Biases in the learning and plasticity of birdsong

Logan James Smith

Department of Biology
McGill University, Montréal
August 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Dedication

For my mom

Table of Contents

Abstract

Communicative behaviors are often important for the survival and reproductive ability of organisms. Many of these communicative behaviours are learned and can be modified through life, and both processes can be influenced by internal and external factors. Discovering these factors that guide the learning and modification of behaviour is fundamental to our understanding of animal behaviour and the neural basis of behaviour. One of the most powerful model systems for this purpose is the songbird. Like humans, many species of songbirds learn their vocal signals during a sensitive period of development. After adulthood, many species will no longer learn new songs, but instead exhibit behavioural modulation and plasticity with subtle song changes that can occur rapidly across social contexts or slowly over long periods of time. In this thesis, we explored various forms of vocal modulation, plasticity and learning in adult and developing songbirds. We first measured longitudinal age-dependent changes in adult Bengalese finches and found that over time, individuals produced faster songs with more stereotyped sequencing. Next, we found that short-term context-dependent modulations to song features could provide predictive insight into subsequent age-dependent changes of those features. Given these results, we used a comparative approach to understand whether the same processes occurred in the closely related zebra finch. We analyzed the context- and age-dependent changes of adult zebra finches and found multiple similarities in the nature and predictability of age-dependent changes between the two species. We next turned to the learning of song to assess the predictability of song development. For this, we used an experimental tutoring approach to ask whether zebra finches are biased to produce certain types of sequences over others. We indeed found that sequence learning biases led to many common song patterns across individuals, and, thus, in our final study, explored the extent to which motor biases could contribute to these learning biases. Altogether, this thesis provides integrative and fundamental data on links between short- and long-term changes to communication and provides a foundation for further understanding of mechanisms that drive behavioral plasticity, development and control.

Résumé

Les comportements de communication sont souvent importants pour la survie et la capacité de reproduction des organismes. Bon nombre de ces comportements de communication sont appris et peuvent être modifiés tout au long de la vie, et les deux processus peuvent être influencés par des facteurs internes et externes. La découverte de ces facteurs qui guident l'apprentissage et la modification du comportement est fondamentale pour notre compréhension du comportement des animaux et de la base neuronale du comportement. L'un des modèles les plus puissants à cet effet est l'oiseau chanteur. Comme les humains, de nombreuses espèces d'oiseaux chanteurs apprennent leurs signaux vocaux au cours d'une période sensible du développement. À l'âge adulte, de nombreuses espèces n'apprendront plus de nouvelles chansons, mais présenteront une modulation et une plasticité comportementales accompagnées de modifications subtiles de la chanson pouvant se produire rapidement dans des contextes sociaux ou lentement, sur de longues périodes. Dans cette thèse, nous avons exploré diverses formes de modulation vocale, de plasticité et d'apprentissage chez des oiseaux chanteurs adultes et en développement. Nous avons d'abord mesuré les changements longitudinaux liés à l'âge des pinsons bengalais adultes en fonction de l'âge et nous avons constaté qu'au fil du temps, les individus produisaient des chansons plus rapides avec un séquençage plus stéréotypé. Ensuite, nous avons constaté que les modulations à court terme dépendant du contexte des caractéristiques de chanson pourraient fournir un aperçu prédictif des modifications ultérieures de ces caractéristiques dépendant de l'âge. Compte tenu de ces résultats, nous avons utilisé une approche comparative pour déterminer si les mêmes processus se produisaient chez le pinson des zèbres, étroitement apparenté. Nous avons analysé les changements liés au contexte et à l'âge des pinsons adultes et constaté de nombreuses similitudes dans la nature et la prévisibilité des changements liés à l'âge entre les deux espèces. Nous nous sommes ensuite tournés vers l'apprentissage de la chanson pour évaluer la prévisibilité du développement de la chanson. Pour cela, nous avons utilisé une approche de tutorat expérimental pour demander si les pinsons zébrés sont biaisés pour produire certains types de séquences par rapport à d'autres. Nous avons en effet constaté que les biais d'apprentissage par séquence conduisaient à de nombreux schémas de

chansons communs chez les individus et, par conséquent, dans notre dernière étude, nous avons examiné dans quelle mesure les biais moteurs pouvaient contribuer à ces biais d’apprentissage. Globalement, cette thèse fournit des données intégratives et fondamentales sur les liens entre les changements de communication à court et à long terme et fournit une base pour mieux comprendre les mécanismes qui régissent la plasticité, le développement et le contrôle du comportement.

Preface

This thesis is presented in a manuscript-based format with a brief general introduction followed by five chapters, each of which has either been published or is in the process of publication in a peer-reviewed journal. Chapters that have already been published (Chapters 1, 2 and 4) have been slightly reformatted into a common layout. Where applicable, online supplementary material has been included as an appendix to the chapter.

Acknowledgements

First and foremost, thank you Jon Sakata. You've been an incredibly supportive supervisor in all aspects of my graduate endeavors. I am so grateful for everything you've taught me over these years and I look forward to continuing our (many) exciting projects and ideas in the future. I couldn't have asked for a better supervisor.

I'd also like to thank my supervisory committee, Rüdiger Krahe and Kris Onishi, who have provided incredibly useful comments and support throughout my PhD. In addition, thank you to Sarah Woolley whose mentorship has meant so much to me. I also thank the many collaborators I've been lucky enough to work with, including: Jonathan Prather, Karagh Murphy, Jonathan Davies, William Pierce, Maxwell Farrell, Ignacio Morales-Castilla, Frédéric Boivin, Kazuhiro Wada, Chihiro Mori, Raghav Rajan, and Shikha Kalra, as well as amazing undergrads including Raina Fan, Jennifer Dai, Herie Sun and Ronald Davies Jr.

I've been incredibly lucky to work with amazing labmates throughout this process. I am especially grateful to Laura Matheson, Danielle Toccalino, Nancy Chen, Helena Barr and Erin Wall who provided me with incredible support in so many ways across the years. There have also been so many other fantastic graduate and undergraduate students who have come through the lab that I am very grateful to have worked with.

The department of biology at McGill has been an incredible place to spend my ten years in Montréal. I am especially appreciative of Lauren Chapman, who introduced me to field work and has included me in her lab community for years, as well as the rest of the Lake

Nabugabo crew: Suzanne Gray, Tyler Moulton, Laura McDonnell and Dalal Hanna. Other faculty in the department have been very supportive as well, in particular, I thank Anna Hargreaves, Andrew Hendry and Gregor Fussmann for going above and beyond with student engagement and sharing their wisdom with me. Finally, the graduate student association (BGSA) has been an incredible source of friendship and support. I couldn't have made it through this without such a great social circle. There are too many people to name, but I'm especially grateful for Beth Nyboer, Dustin Raab, Tim Thurman, Krista Oke, Jacob Ziegler, Anna Cook, Christine Yergeau (who I also thank for translating my abstract!!), Emma Hudgins, Angela Yu, Sophie Koch, Lindsay Potts, Rahul Rote, Kelvin Ka Kay Ip, Marisol Valverde, Vincent Fugère, Claudia Wever, Sami Chaaban, Michelle Gros, Alex Crew, Sarah Baldwin, Mary-Rose Gill and Kiyoko Gotanda.

I've had the great pleasure of living with Anna Luz, Kyle Godden, Reilly MacDonald and Melissa Tamporello during my PhD. Thank you for being the best friends and roommates I could imagine. Beyond them, I've had the best support from too many friends to list, but I'd especially like to thank: Christina Tadiri, Aviva Wolpert, Emily Hacker, Jillian Oliver, Lorraine Chuen, Minhee Kim, Allen Leary, Laura Shevchenko, Maryse Thomas, Jade Tremblay, Sara Kurland, Roberta Hayes, Bergþóra Jónsdóttir, Margot Flores Torre, Valerie Campanelli, Benjamin Hoffman, Felix Muzny, Stephanie Jones, Nathan Lenssen, Lisa Han and Benjamin Kunstman. And most importantly, thank you Dylan DiCicco.

