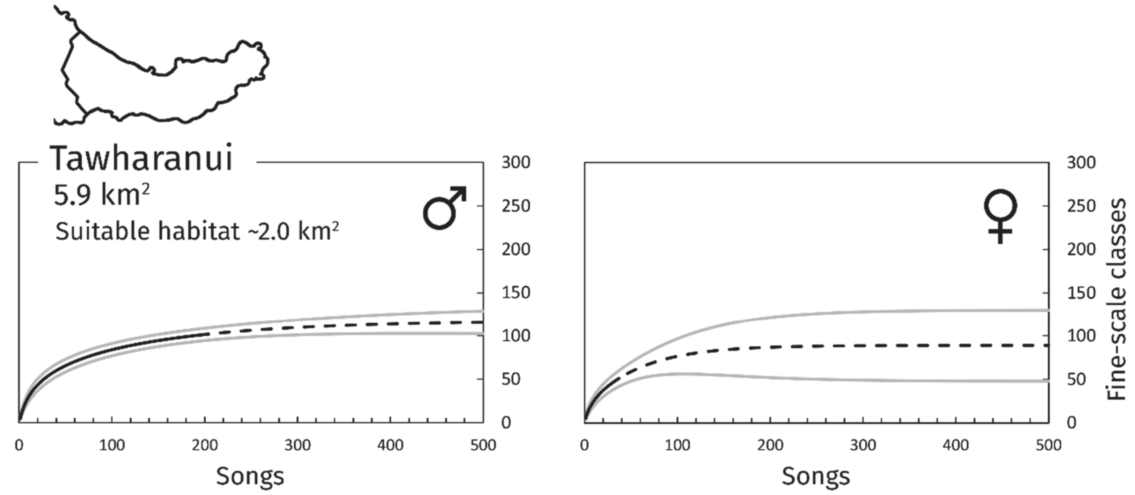
7.3.2 Repertoire size estimates for males and females at each site

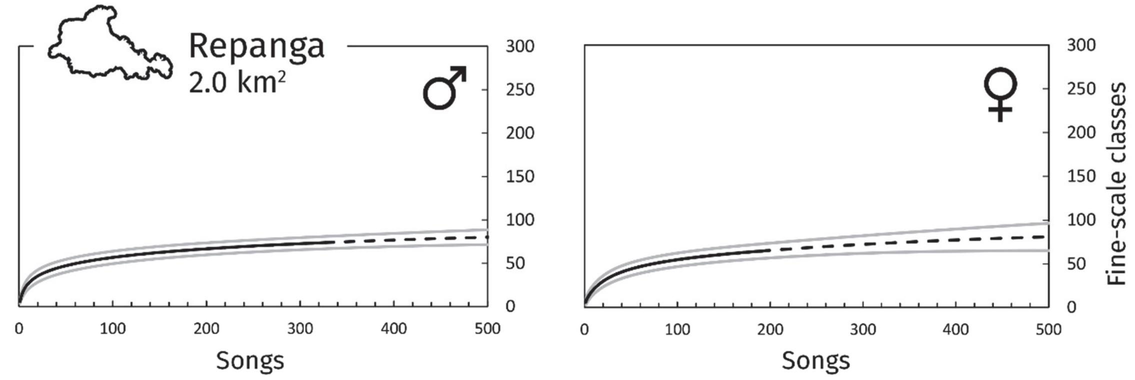
Syllable class accumulation curves for each site-and-sex group are shown in Figure 7.3 (fine-scale classification level) and Figure 7.4 (broad-scale classification level).

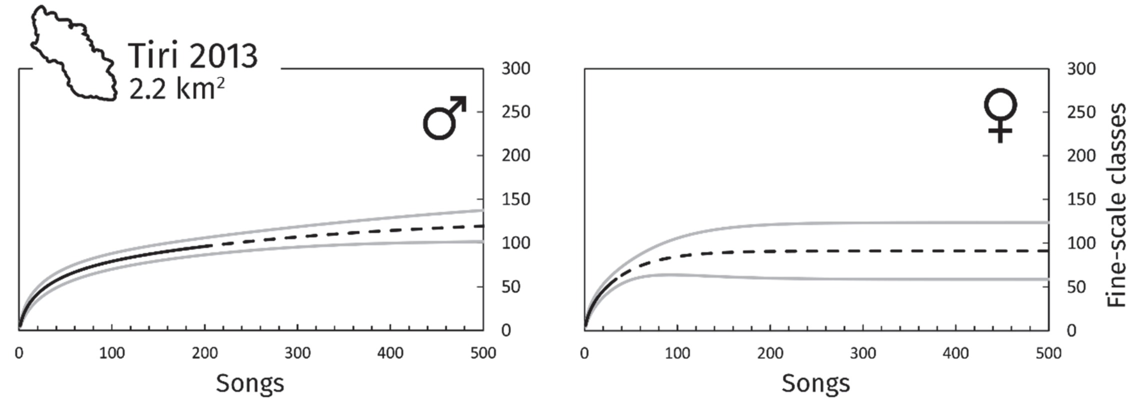

Tawhiti Rahi
1.5 km²
Representing
Poor Knights
2.5 km² total
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes

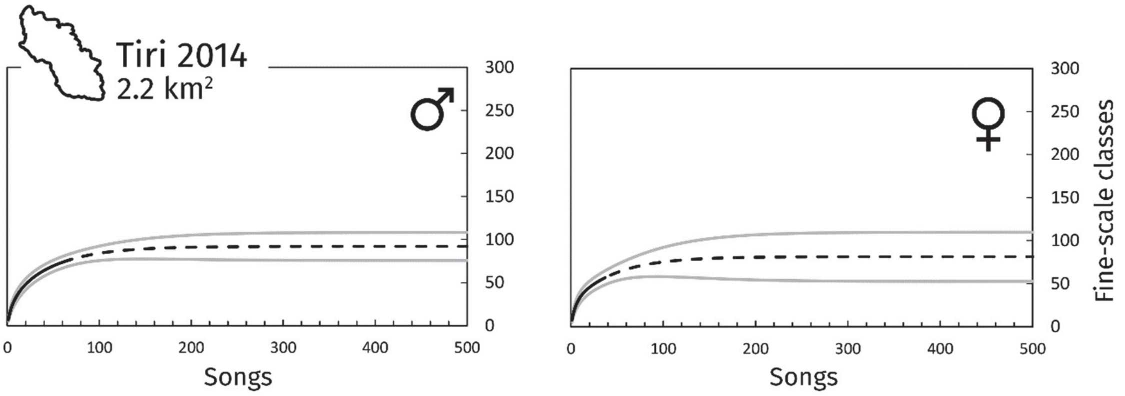

Lady Alice
1.4 km²
Representing
Hen & Chickens
7.3 km² total
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes

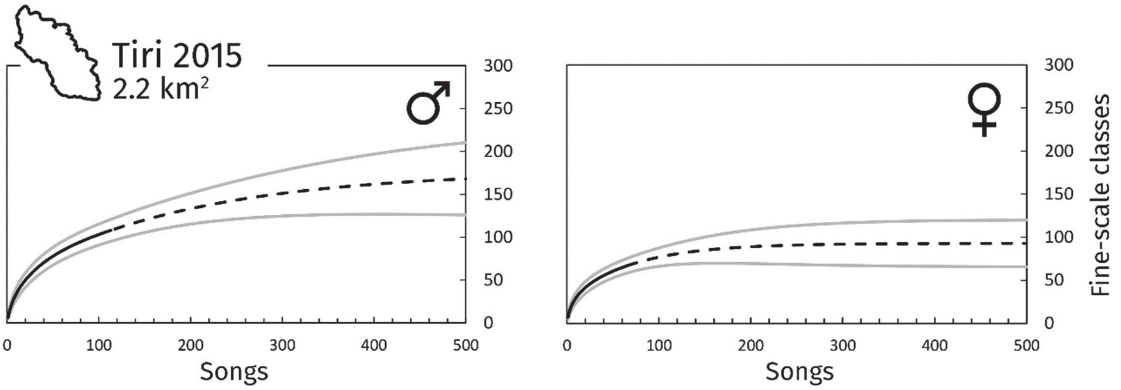

0.4 km²
2.6 km²
Hauturu
28.0 km² total
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes


Tawharanui
5.9 km²
Suitable habitat ~2.0 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes


Repanga
2.0 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes


Tiri 2013
2.2 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes


Tiri 2014
2.2 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes


Tiri 2015
2.2 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes

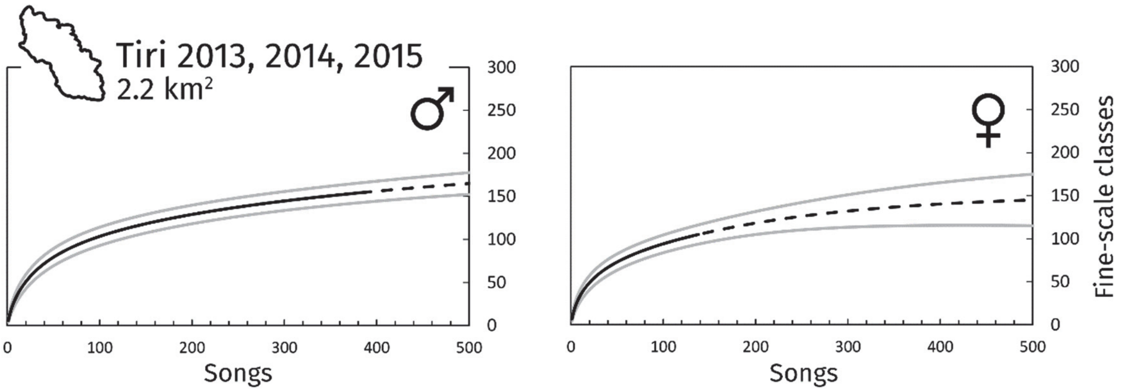
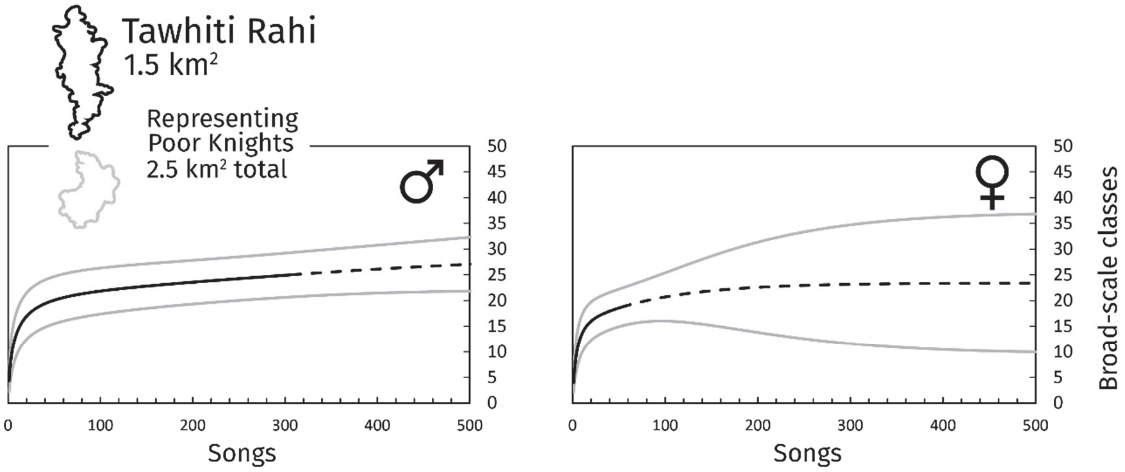

Tiri 2013, 2014, 2015
2.2 km²
♂
♀
0
50
100
150
200
250
300
0
100
200
300
400
500
Songs
Fine-scale classes

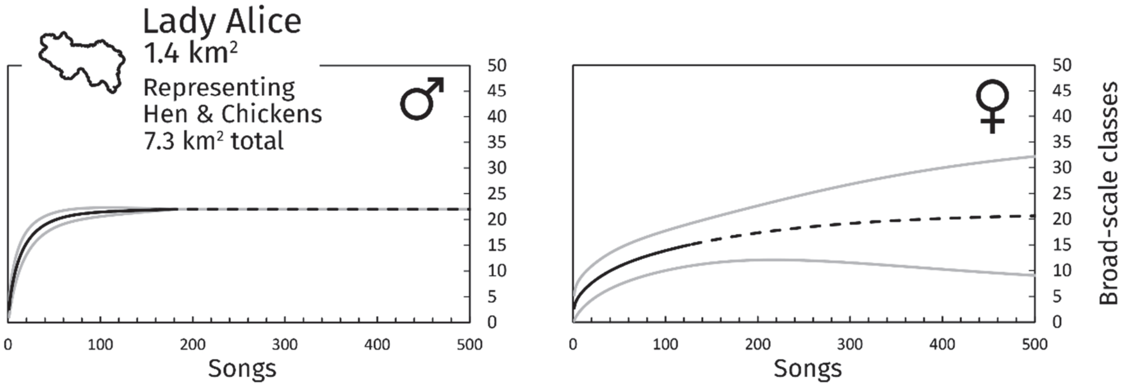
Figure 7.3—Fine-scale syllable class accumulation curve and estimate of fine-scale class richness (*S*) for each site-and-sex group.