Finally, thank you to my family who have been so supportive of everything I've done for my whole life. Thank you to my siblings, Katie and Caleb Dickinson and Tyler and Sara Hansen who are also my best of friends. Thank you to my extended family including many aunts, uncles and cousins for their unending love. Thank you to Barb and James Howard – I hope I would've made you proud Gpa. And an exceptionally big thank you to my mom, the best.

For financial support, I thank my supervisor Jon Sakata, the McGill Department of Biology, the Centre for Research on Brain, Language and Music (who I also thank for fostering an incredibly engaging scientific community), the Jane Stewart Prize, the Heller award and the Center for Studies in Behavioral Neurobiology. Details of financial support and contributions from collaborators are listed in the acknowledgements section of each chapter.

Contribution of authors

I am the first author and Jon T. Sakata is the senior author on all five of the chapters contained in this thesis. For chapters 1-4, we are the only co-authors, and for all chapters we both contributed to the development of the project, the collection of data, the analysis of data and the writing and revising of the manuscript.

For chapter 5, there are three additional co-authors. R. Davies Jr. aided us with the machine learning techniques that we employed. C. Mori and K. Wada conducted surgery and collected data on untutored birds. All three additional co-authors helped to write and edit the manuscript.

Statement of originality

All five chapters in this thesis constitute original scientific progress and have been published or prepared to be published in high quality, peer-reviewed journals. Chapters 1 and 2 were published in the *Journal of Neurophysiology*, Chapter 3 has been resubmitted to the journal *Brain Research* following generally positive comments from reviewers, Chapter 4 was published in *Current Biology*, and Chapter 5 is in preparation to submit to the *Proceedings of the Royal Society B*.

Chapter 1 constitutes the first detailed account of longitudinal song changes within adult Bengalese finches, and Chapter 2 documents the first within-bird relationships between rapid modulation of song and long-term changes to song within any species of songbird. Chapter 3 provides the first demonstration of such relationships between short- and long-term changes to song within a second species, the zebra finch. Chapter 4 provides fundamental data on sequence learning biases in zebra finch, using an experimental tutoring approach to uncover common song properties. And Chapter 5 provides the first comparison of data from the songs of untutored birds to the common patterns uncovered in Chapter 4.

General introduction

Animals must learn many of their most important behaviours. This ranges from developmentally linked behaviours such as locomotion, to quite complex and context-specific behaviours including foraging tactics, tool use, courtship displays, and many forms of communication. Often, these latter, more complex behaviours are learned from other individuals, a process which may "short cut" longer individual learning tactics (Galef and Laland 2005; Reader 2016). Across generations, these socially learned behaviours may lead to distinct cultural traditions, which can be important for group integration and mate choice (Laland and Galef 2009; Tomback and Baker 1984). Furthermore, once behaviours are acquired, they are not static, and can exhibit modulation and plasticity within an individual's lifetime based on external or internal factors (Alcock 2009; Kolb et al. 2003).

Environmental (i.e., external) factors can interact with internal biases to influence which, when and how behaviours are developed or modified. For example, classic studies on imprinting in chicks revealed how these internal biases and environmental conditions can interact. Chicks would generally imprint on the first object they were exposed to; however, when presented with multiple options simultaneously, including a more relevant object like a stuffed junglefowl, chicks were highly biased to imprint on that more relevant object compared to a simple red box (Horn and McCabe 1984; Sluckin and Salzen 1961). Less intuitively, chicks were also biased against imprinting on a yellow box when presented with five differently colored boxes (Schaefer and Hess 1959), a bias that is only revealed in this highly specific environmental context. Other environmental factors can interact with these biases: chicks reared under impoverished visual stimuli are slower to exhibit the species-typical imprinting preference (Bolhuis et al. 1985).

Similar interactions between the environment and internal biases in learning have been observed in more complex, socially-learned behaviours. For example, in many species of songbird, courtship songs are learned through imitation from an adult conspecific during a sensitive period in development (see below), and internal factors can influence this process. For instance, Bengalese and zebra finches are biased to learn from a conspecific tutor over a heterospecific tutor when tutored with both (Clayton 1988), and when tutored with only heterospecific song, both species will produce songs that have some intermediate properties between typical conspecific and heterospecific song (Clayton 1989). In addition, white-crowned sparrows tutored with different song segments presented in isolation display a sequencing bias towards a species-typical song sequence (Soha and Marler 2001). External factors, including social influences, can also impact the song learning process and interact with internal biases. For instance, white-crowned sparrows will generally learn from conspecific but not heterospecific song playbacks, however, heterospecific song content can be learned if the sparrow is reared with a live heterospecific tutor (Baptista and Petrinovich 1984; Marler 1970a).

In this thesis we leverage the power of the birdsong model system to explore unanswered questions regarding the extent to which complex and learned behaviours are modified outside of development, how external factors (here, social influences) can predict or change these modifications, as well as how internal factors can guide learning during development. This research is particularly interesting as a general model system for vocal learning, plasticity and communication, including human speech (see below).

Sensitive periods, modulation and plasticity of vocal communication

Sensitive periods, epochs during which the effects of experience on behavior and the brain are particularly strong, regulate the extent of learning and plasticity in many sensory and behavioural systems (Knudsen 2004). For example, foundational research in the visual system has shown how perturbances in vision during a very narrow window of time in development can lead to life-long impairments in visual acuity (Hensch 2005). Similar high levels of sensitivity during development have also been shown in the auditory and visual systems of owls (Knudson et al. 1984, Knudson and Knudson 1989). However, it was found that adults housed in environments with rich sensory stimulation were able to overcome behavioural deficits from experimental manipulations during the early sensitive period (Brainard and Knudsen 1998). Specifically, the window of time when owls were able to recover from the experimental manipulation of visual input was longer when animals were housed in visually rich environments compared to visually impoverished environments. This highlights how external factors can influence sensitive periods and modulate when behavioural plasticity can occur (Knudsen 2004).

Sensitive periods for learning have also long been studied for communicative behaviors including human language and birdsong. Children are abundantly better language learners than adults (reviewed in Newport 1990), and recent research has begun to reveal possible neural changes responsible for this difference, as well as the timelines for distinct sensitive periods related to different aspects of language (reviewed in Kuhl 2010). Similarly, many species of songbird only learn their songs during distinct sensitive periods (Konishi 1985; Marler 1970a; Mooney 1999). In species like the zebra finch and Bengalese finch (those studied in this thesis), song learning begins with a 'sensory phase' early in development when a young bird memorizes the song of a tutor (usually father). This is followed by a 'sensorimotor phase' of vocal exploration (akin to human babbling) and practice where the young bird compares his own song to the memorized template. Eventually the bird undergoes a 'crystallization phase' when the song sets in place and in which new song learning can no longer occur (Brainard and Doupe 2002; Doupe and Kuhl 1999).

After such sensitive periods close, many behaviours can still undergo various types of modulation and plasticity. For example, speech and birdsong are modulated depending on the social environment. The field of sociolinguistics is dedicated to uncovering the numerous social factors that influence our speech (e.g., Labov 1972), and the concepts of style and style-shifting refer to how an individual will tailor their speech patterns in response to the external social context (e.g., Eckert and Rickford 2001). Similarly, songbirds like the zebra finch modulate their song during a courtship context (Sossinka and Böhner 1980) and these modulations can have important influences on others (Riebel 2009; Woolley and Doupe 2008). These examples of the regulation of social behaviour are part of a broader 'social competence' ability regarding how to display social behaviours depending on the social context to maximize Darwinian fitness (Taborsky and Oliveira 2012).

Experimental manipulations of sensory feedback have also revealed various forms of modulation and plasticity for behaviours beyond the sensitive period of learning. For example, manipulations of auditory feedback in adult humans and birds result in compensation of vocal outputs (Houde and Jordan 1998; Sober and Brainard 2009). Adult humans and songbirds also demonstrate deterioration of vocal signals after the loss of auditory feedback (Konishi 2004; Lane and Webster 1991; Nordeen and Nordeen 1992; Sakata and Brainard 2006; Woolley and Rubel 1997; Zevin et al. 2004). Further, targeted manipulations of auditory feedback or neural reward systems can drive adaptive changes to song structure, sequencing or timing in adult birds (Ali et al. 2013; Andalman and Fee 2009; Charlesworth et al. 2012; Hisey et al. 2018; Tumer and Brainard 2007; Warren et al. 2012; Xiao et al. 2018).

Finally, behaviours can undergo long-term plasticity that results in long lasting changes. For example, trained musicians and singers will continue to improve through practice, which is characterized by increases in the speed, consistency and accuracy of motor performance (Drake and Palmer 2000; Ericsson 2008; Keith and Ericsson 2007). Relatively less is known about natural changes to behaviours on a longitudinal basis, and the factors that predict and regulate such changes within individuals. In multiple species of birds, vocalizations tend to become more stereotyped as individuals mature during young adulthood (Botero et al. 2009; Kao and Brainard 2006; de Kort et al. 2009; Ota and Soma 2014; Pytte et al. 2007; Rivera-Gutierrez et al. 2010), which may relate to changes in the rate of neurogenesis (Pytte et al. 2007). Subsequently, variability tends to increase in aging (senescing) birds (Cooper et al. 2012; Rivera-Gutierrez et al. 2012), but the factors that could predict individual differences in these age-dependent changes and the degree to which such changes to song relate to other forms of song modulation remain poorly understood. In our first three chapters, we document such longitudinal changes to song in two species of songbirds, and relate these changes to social context-dependent song modulation.

Biases in vocal learning

Internal biases or 'filters' can influence how vocal behaviours are learned. As mentioned earlier, songbirds are often biased to learn conspecific over heterospecific song, and can also exhibit constraints for species-specific sequences of song elements (Clayton 1988; Gardner et al. 2005; Soha and Marler 2001). While some species of songbird, such as song sparrows, learn a relatively fixed, species-specific sequence of song elements, other species like the zebra finch and Bengalese finch seem able to produce many (if not all) possible sequences of conspecific song elements (Brenowitz and Beecher 2005; Lipkind et al. 2013; Zann 1996). Despite this flexibility in songbirds like zebra finches, some common sequence patterns have been observed in both wild and lab populations (Lachlan et al. 2016; Zann 1996). However, whether these common patterns are due to cultural propagation of song, or biological biases in learning is unclear.

Similar to these common patterns observed in zebra finches, linguists have long identified common patterns that occur across languages, and the contributions of cultural propagation relative to biases in learning have been discussed extensively (Bickel 2013; Chomsky 1980; Greenberg 1963; Yang et al. 2017). One strong piece of evidence for biases in language learning is the process of "creolization": children who are exposed to novel, simplified forms of languages (e.g., "pidgins") reshape and regularize this language into one that is more likely to contain common linguistic patterns (Senghas and Coppola 2001; Thomason and Kaufman 1988). Similarly, juvenile zebra finches that are tutored by conspecifics that produce acoustically atypical songs produce songs with more typical acoustic structure as adults (Fehér et al. 2009). This reshaping of the input during learning suggests that individuals are endowed with particular predispositions for a more species-typical learning outcome. In chapter four, we experimentally revealed specific sequencing biases in the zebra finch.

To fully understand such biases in learning, we must also uncover their biological origins. It is likely that a combination of auditory and motor biases in neural circuitry could act throughout the learning process. Indeed, there are many lines of evidence that the motor system influences and constrains the acquisition of speech, music and birdsong

(MacNeilage et al. 2000; Patel 2010; Podos 1996; Tierney et al. 2011; Zollinger and Suthers 2004). On the other hand, innate "tuning" for species-specific acoustic features has also been shown in many forms (Araki et al. 2016; Doupe and Kuhl 1999; Gomez et al. 2014; Marler and Peters 1977). For example, within human speech there are particular acoustic cues that appear prominently across languages which can help listeners segment the speech stream (Bion et al. 2011; Jusczyk et al. 1999). Altogether, these data demonstrate the complex interactions of neural systems that can lead to biases in learning. One key avenue of research to understand biases in learning is to experimentally reveal the distinct contributions of motor and auditory biases that can lead to common patterns in behaviour, and this is a process that we begin in chapter five as we uncover motor contributions to the patterns we observed in chapter four.

Birdsong as a model for human speech

As outlined in the previous sections, there are numerous similarities between human speech and birdsong. These similarities, and the experimental possibilities with birds, have led many to consider songbirds as an ideal model system for gaining insights into human speech (e.g., Doupe and Kuhl 1999; Marler 1970b).

Some of the first results to highlight these similarities include the existence of geographical variation in song (i.e., 'dialects'), and the fact that cross-fostered individuals learn the dialect of their foster parents, rather than their biological parents (Marler and Tamura 1962, 1964). In addition, recent research has shown that developing songbirds and humans both acquire sequence combinations in a similar stepwise manner across development (Lipkind et al. 2013). Also, similar to how humans modulate their speech depending on their audience (e.g., infants, other adults), zebra finches produce slightly different songs directed towards juveniles than to other types of audiences (Chen et al. 2016; Fernald et al. 1989; Werker et al. 1994). Finally, birds have been shown to possess some similar auditory predispositions as humans when perceiving and categorizing sound sequences (ten Cate and Spierings 2018).

Together, these results suggest that similar mechanisms may underlie vocal communication in birds and humans, and that experimental inquiries into birdsong that cannot be performed on humans may provide insight into human speech. For example, the debate as to the contributions of cultural propagation relative to individual learning biases for common patterns observed across languages is very difficult (maybe impossible) to experimentally and rigorously test in humans. However, with songbirds we have the opportunity to carefully control the auditory and social environments of developing songbirds to reveal internal biases in vocal learning, devoid of any cultural influences (see chapter four). Such studies in songbirds can reveal possible biological foundations for the types of learning biases that could exist for human speech acquisition.

Concluding remarks

In this thesis we have leveraged the experimental opportunities that birdsongs provide to investigate a complex, learned, and readily quantifiable measure of animal communication. In particular, we address how the acquisition and plasticity of birdsong is influenced by external factors and internal learning biases. In chapter one we investigated long-term song plasticity in the Bengalese finch, and in chapter two we identified how the social modulation of song can predict these long-term changes across individuals. For chapter three, we revisited questions in chapters one and two in a second songbird species, the zebra finch, to understand how “universal” these patterns might be across species. In chapter four, we conducted a large-scale tutoring experiment designed to uncover sequence learning biases in zebra finches. Finally, in chapter five we began to uncover the possible mechanisms (motor biases) behind the learning biases observed in chapter four.

References

Alcock J. *Animal Behavior*. 9th ed. Sunderland, MA: Sinauer Associates, Inc, 2009.

Ali F, **Otchy TM**, **Pehlevan C**, **Fantana AL**, **Burak Y**, **Ölveczky BP**. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80: 494–506, 2013.

Andalman AS, **Fee MS**. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci* 106: 12518–12523, 2009.