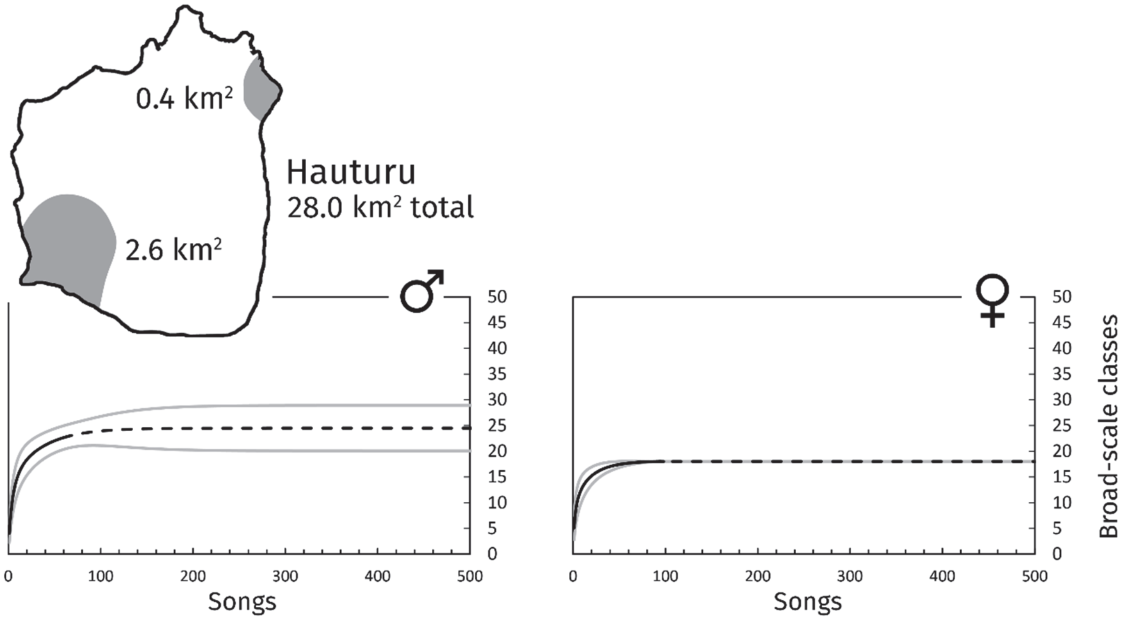
(***Above***) The black inner line shows the estimated true number of fine-scale classes (*S*) for the number of songs on the *x* axis. The dashed region of the line indicates the extrapolated estimate beyond the sample data. The grey lines show the 95% confidence interval for the estimate. Calculated with the statistical software *Estimate S* (Colwell, 2013) using 100 runs.

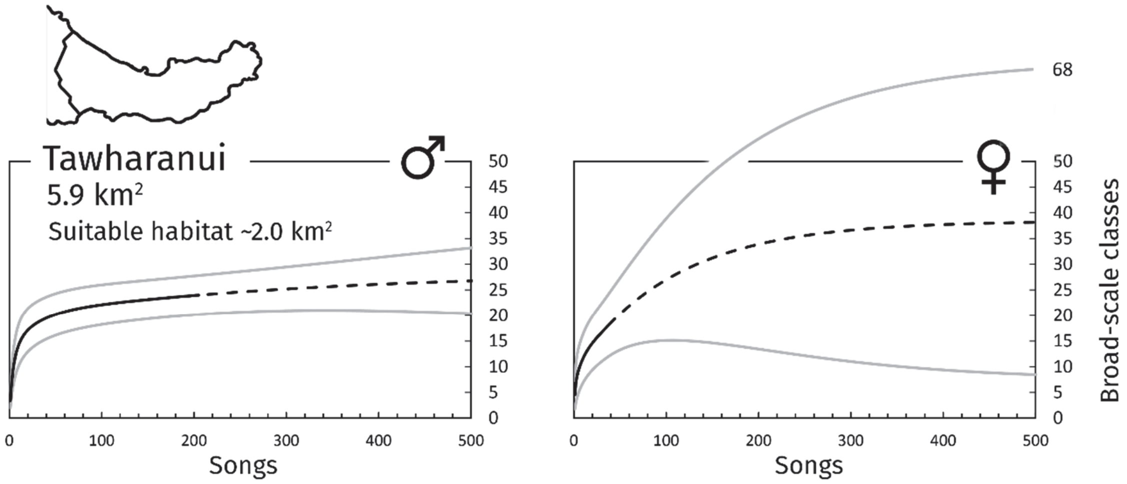
Figure 7.4—Broad-scale syllable class accumulation curve and estimate of broad-scale class richness (*S*) for each site-and-sex group.