Araki M, **Bandi MM**, **Yazaki-Sugiyama Y**. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–1287, 2016.

Baptista LF, **Petrinovich L**. Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Anim Behav* 32: 172–181, 1984.

Bickel B. Distributional biases in language families. In: *Language typology and historical contingency*. 2013, p. 415–444.

Bion RAH, **Benavides-Varela S**, **Nespor M**. Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Lang Speech* 54: 123–140, 2011.

Bolhuis JJ, **Johnson MH**, **Horn G**. Effects of early experience on the development of filial preferences in the domestic chick. *Dev Psychobiol* 18: 299–308, 1985.

Botero CA, **Rossman RJ**, **Caro LM**, **Stenzler LM**, **Lovette IJ**, **de Kort SR**, **Vehrencamp SL**. Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Anim Behav* 77: 701–706, 2009.

Brainard MS, **Doupe AJ**. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, **Knudsen EI**. Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J Neurosci* 18: 3929–3942, 1998.

Brenowitz EA, **Beecher MD**. Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends Neurosci* 28: 127–132, 2005.

ten Cate C, **Spierings M**. Rules, rhythm and grouping: auditory pattern perception by birds. *Anim Behav* 1–9, 2018.

Charlesworth JD, **Warren TL**, **Brainard MS**. Covert skill learning in a cortical-basal ganglia circuit. *Nature* 486: 251–255, 2012.

Chen Y, **Matheson LE**, **Sakata JT**. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc Natl Acad Sci* 113: 6641–6646, 2016.

Chomsky N. Rules and representations. *Behav Brain Sci* 3: 1, 1980.

Clayton NS. Song tutor choice in zebra finches and Bengalese finches: the relative importance of visual and vocal cues. *Behaviour* 104: 281–299, 1988.

Clayton NS. The effects of cross-fostering on selective song learning in Estrildid finches. *Behaviour* 109: 163–174, 1989.

Cooper BG, **Méndez JM**, **Saar S**, **Whetstone AG**, **Meyers R**, **Goller F**. Age-related changes in the Bengalese finch song motor program. *Neurobiol Aging* 33: 564–568, 2012.

Doupe AJ, **Kuhl PK**. Birdsong and humman speech: Common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Drake C, **Palmer C**. Skill acquisition in music performance: relations between planning and temporal control. *Cognition* 74: 1–32, 2000.

Eckert P, **Rickford J**. *Style and sociolinguistic variation*. 2001.

Ericsson KA. Deliberate practice and acquisition of expert performance: A general overview. *Acad Emerg Med* 15: 988–994, 2008.

Fehér O, **Wang H**, **Saar S**, **Mitra PP**, **Tchernichovski O**. De novo establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–568, 2009.

Fernald A, **Taeschner T**, **Dunn J**, **Papousek M**, **de Boysson-Bardies B**, **Fukui I**. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J Child Lang* 16: 477, 1989.

Galef BG, **Laland KN**. Social learning in animals: empirical studies and theoretical models. *Bioscience* 55: 489, 2005.

Gardner TJ, **Naef F**, **Nottebohm F**. Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308: 1046–1049, 2005.

Gomez DM, **Berent I**, **Benavides-Varela S**, **Bion RAH**, **Cattarossi L**, **Nespor M**, **Mehler J**. Language universals at birth. *Proc Natl Acad Sci* 111: 5837–5841, 2014.

Greenberg J. Some universals of grammar with particular reference to the order of meaningful elements. *Universals Lang* 2: 73–113, 1963.

Hensch TK. Critical period plasticity in local cortical circuits. *Nat Rev Neurosci* 6: 877–888, 2005.

Hisey E, **Kearney MG**, **Mooney R**. A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat Neurosci* 21: 589–597, 2018.

Horn G, **McCabe BJ**. Predispositions and preferences. Effects on imprinting of lesions to the chick brain. *Anim Behav* 32: 288–292, 1984.

Houde JF, **Jordan MI**. Sensorimotor Adaptation in Speech Production. *Science* 279: 1213–1216, 1998.

Jusczyk PW, **Houston DM**, **Newsome M**. The beginnings of word segmentation in English-learning infants. *Cogn Psychol* 39: 159–207, 1999.

Kao MH, **Brainard MS**. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96: 1441–1455, 2006.

Keith N, **Ericsson KA**. A deliberate practice account of typing proficiency in everyday typists. *J Exp Psychol Appl* 13: 135–145, 2007.

Knudsen EI. Sensitive periods in the development of the brain and behavior. *J Cogn Neurosci* 16: 1412–1425, 2004.

Kolb B, **Gibb R**, **Robinson TE**. Brain plasticity and behavior. *Curr Dir Psychol Sci* 12: 1–5, 2003.

Konishi M. Birdsong: from behavior to neuron. *Annu Rev Neurosci* 8: 125–170, 1985.

Konishi M. The role of auditory feedback in birdsong. *Ann N Y Acad Sci* 1016: 463–475, 2004.

de Kort SR, **Eldermire ERB**, **Valderrama S**, **Botero CA**, **Vehrencamp SL**. Trill consistency is an age-related assessment signal in banded wrens. *Proc R Soc B Biol Sci* 276: 2315–2321, 2009.

Kuhl PK. Brain Mechanisms in Early Language Acquisition. *Neuron* 67: 713–727, 2010.

Labov W. *Sociolinguistic patterns*. University of Pennsylvania Press, 1972.

Lachlan RF, **van Heijningen CAA**, **ter Haar SM**, **ten Cate C**. Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front Psychol* 7: 980, 2016.

Laland K, **Galef B**. *The question of animal culture*. Harvard University Press, 2009.

Lane H, **Webster JW**. Speech deterioration in postlingually deafened adults. *J Acoust Soc Am* 89: 859–866, 1991.

Lipkind D, **Marcus GF**, **Bemis DK**, **Sasahara K**, **Jacoby N**, **Takahasi M**, **Suzuki K**, **Feher O**, **Ravbar P**, **Okanoya K**, **Tchernichovski O**. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013.

MacNeilage PF, **Davis BL**, **Kinney A**, **Matyear CL**. The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Dev* 71: 153–163, 2000.

Marler P. A comparative approach to vocal learning: Song development in white-crowned sparrows. *J Comp Physiol Psychol* 71: 1–25, 1970a.

Marler P. Birdsong and human speech: could there be parallels? *Am Sci* 58: 669–674, 1970b.

Marler P, **Peters S**. Selective vocal learning in a sparrow. *Science* 198: 519–521, 1977.

Marler P, **Tamura M**. Song "dialects" in three populations of white-crowned sparrows. *Condor* 64: 368–377, 1962.

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

Mooney R. Sensitive periods and circuits for learned birdsong. *Curr Opin Neurobiol* 9: 121–127, 1999.

Newport E. Maturational constraints on language learning. *Cogn Sci* 14: 11–28, 1990.

Nordeen KW, **Nordeen EJ**. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57: 58–66, 1992.

Ota N, **Soma M**. Age-dependent song changes in a closed-ended vocal learner: elevation of song performance after song crystallization. *J Avian Biol* 45: 566–573, 2014.

Patel A. *Music, language, and the brain*. Oxford, UK: Oxford University Press, 2010.

Podos J. Motor constraints on vocal development in a songbird. *Anim Behav* 51: 1061–1070, 1996.

Pytte CL, **Gerson M**, **Miller J**, **Kirn JR**. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Reader SM. Animal social learning: associations and adaptations. *F1000Research* 5: 2120, 2016.

Riebel K. Song and female mate choice in zebra finches: a review. In: *Advances in the Study* *of Behavior*. Elsevier Inc., 2009, 197–238.