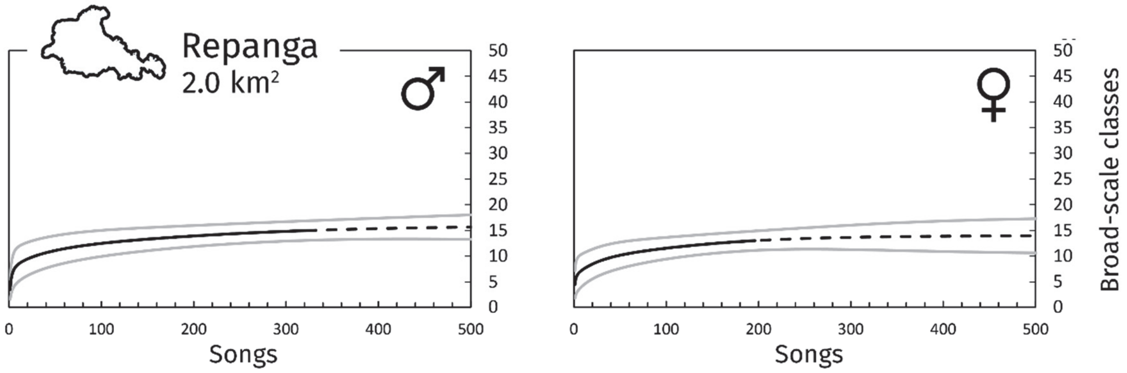
(***Below***) The black inner line shows the estimated true number of broad-scale syllable 'family' classes (*S*) for the number of songs on the *x* axis. The dashed region of the line indicates the extrapolated estimate beyond the sample data. The grey lines show the 95% confidence interval for the estimate. Calculated with the statistical software *Estimate S* (Colwell, 2013) using 100 runs. Some graphs show saturation, with convergent confidence intervals (e.g. Tiri 2014 male graph). The Tiri 2013 female data was insufficient for correct calculation, as evidenced by the logically-impossible confidence interval, which should not include 0 (number of 'species' is always positive).

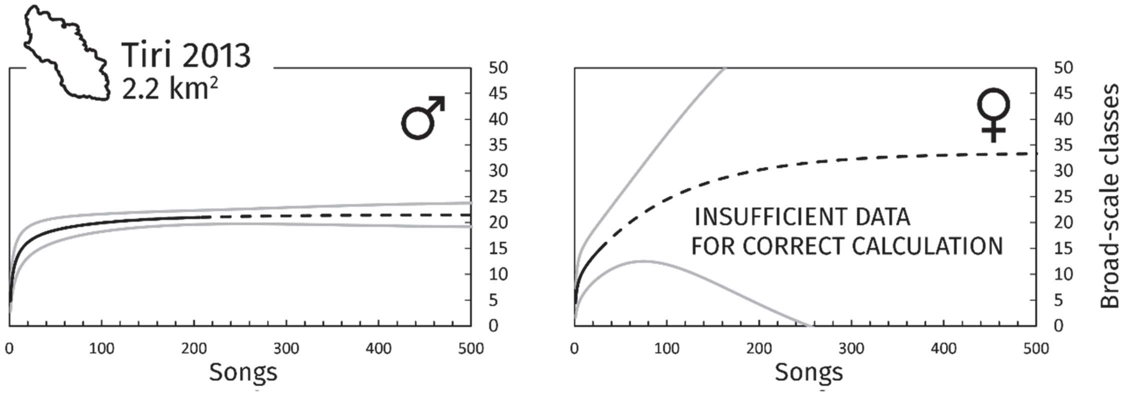

Tawhiti Rahi
1.5 km²
Representing
Poor Knights
2.5 km² total
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50

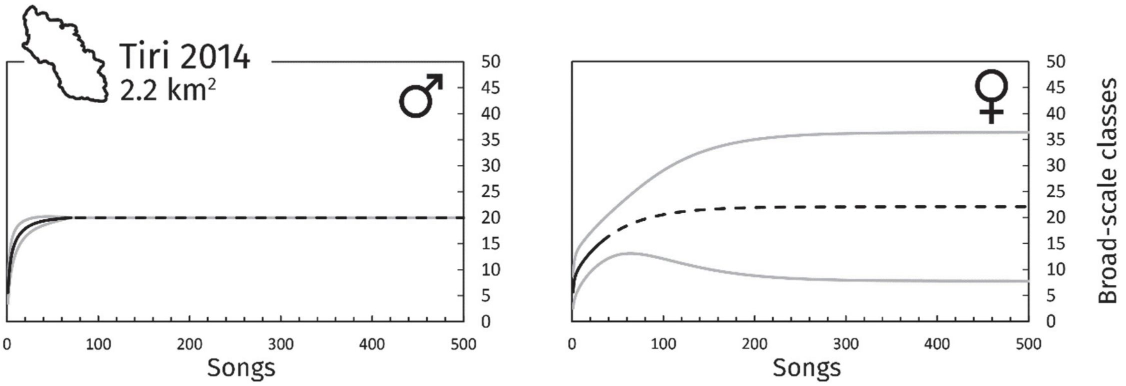

Lady Alice
1.4 km²
Representing
Hen & Chickens
7.3 km² total
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50

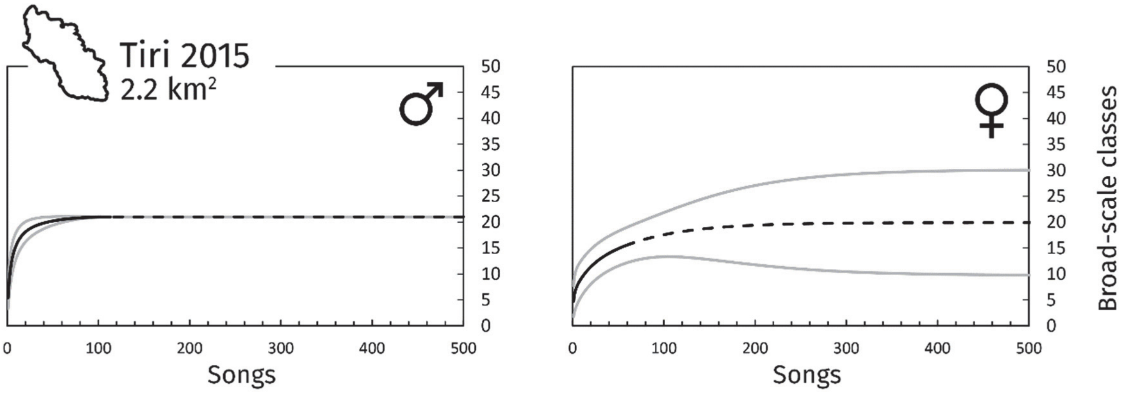

0.4 km²
Hauturu
28.0 km² total
2.6 km²
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50