Rivera-Gutierrez HF, **Pinxten R**, **Eens M**. Multiple signals for multiple messages: great tit, Parus major, song signals age and survival. *Anim Behav* 80: 451–459, 2010.

Rivera-Gutierrez HF, **Pinxten R**, **Eens M**. Tuning and fading voices in songbirds: age-dependent changes in two acoustic traits across the life span. *Anim Behav* 83: 1279–1283, 2012.

Sakata JT, **Brainard MS**. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006.

Schaefer H, **Hess H**. Color preferences in imprinting objects. *Z Tierpsychol* 16: 161–172, 1959.

Senghas A, **Coppola M**. Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychol Sci* 12: 323–328, 2001.

Sluckin W, **Salzen EA**. Imprinting and perceptual learning. *Q J Exp Psychol* 13: 65–77, 1961.

Sober SJ, **Brainard MS**. Adult birdsong is actively maintained by error correction. *Nat Neurosci* 12: 927–931, 2009.

Soha JA, **Marler P**. Vocal syntax development in the white-crowned sparrow (Zonotrichia leucophrys). *J Comp Psychol* 115: 172–180, 2001.

Sossinka R, **Böhner J**. Song types in the zebra finch. *Z Tierpsychol* 53: 123–132, 1980.

Taborsky B, **Oliveira RF**. Social competence: An evolutionary approach. *Trends Ecol Evol* 27: 679–688, 2012.

Thomason S, **Kaufman T**. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press, 1988.

Tierney AT, **Russo FA**, **Patel AD**. The motor origins of human and avian song structure. *Proc Natl Acad Sci* 108: 15510–15515, 2011.

Tomback DF, **Baker MC**. Assortative mating by white-crowned sparrows at song dialect boundaries. *Anim Behav* 32: 465–469, 1984.

Tumer EC, **Brainard MS**. Performance variability enables adaptive plasticity of ‘crystallized’ adult birdsong. *Nature* 450: 1240–1244, 2007.

Warren TL, **Charlesworth JD**, **Tumer EC**, **Brainard MS**. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32: 15414–15425, 2012.

Werker JF, **Pegg JE**, **McLeod PJ**. A cross-language investigation of infant preference for infant-directed communication. *Infant Behav Dev* 17: 323–333, 1994.

Woolley SC, **Doupe AJ**. Social context-induced song variation affects female behavior and gene expression. *PLoS Biol* 6: 0525–0537, 2008.

Woolley SMN, **Rubel EW**. Bengalese finches (*Lonchura striata domestica*) depend upon auditory feedback for the maintenance of adult song. *J Neurosci* 17: 6380–6390, 1997.

Xiao L, **Chattree G**, **Oscos FG**, **Cao M**, **Wanat MJ**, **Roberts TF**. A basal ganglia circuit sufficient to guide birdsong learning. *Neuron* 98: 208-221.e5, 2018.

Yang C, **Crain S**, **Berwick RC**, **Chomsky N**, **Bolhuis JJ**. The growth of language: Universal Grammar, experience, and principles of computation. *Neurosci Biobehav Rev*, 2017.

Zann RA. *The zebra finch: A synthesis of field and laboratory studies*. Oxford, UK: Oxford University Press, 1996.

Zevin JD, **Seidenberg MS**, **Bottjer SW**. Limits on reacquisition of song in adult zebra finches exposed to white noise. *J Neurosci* 24: 5849–5862, 2004.

Zollinger SA, **Suthers RA**. Motor mechanisms of a vocal mimic: implications for birdsong production. *Proc R Soc B Biol Sci* 271: 483–491, 2004.

Linking statement to chapter 1

Our first objective was to document longitudinal, age-dependent plasticity to song in adult songbirds, with particular interest in the complex sequencing produced by Bengalese finches. Because age-dependent changes had not been documented in this species, chapter 1 fills in this gap of knowledge by analyzing the songs of individual birds in a longitudinal manner.

Chapter 1

Vocal motor changes beyond the sensitive period for song plasticity

Logan S. James and Jon T. Sakata

Published: *Journal of Neurophysiology* 2014

Abstract

Behavior is critically shaped during sensitive periods in development. Birdsong is a learned vocal behavior that undergoes dramatic plasticity during a sensitive period of sensorimotor learning. During this period, juvenile songbirds engage in vocal practice to shape their vocalizations into relatively stereotyped songs. By the time songbirds reach adulthood, their songs are relatively stable and thought to be "crystallized." Recent studies, however, highlight the potential for adult song plasticity and suggest that adult song could naturally change over time. As such, we investigated the degree to which temporal and spectral features of song changed over time in adult Bengalese finches. We observed that the sequencing and timing of song syllables became more stereotyped over time. Increases in the stereotypy of syllable sequencing were due to the pruning of infrequently produced transitions and, to a lesser extent, increases in the prevalence of frequently produced transitions. Changes in song tempo were driven by decreases in the duration and variability of intersyllable gaps. In contrast to significant changes to temporal song features, we found little evidence that the spectral structure of adult song syllables changed over time. These data highlight differences in the degree to which temporal and spectral features of adult song change over time and support evidence for distinct mechanisms underlying the control of syllable sequencing, timing, and structure. Furthermore, the observed changes to temporal song features are consistent with a Hebbian framework of behavioral plasticity and support the notion that adult song should be considered a form of vocal practice.

Introduction

Sensitive periods regulate the extent of learning and plasticity in a variety of sensory and behavioral systems. For example, the visual system is critically shaped by environmental stimuli during a focal period in development (reviewed in Hensch 2005), and the acquisition of complex behaviors such as language and music are heightened during sensitive periods in human development (Kuhl 2010; Trainor 2005). Sensitive periods are characterized by enhanced levels of plasticity relative to periods preceding the opening and following the closing of the sensitive period (Erzurumlu and Gaspar 2012; Hensch 2005).

However, various degrees of plasticity have been noted beyond the closure of the sensitive period for sensory systems such as vision and complex behaviors such as language (Creutzfeldt and Heggelund 1975; Kuhl 2010).

Like speech and language, birdsong is a vocal communication signal that is learned during a sensitive period in development (Doupe and Kuhl 1999). As young fledglings, songbirds such as Bengalese finches memorize the songs of adult tutors and then engage in vocal practice to hone their initially noisy and variable vocalizations into song elements ("syllables") and sequences that resemble those memorized from adults (Brainard and Doupe 2000, 2002; reviewed in Doupe and Kuhl 1999; Immelmann 1969; Mooney 2009; Price 1979; Tchernichovski et al. 2001). To accurately develop a species-typical song, juvenile songbirds must hear and memorize song during a sensitive window in development ("sensory phase") and engage in vocal practice during a distinct sensitive period ("sensorimotor phase"; Brainard and Doupe 2000, 2002; Mooney 2009; Tschida and Mooney 2012). As juveniles practice their song during the sensorimotor phase, the sequencing, timing, and structure of song syllables become more stereotyped (reviewed in Doupe and Kuhl 1999; Glaze and Troyer 2013; Johnson et al. 2002; Mooney 2009; Ölveczky et al. 2005; Podos et al. 2009; Tchernichovski et al. 2001).

The sensitive period for song plasticity is thought to close around 3–4 mo of age, when songbirds reach sexual maturation (reviewed in Brainard and Doupe 2000, 2002). This notion is supported by the finding that the gross structure of adult song does not dramatically change over time (reviewed in Brainard and Doupe 2002). Furthermore, neural plasticity mechanisms are progressively lost across juvenile development. For example, dendritic spines in HVC (proper name) stabilize and spines in LMAN (lateral magnocellular nucleus of the anterior nidopallium) are pruned following juvenile song learning (Nixdorf-Bergweiler et al. 1995; Roberts et al. 2010). Furthermore, the expression of cellular mediators of plasticity such as NMDA receptors, CAMK2A, FOXP2, and MAP2K1 in brain nuclei important for song learning decline across juvenile development (Aamodt et al. 1992; Balmer et al. 2009; Haesler et al. 2004; Kato and Okanoya 2010).