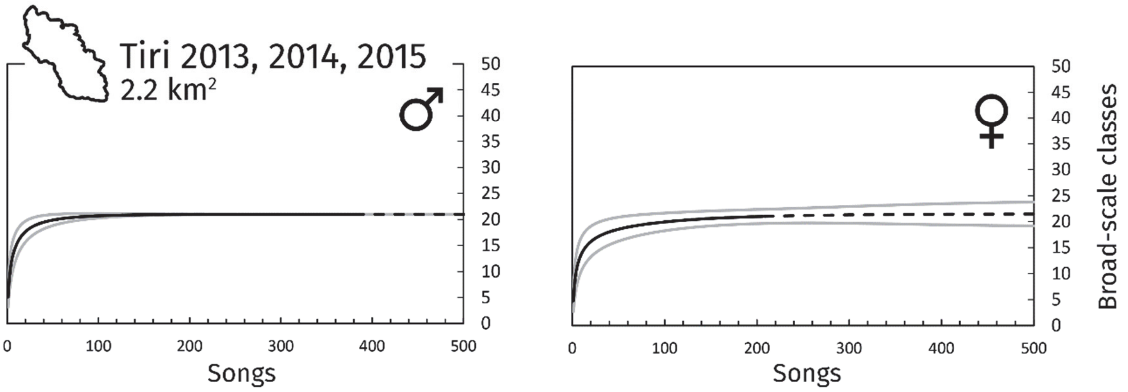

Tawharanui
5.9 km²
Suitable habitat ~2.0 km²
♂
♀
68
0
5
10
15
20
25
30
35
40
45
50
0
100
200
300
400
500
Songs
Broad-scale classes


Repanga
2.0 km²
♂
♀
0
5
10
15
20
25
30
35
40
45
50
0
100
200
300
400
500
Songs
Broad-scale classes


Tiri 2013
2.2 km²
♂
♀
INSUFFICIENT DATA
FOR CORRECT CALCULATION
0
5
10
15
20
25
30
35
40
45
50
0
100
200
300
400
500
Songs
Broad-scale classes


Tiri 2014
2.2 km²
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50


Tiri 2015
2.2 km²
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50


Tiri 2013, 2014, 2015
2.2 km²
♂
♀
Songs
Songs
Broad-scale classes
0
100
200
300
400
500
0
5
10
15
20
25
30
35
40
45
50

7.3.3 Prevalence distributions

Lists of classes shared between sexes, and their support values, are shown in Table 7.3 and Table 7.4 (fine-scale and broad-scale classification levels, respectively).

Table 7.3—Fine-scale syllable classes shared between sexes, and their support values.

Only shared classes occurring in at least two songs in each sex have been included.

Classes shared between sexes	Male support	Female support
Tawhiti Rahi		
Chiggle	0.01	0.08
Downsqueak(laser)	0.01	0.13
Pipe(low)_Weeyoo	0.14	0.10
SimUpDown_Upsqueak	0.01	0.08
Stutter	0.09	0.11
Woah	0.05	0.33
Lady Alice		
Crescendo(rough)	0.08	0.03
Down(LAI2)	0.03	0.06
Down(LAI3)	0.04	0.41
Downsweep(roughtosmooth)	0.02	0.90
Pipe(A#[6])	0.04	0.09
Pipe(A#[6]+)	0.07	0.04
Stutter	0.12	0.05
Hauturu		
Cough(clickhigh)	0.04	0.10
DinnerplateFlatArm	0.07	0.15
Down(LBI1)	0.03	0.06
Down(slightcurve+click)	0.04	0.10
Down(slightcurve1)	0.15	0.12
Down_Pipe(G#[6])	0.10	0.16
Down_Pipe(G[6]+)	0.07	0.26
Pipe(A#[6])	0.09	0.16
Pipe(A#[6]+)	0.10	0.31
Pipe(A#[6]bent)	0.10	0.05
Pipe(A[6])	0.12	0.03
Pipe(A[6]+)	0.04	0.06
Pipe(B[6])	0.13	0.14
Pipe(D#[6])	0.04	0.08
Pipe(D[6])	0.06	0.04

Pipe(D[6]+)	0.07	0.05
Pipe(D[7])	0.10	0.18
Pipe(D[7]+)	0.07	0.03
Pipe(E[6])	0.04	0.04
Pipe(F#[6])	0.06	0.18
Pipe(F#[6]+)	0.06	0.08
Pipe(F[6])	0.06	0.10
Stepdown(G[7]-D#[6])	0.04	0.04
Stutter	0.28	0.56
Stutter(high)	0.04	0.16
Stutter_Stutter(LBITAW+nownow)	0.07	0.05
Upsqueak(cleanglass)	0.06	0.09
Upsqueak(harmonic)	0.09	0.05
Tawharanui		
Downsqueak(chumpbit)_Pipe(high)	0.08	0.11
Pipe(A#[6])	0.01	0.24
Pipe(B[6]+)	0.02	0.16
Pipe(G[6]+)	0.02	0.05
Stutter	0.02	0.11
Stutter(click)	0.02	0.14
Stutter(TAW)	0.27	0.89
Upsqueak(kitten)	0.19	0.08
Repanga		
Alarmy(short+highbit)	0.03	0.02
Alarmy(short+up)	0.31	0.02
Alarmy(twobitCUV+beware)	0.11	0.03
Chiggle	0.02	0.04
CUVsqueak	0.33	0.06
Pipe(A#[6]+)	0.04	0.02
Pipe(A[6]+)	0.01	0.08
Pipe(B[7])	0.01	0.05
Pipe(C#[7])	0.02	0.07
Pipe(C[7]+)	0.01	0.07
Pipe(E[6]+)	0.38	0.06
Pipe(F#[6])	0.06	0.02
Pipe(F#[6]+)	0.07	0.06
Pipe(F[6])	0.30	0.09
Pipe(F[6]+)	0.06	0.02
Pipe(F[7]+)	0.01	0.02
Stutter(alarmy)	0.01	0.39
Stutter(decurved)	0.27	0.83

Tiri 2013		
Peaksqueak(cutting)	0.02	0.26
Peaksqueak(whistley)	0.03	0.06
Pipe(A[6]+)	0.01	0.03
Pipe(C#[6])	0.03	0.03
Pipe_Downsqueak	0.08	0.06
Stutter	0.35	0.91
Tiri 2014		
Downsqueak(hook)	0.41	0.11
Peaksqueak(cutting)	0.07	0.11
Pipe(A[6]+)	0.01	0.05
Pipe(C#[6]+)	0.01	0.08
Pipe(D#[6])	0.03	0.03
Pipe_Downsqueak	0.03	0.03
Stutter	0.39	0.95
Tiri 2015		
Downsqueak(hook)	0.54	0.02
Peaksqueak(cutting)	0.11	0.11
Peaksqueak(whistley)	0.04	0.02
Pipe(D#[6])	0.02	0.05
Pipe(E[6])	0.17	0.09
Pipe(E[6]+)	0.02	0.05
Pipe_Downsqueak	0.02	0.06
Stutter	0.17	0.94
Tiri 2013, 2014, 2015		
Chiup	<0.01	0.59
Downsqueak(hook)	0.35	0.04
Peaksqueak(cutting)	0.05	0.15
Peaksqueak(whistley)	0.05	0.02
Pipe(A[6]+)	0.04	0.02
Pipe(C#[6])	0.02	0.02
Pipe(D#[6])	0.01	0.03
Pipe(E[6])	0.01	0.05
Pipe(E[6]+)	< 0.01	0.02
Pipe(G#[6])	0.01	0.01
Pipe_Downsqueak	0.05	0.05
Stutter	0.31	0.93

Table 7.4—Broad-scale syllable classes shared between sexes, and their support values.

Only shared classes occurring in at least two songs in each sex have been included.

Classes shared between sexes	Male support	Female support
Tawhiti Rahi		
Bip	0.37	0.06
Chiggle	0.05	0.07
Cough	0.59	0.09
Down_Pipe	0.04	0.35
Downsqueak	< 0.01	0.09
Peaksqueak	0.18	0.46
Pipe	0.80	0.35
Pipe_Down	0.12	0.76
RoughBuzz	0.17	0.37
Stepdown	0.02	0.15
Stutter	0.11	0.09
Upsqueak	0.02	0.13
Lady Alice		
Alarmy	0.09	0.03
Down	0.12	0.98
Pipe	0.78	0.49
Pipe_Down	0.04	0.86
RoughBuzz	0.08	0.04
Stepdown	0.05	0.17
Stepup	0.05	0.02
Stutter	0.12	0.03
Trill	0.47	< 0.01
Hauturu		
Alarmy	0.08	0.05
Cough	0.23	0.11
Down	0.26	0.40
Down_Pipe	0.23	0.77
Downsqueak	0.09	0.26
Flatsqueak	0.22	0.05
Pipe	0.82	0.93
Pipe_Down	0.17	0.30
Shrieky	0.15	0.33
Stepdown	0.29	0.70
Stutter	0.29	0.65
Stutter(repeats)	0.06	0.05
Upsqueak	0.09	0.14
Tawharanui		
Alarmy	0.27	0.16
Down	0.03	0.54
Downsqueak	0.27	0.51
Downsqueak_Pipe	0.20	0.11
Pipe	0.54	0.70
Stepup	0.16	0.14
Stutter	0.26	0.89
Repanga		
Alarmy	0.39	0.45
Chiggle	0.02	0.03
Chortle	0.33	0.89
Peaksqueak	0.30	0.05
Pipe	0.83	0.84

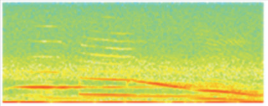
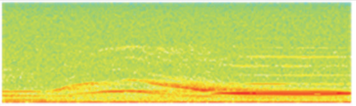
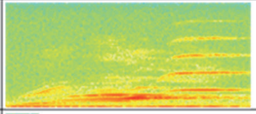
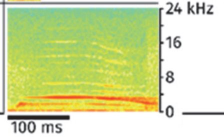
Stutter	0.25	0.85
Tiri 2013		
Chump	0.61	0.39
Downsqueak	0.44	0.48
Flatsqueak	0.14	0.12
Peaksqueak	0.05	0.33
Pipe	0.55	1.00
Stepdown	0.01	0.55
Stutter	0.36	0.91
Waah	0.85	0.03
Tiri 2014		
Chortle	0.27	0.44
Chump	0.54	0.78
Downsqueak	0.74	0.64
Flatsqueak	0.07	0.36
Peaksqueak	0.21	0.11
Pipe	0.50	0.94
Stepdown	0.10	0.81
Stutter	0.39	0.97
Waah	0.86	0.06
Tiri 2015		
Chortle	0.24	0.03
Chump	0.39	0.60
Downsqueak	0.79	0.48
Flatsqueak	0.10	0.06
Peaksqueak	0.10	0.13
Pipe	0.55	0.98
Pipe_Downsqueak	0.13	0.06
Stepdown	0.02	0.95
Stutter	0.18	0.94
Stutter(repeats)	0.09	0.02
Tiri 2013, 2014, 2015		
Chortle	0.18	0.14
Chump	0.54	0.59
Downsqueak	0.59	0.52
Flatsqueak	0.12	0.16
Peaksqueak	0.09	0.18
Pipe	0.54	0.97
Pipe_Downsqueak	0.11	0.04
Stepdown	0.03	0.80
Stutter	0.32	0.94
Stutter(repeats)	0.03	0.02
Waah	0.84	0.02

7.3.4 Catalogue of syllable types at each site common to both sexes

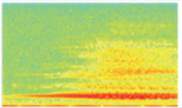
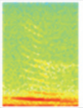
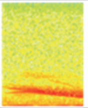
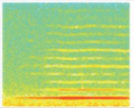
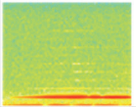
Table 7.5—List of fine-scale syllable classes sung by both males and females at each site.

(***Below***) The list for each site corresponds to the male-female bar overlap regions in Figure 5.7 (Chapter 5). Scalebars are provided.