However, recent experiments demonstrate that adult birdsong retains some plasticity. For example, targeted auditory manipulations drive adaptive plasticity in the sequencing, timing, and structure of song syllables (Ali et al. 2013; Andalman and Fee 2009; Charlesworth et al. 2011, 2012; Tumer and Brainard 2007; Warren et al. 2011, 2012). Furthermore, experimental manipulations of auditory feedback cause acute and plastic changes to adult song (e.g., Brainard and Doupe 2001; Osmanski and Dooling 2009; Sakata and Brainard 2006; Sober and Brainard 2009). These data indicate that vocal plasticity mechanisms persist in the adult songbird brain and have led to the suggestion that adult song should continue to represent a form of vocal practice (reviewed in Brainard and Doupe 2013).

Despite demonstrations that experimental manipulations can drive vocal plasticity, there is limited evidence that adult song naturally changes over time. Natural changes to vocal performance (e.g., increased song consistency) would complement experimental studies and support the notion that plasticity mechanisms are engaged during adult song production. To this end, we analyzed the degree to which the temporal and spectral structure of adult Bengalese finch song changes over time. Using a repeated-measures design, we compared how the sequencing, timing, and structure of song syllables changed across months and years in adult Bengalese finches. We explicitly tested the hypothesis that adult song, like juvenile song, becomes more stereotyped and consistent with age.

Methods

Animals and data collection

Bengalese finches were either born and raised in a breeding colony at the University of California, San Francisco (UCSF; n = 10) or purchased from vendors (Exotic Wings and Things, Ontario, Canada; n = 12). Birds purchased from outside vendors were 2–3 mo of age when shipped to McGill University. All birds were housed on a 14:10-h light-dark cycle with other birds in group cages and provided food and water ad libitum. All procedures were approved by the UCSF Institutional Animal Care and Use Committee or by the McGill

University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

For song recordings, birds were housed individually in sound-attenuating chambers (Acoustic Systems, Austin, TX; TRA Acoustics, Ontario, Canada). Song was recorded using an omnidirectional microphone (Countryman Associates, Menlo Park, CA) positioned above the male's cage. Computerized, song-activated recording systems were used to detect and digitize song (EvTAF, E. Tumer, UCSF, digitized at 32 kHz; or Sound Analysis Pro v.1.04, http://ofer.sci.ccny.cuny.edu/html/sound_analysis.html, digitized at 44.1 kHz). Recorded songs were digitally filtered at 0.3–8 kHz for off-line analysis using software custom written in the MATLAB programming language (The MathWorks, Natick, MA).

To analyze the natural vocal motor changes to adult birdsong, the songs of adult male Bengalese finches ($n = 22$) were recorded at two ages (younger adult vs. older adult) following sexual maturation. Younger adult recordings were conducted when males were, on average, 6 mo old (range: 4–11 mo). At this age, song is relatively stable compared with juvenile song, and new songs cannot be learned in normally reared birds (reviewed in Okanoya 2004). Older adult recordings were conducted when birds were, on average, 23 mo old (range: 11–55 mo). Recordings were separated, on average, by 17 mo (range: 6–47 mo). Between recording sessions, males were housed in either all-male cages ($n = 15$) or breeding cages ($n = 7$). The direction and magnitude of vocal motor changes were not affected by housing condition ($P > 0.05$ for all comparisons); therefore, all males were analyzed without regard to housing. At least 30 songs from random times across a single day were selected and analyzed at each age for all but 2 birds for which only 25 and 22 songs were recorded as older adults. All songs used for the analysis of age-dependent changes were spontaneous songs produced in isolation (“undirected song”).

Bengalese finches were, on average, ~2 yr old at the time of the older recording. Because Bengalese finches can live up to 10 yr, our older recordings were not of senescing Bengalese finches. A previous study demonstrated age-related declines in song performance in Bengalese finches that were >6 yr of age (Cooper et al. 2012).

Song analysis

Bengalese finch song consists of acoustic elements arranged in specific sequences. For purposes of description and analysis, we use the term "syllable" to refer to individual acoustic elements that are separated from each other by at least 5 ms of silence (e.g., Okanoya and Yamaguchi 1997). Labeling and analysis of syllables were identical to those of previously published studies (Sakata and Brainard 2006; Sakata et al. 2008; Warren et al. 2012). Briefly, following amplitude-based syllable segmentation (MATLAB), we labeled syllables manually based on visual inspection of spectrograms using custom-written MATLAB software. Bengalese finch song consists of syllables arranged in stereotyped and variable ("branch point") sequences (Fig. 1A) (Jin 2009; Okanoya 2004; Okanoya and Yamaguchi 1997; Sakata and Brainard 2006; Wohlgemuth et al. 2010; Woolley and Rubel 1997). An example of a branch point sequence is provided in Fig. 1, in which the sequence "abc" can be followed by either "d" or "q". This variability is not simply biological noise but reflects a controlled aspect of song (e.g., Sakata and Brainard 2006; Sakata et al. 2008; Warren et al. 2012).

To quantify age-dependent changes in syllable sequencing, we analyzed how sequence transition probabilities and transition entropies of branch point sequences changed with age. We analyzed the probability of different syllable transitions immediately following a specific sequence of syllables. Typically, there are two to five transitions at individual branch points. For each branch point, this sequence variability was quantified as the transition entropy:

$$\text{transition entropy}=\sum -p_i \cdot \log_2(p_i),$$

where the sum is over all possible transitions, and pi is the probability of the ith transition across all songs (Gentner and Hulse 1998; Gil and Slater 2000; Sakata and Brainard 2006; Sakata et al. 2008). Branch points with transitions that are more variable (i.e., closer to uniform probability) have higher transition entropy scores. Sequences in which the most prevalent (i.e., dominant) transition occurred >95% of the time were not considered branch points. Instances in which song was terminated immediately following the branch point were not included in the calculation of entropy. Only branch points that occurred at least 15 times were analyzed (range: 17–442; mean ± SE: 120.9 ± 5.2; e.g., Sakata and Brainard 2009). Age-dependent changes in syllable sequencing within stereotyped sequences were rarely observed and, hence, not analyzed.

We labeled songs from recordings of younger and older adults and identified branch point sequences for each time point independently. We independently identified branch point sequences in younger and older recordings so as not to bias our results (e.g., by analyzing age-dependent changes only in branch points identified from young adult song, one neglects the possibility that stereotyped sequences become variable). We also paid close attention to long range statistics (i.e., history dependence) in Bengalese finch song (e.g., Fujimoto et al. 2011; Katahira et al. 2011; Warren et al. 2012). For example, if a branch point sequence “cd” was preceded either by an “a” or a “b”, we examined whether transition probabilities were different across “acd” and “bcd” sequence contexts. If transition probabilities were significantly different across sequence contexts, we analyzed the sequences separately; otherwise, we pooled across sequence contexts.

Age-dependent changes in song tempo were analyzed using methods comparable to previous studies (Kao and Brainard 2006; Sakata and Brainard 2006, 2009; Sakata et al. 2008). Specifically, we identified sequences that were commonly produced at both time points and measured the duration from the onset of the first syllable of the sequence to the onset of the last syllable in the sequence (“sequence duration”). We computed the mean and variance of these sequence durations. Only one sequence was selected and analyzed per bird, and we used the interquartile range (IQR; the distance between the 25th and 75th percentiles) as our measure of variance because it is robust to outliers (e.g., Leblois et al. 2010; Samuels and Witmer 2002; Thompson et al. 2011). To compare song tempo across ages and recording conditions, we normalized waveforms within each recording. To normalize the data, we extracted the waveform of focal sequences then resampled (1 kHz), smoothed (5-ms square window), and rectified the waveform. Thereafter, we normalized the waveform between 0 and 1 for both the young and older adult recordings. After this normalization, the amplitude traces across ages were comparable, and we used a common threshold on these normalized traces to find syllable onsets and offsets at each age.