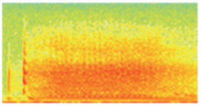
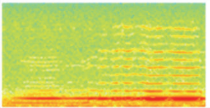
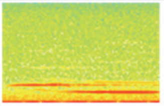
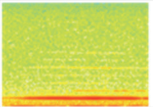
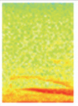
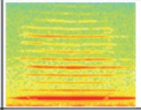
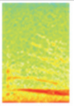
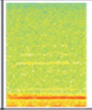
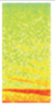
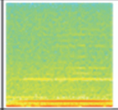
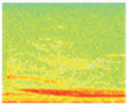
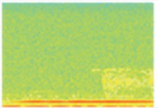
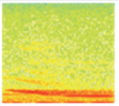
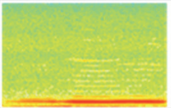
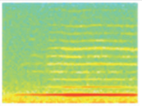
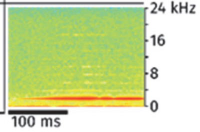
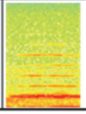
Tawhiti Rahi

Fine-scale syllable class	Spectrogram
Chiggle	
Downsqueak(laser)	
Pipe(low)_Weeyoo	
SimUpDown_Upsqueak	
Stutter	
Woah	

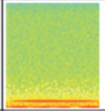
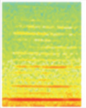
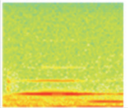
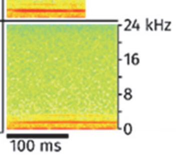
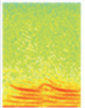
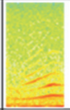
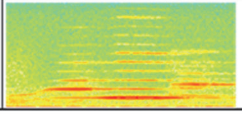
Lady Alice Island

Fine-scale syllable class	Spectrogram
Crescendo(rough)	
Down(LAI2)	
Down(LAI3)	
Downsweep(roughtosmooth)	
Pipe(A#[6])	
Pipe(A#[6]+)	
Stutter	

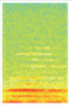
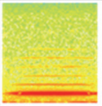
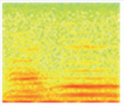
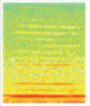
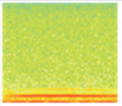
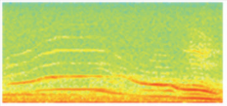
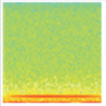
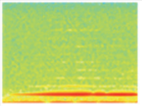
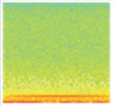
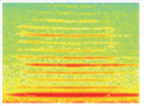
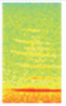
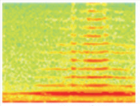
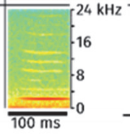
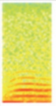
Hauturu

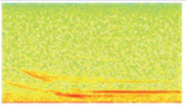
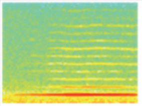
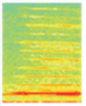
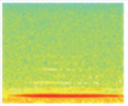
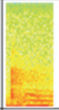
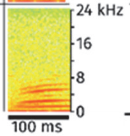
Fine-scale syllable class	Spectrogram	Fine-scale syllable class	Spectrogram
Cough(clickhigh)		Pipe(A#[6]bent)	
DinnerplateFlatArm		Pipe(A[6])	
Down(LBI1)		Pipe(A[6]+)	
Down(slightcurve+click)		Pipe(B[6])	
Down(slightcurve1)		Pipe(D#[6])	
Down_Pipe(G#[6])		Pipe(D[6])	
Down_Pipe(G[6]+)		Pipe(D[6]+)	
Pipe(A#[6])		Pipe(D[7])	
Pipe(A#[6]+)		Pipe(D[7]+)	

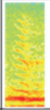
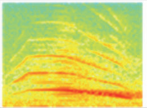
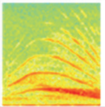
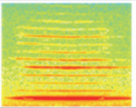
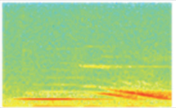
Hauturu *(cont.)*

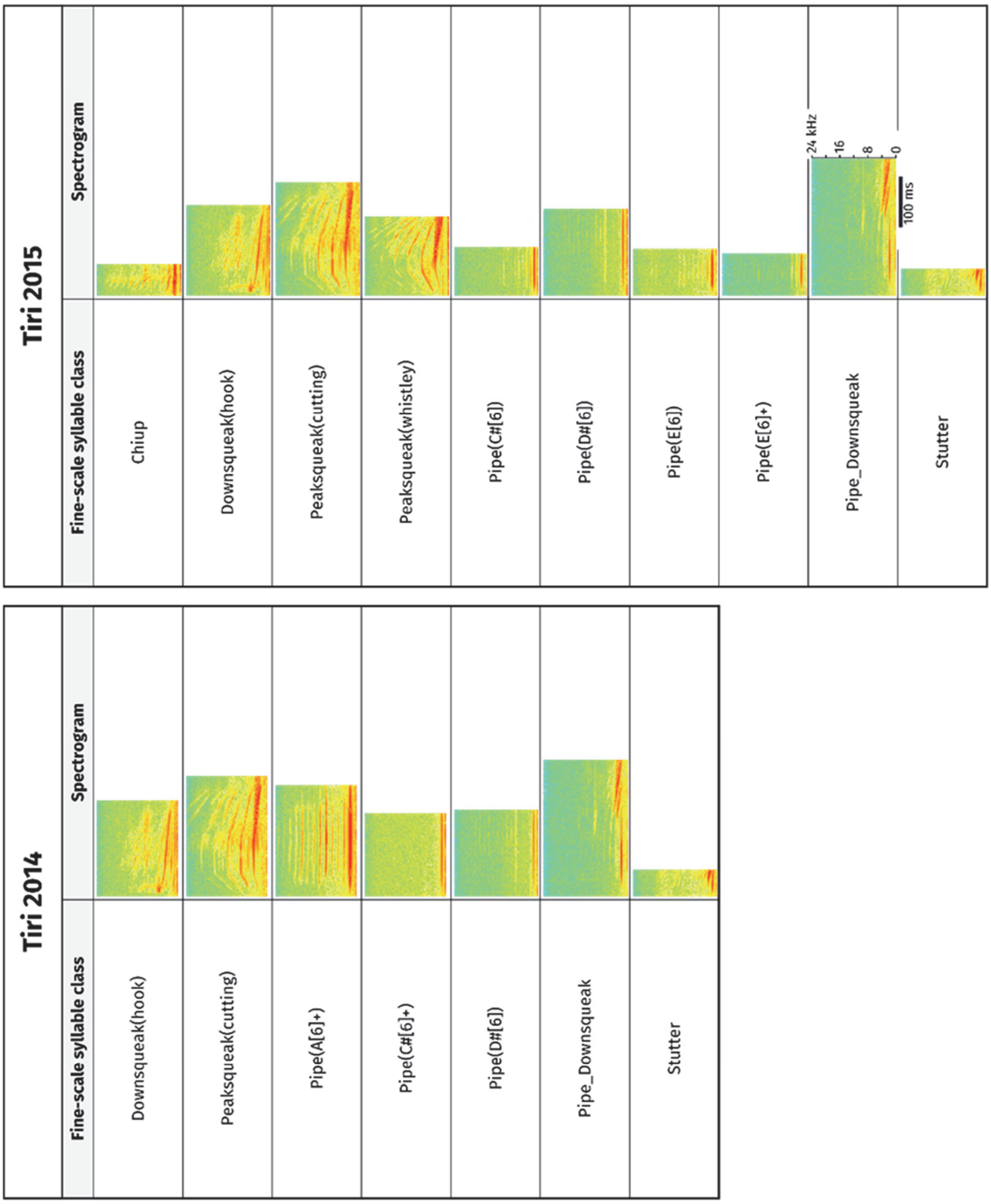
Fine-scale syllable class	Spectrogram	Fine-scale syllable class	Spectrogram
Pipe(E[6])		Pipe(F[6])	
Pipe(F#[6])		Stepdown(G[7]-D#[6])	
Pipe(F#[6]+)		Stutter	
		Stutter(high)	
		Stutter_Stutter(LBITAW+nownow)	
		Upsqueak(cleanglass)	
		Upsqueak(harmonic)	

Repanga

Fine-scale syllable class	Spectrogram	Fine-scale syllable class	Spectrogram
Alarmy(short+highbit)		Pipe(C[7]+)	
Alarmy(short+up)		Pipe(E[6]+)	
Alarmy(twobitCUV+beware)		Pipe(F#[6])	
Chiggle		Pipe(F#[6]+)	
CUVsqueak		Pipe(F[6])	
Pipe(A#[6]+)		Pipe(F[6]+)	
Pipe(A[6]+)		Pipe(F[7]+)	
Pipe(B[7])		Stutter(alarmy)	
Pipe(C#[7])		Stutter(decurved)	

Tawharanui	
Fine-scale syllable class	**Spectrogram**
Downsqueak(chumpbit)_Pipe(high)	
Pipe(A#[6])	
Pipe(B[6]+)	
Pipe(G[6]+)	
Stutter	
Stutter(click)	
Stutter(TAW)	
Upsqueak(kitten)	

Tiri 2013	
Fine-scale syllable class	**Spectrogram**
Chiup	
Peaksqueak(cutting)	
Peaksqueak(whistley)	
Pipe(A[6]+)	
Pipe(C#[6])	
Pipe_Downsqueak	
Stutter	



Tiri 2014

Fine-scale syllable class	Spectrogram
Downsqueak(hook)	
Peaksqueak(cutting)	
Pipe(A[6]+)	
Pipe(C#[6]+)	
Pipe(D#[6])	
Pipe_Downsqueak	
Stutter	

Tiri 2015

Fine-scale syllable class	Spectrogram
Chiup	
Downsqueak(hook)	
Peaksqueak(cutting)	
Peaksqueak(whistley)	
Pipe(C#[6])	
Pipe(D#[6])	
Pipe(E[6])	
Pipe(E[6]+)	
Pipe_Downsqueak	
Stutter	

Tiri 2013, 2014, 2015

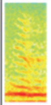
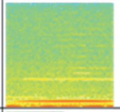
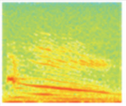
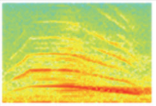
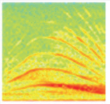
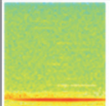
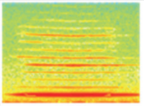
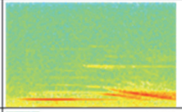
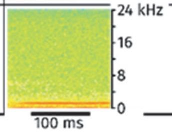
Fine-scale syllable class	Spectrogram	Fine-scale syllable class	Spectrogram
Chiup		Pipe(D#[6])	
Downsqueak(hook)		Pipe(E[6])	
Peaksqueak(cutting)		Pipe(E[6]+)	
Peaksqueak(whistley)		Pipe(G#[6])	
Pipe(A[6]+)		Pipe_Downsqueak	
Pipe(C#[6])		Stutter	
Pipe(C#[6]+)			

Table 7.6—List of broad-scale syllable classes (syllable 'families') sung by both males and females at each site.

The list for each site corresponds to the male-female bar overlap regions in Figure 5.7 (Chapter 5). For spectrogram examples of the families see catalogue of class types in Table 7.2

Tawhiti Rahi	Lady Alice	Hauturu	Tawharanui	Repanga
Bip	Alarmy	Alarmy	Alarmy	Alarmy
Chiggle	Down	Cough	Down	Chiggle
Cough	Pipe	Down	Downsqueak	Chortle
Down_Pipe	Pipe_Down	Down_Pipe	Downsqueak_Pipe	Peaksqueak
Downsqueak	RoughBuzz	Downsqueak	Pipe	Pipe
Peaksqueak	Stepdown	Flatsqueak	Stepup	Stutter
Pipe	Stepup	Pipe	Stutter	
Pipe_Down	Stutter	Pipe_Down		
RoughBuzz	Trill	Shrieky		
Stepdown		Stepdown		
Stutter		Stutter		
Upsqueak		Stutter(repeats)		
		Upsqueak		

Tiri 2013	Tiri 2014	Tiri 2015	Tiri 2013, 2014, 2015
Chump	Chortle	Chortle	Chortle
Downsqueak	Chump	Chump	Chump
Flatsqueak	Downsqueak	Downsqueak	Downsqueak
Peaksqueak	Flatsqueak	Flatsqueak	Flatsqueak
Pipe	Peaksqueak	Peaksqueak	Peaksqueak
Stepdown	Pipe	Pipe	Pipe
Stutter	Pipe_Downsqueak	Pipe_Downsqueak	Pipe_Downsqueak
Stutter(repeats)	Stepdown	Stepdown	Stepdown
Waah	Stutter	Stutter	Stutter
	Stutter(repeats)	Stutter(repeats)	Stutter(repeats)
	Waah		Waah

7.3.5 Syllable classes shared between sites

Lists of syllable classes shared between sites are available as an *Excel* spreadsheet at this link: http://tiny.cc/BellbirdRepertoires