Changes to song tempo could be driven by changes to the duration of syllables, the duration of silent periods between syllables ("gaps"), or both (Andalman et al. 2011; Cooper and Goller 2006; Glaze and Troyer 2006, 2013; Long and Fee 2008; Thompson et al. 2011). To assess the degree to which syllables and gaps contribute to changes in song tempo, we computed the mean and IQR of syllable and gap durations for all syllables and gaps within a sequence. Specifically, for each rendition of the sequence, we summed the durations of all syllables and all gaps within the sequence and then computed the mean and IQR of these summed durations across all renditions of that sequence.

To analyze changes to syllable structure, we measured the fundamental frequency (FF) of syllables that had distinct and stable harmonic structure across renditions (see Kao and Brainard 2006; Kao et al. 2005; Sakata et al. 2008). For each syllable we calculated the autocorrelation of a segment of the sound waveform. The FF was defined as the distance (in Hz) between the zero-offset peak and the highest peak in the autocorrelation function. To improve the resolution of frequency estimates, we performed a parabolic interpolation of the peak of the autocorrelation function, which reduces the magnitude of error in FF measurements (de Cheveigné and Kawahara 2002). The use of parabolic interpolations of the autocorrelation function improves the accuracy of FF measurements, although this can be impacted when background noise is high or if the signal contains strong high-frequency components. Each example of a syllable was visually screened to ensure that only examples devoid of sound artifacts that could affect FF calculations (e.g., sound of movement, female calls in background) were used in the analysis. The FF of syllables varied from rendition to rendition, and we characterized this variation using the IQR (e.g., Leblois et al. 2010; Thompson et al. 2011).

To further investigate age-dependent changes to syllable structure, we analyzed changes in the mean and variability of seven other acoustic features for each syllable in the bird's repertoire (mean frequency, density of the spectral entropy, syllable duration, loudness entropy, frequency slope, amplitude slope, and spectrotemporal entropy: e.g., Sakata and Brainard 2006).

Statistical analyses

Analyses were done on a syllable or sequence basis using paired t-tests or mixed-effects models. We measured multiple distinct examples of song features (e.g., FFs of multiple syllables, transition entropies of multiple branch points) for a number of Bengalese finches. For paired t-tests, syllables or sequences were considered individually. For the mixed-effects models (e.g., Reed and Kaas 2010), independent variables were age, bird identification (ID), and syllable ID or sequence ID nested within bird ID. Bird ID, syllable ID, and sequence ID were random variables. This model statistically controls for the fact that multiple syllables or sequences from a single individual are present in the data set. The results from the paired t-test and mixed-effects model were comparable, and results from both tests are presented.

For all tests examining age-dependent changes in song, we set $\alpha = 0.05$ (2-tailed). Analyses were conducted using JMP 8 (SAS Institute, Cary, NC) and MATLAB.

Results

Stability of adult song over shorter timescales (1–2 mo)

The songs of juvenile songbirds change dramatically over the course of 1–2 mo. During juvenile song development, significant changes to the structure, timing, and sequencing of song syllables can occur within days, underscoring the heightened levels of plasticity during this period (Derégnaucourt et al. 2005; Lipkind et al. 2013; Tchernichovski et al. 2001). In contrast, studies of adult songbirds indicate that song organization changes very little over months for adult Bengalese finches (>3 mo old; e.g., Hampton et al. 2009; Okanoya and Yamaguchi 1997; Sakata and Brainard 2006; Warren et al. 2012; Woolley and Rubel 1997).

To assess the degree to which our population of adult Bengalese finches demonstrated such song stability, we compared the sequencing, timing, and structure of syllables over 1–2 mo when Bengalese finches were 4–6 mo old (n = 12 birds). Consistent with previous findings, we found that the organization of adult Bengalese finch song did not significantly change over 1–2 mo. For example, stereotyped sequences remained stereotyped, and branch point sequences continued to be variable in sequencing. To quantify and compare sequence variability at branch points, we computed the transition entropy of distinct branch point sequences (n = 45) and found no significant change across 1–2 mo (mixed-effects model: $F_{1,14.91} = 1.85$; $P = 0.1939$; paired t-test: $t_{44} = 1.50$, $P = 0.1412$). We also found that the duration of syllable sequences (paired t-test: $t_{11} = 1.44$, $P = 0.1783$), mean FF of syllables with flat harmonic structure (mixed-effects model: $F_{1,10.19} = 0.12$; $P = 0.7342$; paired t-test: $t_{26} = 0.45$, $P = 0.6577$), and the variability of FF did not significantly change over 1–2 mo (mixed-effects model: $F_{1,7.84} = 0.12$; $P = 0.7388$; paired t-test: $t_{26} = 0.30$, $P = 0.7665$). These lack of significant changes are not due to insufficient statistical power, because comparable sample sizes have been sufficient to detect other experimental effects on song (e.g., Sakata et al. 2008; Sakata and Brainard 2006; Hampton et al. 2009). These data confirm that the songs of our young adult Bengalese finches are stable across a few months.

Vocal motor changes across longer timescales: syllable repertoire and sequencing

Whereas the previous analysis suggests that adult song is stable, it is possible that adult song retains some plasticity after sexual maturation but that the rate of adult vocal motor change is relatively slow compared with the rate of juvenile vocal motor plasticity. Consequently, it is possible that significant changes to adult vocal performance could be observed across longer periods of time. To test this, we compared the songs of young adult Bengalese finches (4–11 mo of age) with the songs produced by these same males 6–47 mo later (mean interval: 17 mo).

Broadly speaking, as observed in previous studies (e.g., Okanoya and Yamaguchi 1997; Sakata and Brainard 2006; Warren et al. 2012; Woolley and Rubel 1997), the gross structure and organization of songs remained stable, even after years (Fig. 1, A and B). For example, the songs of individual males typically consisted of the same syllables across age: 96% (259/269) of the syllables observed in the songs of young adult males continued to be produced as older adults. Some syllables were “lost” or “gained” over time, but these represented very minor changes to syllable repertoire. The 10 syllables that were “pruned” from the songs of younger adults tended to be syllables that were infrequently produced when birds were young adults; on average, pruned syllables constituted 3.1 ± 0.9% (mean ± SE) of the young male's repertoire. Five syllables were found in the songs of older adult males but not in the songs produced as young adults, and these gained syllables represented 5.9 ± 3.3% of the older male's repertoire. Therefore, in the rare instances in which the syllable repertoire of a bird changed over time, the magnitude of change was generally small. As such, the syllable repertoire of an adult Bengalese finch is generally stable across years.

In contrast to the stability of adult syllable repertoire, the sequencing and timing of song syllables changed significantly over time. Whereas stereotyped sequences remained the same over time, syllable sequencing at branch points showed consistent changes from young adult to older adult recordings. In particular, syllable sequencing at branch points became more stereotyped as young adults aged into older adults. In the example depicted in Fig. 1, A–C, the branch point sequence “cd” was followed by “b” 11% of the time and by “e” 89% of the time when the bird was 5 mo old. However, 13 mo later, only the transition to “e” was produced following “cd”. In another bird, the branch point syllable “X” was followed by the syllable “a” 2% of the time, “b” 44% of the time, and “c” 54% of the time when the bird was 4 mo old (Fig. 1D). Twenty-five months later, the transitions to “b” and “c” were retained but the transition to “a” was no longer observed. For both of these examples, because the number of possible transitions decreased over time, syllable sequencing became more stereotyped and transition entropy decreased as the birds aged.

Decreases in transition entropy were common across branch point sequences. Of the 98 branch point sequences (in 22 males), 69 branch points (70%) showed decreases in transition entropy over time. Across all branch points, transition entropy decreased by ~30% from the young adult to older adult recordings (Fig. 2A; mixed-effects model: $F_{1,18.8}$ =

19.17, P = 0.0003; paired t-test: $t_{97} = 5.56$, P < 0.0001). This indicates that syllable sequencing of adult Bengalese finch song became more stereotyped over time.

To understand the nature of changes to syllable sequencing that caused the decrease in transition entropy, we examined specific changes to transition probabilities at branch points. Transition entropy is affected by the number of distinct transitions as well as the degree of skew in transition probabilities. As such, decreases in transition entropy could be driven by decreases in the number of possible transitions over time as well as increases in the degree of "dominance" of individual transitions. Such changes would be consistent with Hebbian models of synaptic and behavioral plasticity (Abbott and Nelson 2000; Dan and Poo 2004; Hebb 1949).

We first examined the degree to which the number of distinct transitions at branch point sequences changed over time. Overall, significantly fewer transitions were produced at branch points in the songs of older adults than younger adults (mixed-effects model: $F_{1,18.8} = 9.4$, P = 0.0063; paired t-test: $t_{97} = 6.67$, P < 0.0001). This indicates that transitions were pruned from individual branch points. Indeed, 66% of branch point sequences lost at least one transition. Most of the transitions that were lost were produced infrequently when birds were young adults (n = 89 transitions; median transition probability = 6.9%; Fig. 2B). For example, 61% of the transitions that were pruned as the birds aged had transition probabilities <10% when birds were young adults. A significantly smaller proportion of retained transitions had transition probabilities <10% as young adults (38%; Fig. 2C; likelihood ratio test, $\chi_{12} = 13.8$, P = 0.0002). Therefore, individual transitions were often pruned from young adult to older adult song, and these lost transitions tended to be infrequently produced when birds were young adults.

The pruning of transitions was not due to the pruning of syllables from the male's repertoire. As mentioned above, syllables were rarely lost over time. Furthermore, the loss of a syllable from the male's repertoire caused the loss of a transition in only 6.7% of the cases. In most cases, transitions that were pruned at a branch point consisted of syllables that were embedded in multiple sequences in the young adult song such that the loss of that syllable transition did not represent a complete loss of that syllable.

The pruning of sequence transitions played a significant role in the overall change in transition entropy. The change in transition entropy was significantly larger for branch point sequences with pruned transitions than for branch point sequences without pruned transitions (Fig. 2D; t-test: $t_{97} = 4.52$, $P < 0.0001$). Furthermore, the magnitude of change in transition entropy was significantly different from zero for branch points with at least one pruned transition (t-test: H_0: mean difference = 0; $t_{64} = 6.84$, $P < 0.0001$) but not for branch points without pruned transitions (t-test: H_0: mean difference = 0; $t_{32} = 0.26$, $P = 0.7972$).

Aside from decreases in the number of transitions, increases in the prevalence of dominant (i.e., the most prevalent) transitions could also lead to decreases in transition entropy. In its most extreme form, the dominant transition could become the sole transition, thereby eliminating sequence variability. There were 14 branch point sequences that became completely stereotyped over time, and in 12 of these 14 cases it was the dominant transition that became the sole transition (sign test: $P = 0.0129$). This finding is consistent with the notion that increases in the skew of transition probabilities contribute to the age-dependent change in sequence variability.

We next broadened the analysis to all branch point sequences, not just those wherein sequence variability was lost, and assessed the consistency and importance of changes in the transition probability of dominant transitions. Transition probabilities of dominant transitions did not consistently increase over time. In only 48% of the cases did the dominant transition increase over time, which was not significantly different from 50% (sign test: $P = 0.7620$). However, increases in the transition probabilities of dominant transitions did contribute to the decrease in transition entropy. Transition entropy decreased significantly more for branch point sequences in which the dominant transition increased in prevalence than for branch point sequences in which the dominant transition decreased in prevalence over time (Fig. 2E; t-test: $t_{97} = 4.32$, $P < 0.0001$). Indeed, the decrease in transition entropy over time was significant for branch point sequences in which the dominant transition increased in prevalence (t-test: H_0: mean = 0; $t_{46} = 9.64$, $P < 0.0001$) but not for branch point sequences in which the dominant transition decreased in prevalence (t-test: H_0: mean = 0; $t_{50} = 1.14$, $P = 0.2617$).

The pruning of transitions and increases in the transition probabilities of dominant transitions affected transition entropy in similar ways. To test whether transition pruning and increases in the prevalence of dominant transitions independently contributed to the change in transition entropy, we analyzed the association between transition pruning and increases in dominance. We found that transition pruning and increases in the transition probabilities of dominant transitions were not significantly associated with each other (likelihood ratio test: $\chi_{12} = 0.61$, $P = 0.4338$); for instance, branch points with at least one pruned transitions were equally likely to have dominant transitions that increased or decreased with time. Moreover, a two-way ANOVA with the presence of a pruned transition (yes or no) and the direction of change of the dominant transition (up or down) as independent factors confirmed that both transition pruning ($F_{1,94} = 19.8$, $P < 0.0001$) and increases in dominant transitions ($F_{1,94} = 18.5$, $P < 0.0001$) independently and additively contributed to the decrease in transition entropy. Branch points with both pruned transitions and increases in the transition probability of dominants had the largest decreases in transition entropy; branch points with either pruned transitions or increases in the prevalence dominant transitions had intermediate decreases in entropy; and branch point sequences with no pruned transitions and decreases in the transition probability of dominant transitions tended to show increases in transition entropy (Fig. 2F).

Taken together, these analyses suggest that syllable sequencing becomes more stereotyped over time and that the loss of infrequently produced transitions and, to a lesser extent, increases the dominance of dominant transitions independently contribute to the age-dependent increase in stereotypy.

Vocal motor changes across longer timescales: syllable timing

Song tempo becomes faster and more stereotyped over time in juvenile as well as adult zebra finches (Brainard and Doupe 2001; Glaze and Troyer 2013; Kao and Brainard 2006; Kojima and Doupe 2011). Consistent with these reports, we found that adult Bengalese finch song becomes faster and more stereotyped over time. For instance, in the example depicted in Fig. 3A, the duration of the sequence “cddee” decreased from 493 ms when the male was 5 mo old to 453 ms when he was 25 mo old. Generally speaking, the durations of frequently produced sequences decreased by 3.4% from the young adult recordings to older adult recordings (paired t-test: $t_{21} = 3.03$, $P = 0.0064$). Furthermore, the variability of sequence durations (IQR) was lower for the songs of older adults than for those of younger adults (paired t-test: $t_{21} = 1.78$, $P = 0.0904$).

We repeated the analysis of age-dependent tempo changes using just the first instance of the sequence within a song because song tempo has been found to decrease over the course of a song bout (Chi and Margoliash 2001; Cooper and Goller 2006; Glaze and Troyer 2006). Furthermore, because the duration of song bouts decreased with age (paired t-test: $t_{21} = 3.33$, $P = 0.0032$), age-dependent changes to song tempo could be driven by age-dependent changes in song duration. Consistent with the previous analysis, we found significant decreases in sequence durations (paired t-test: $t_{21} = 3.08$, $P = 0.0057$) and in the variability of sequence durations (paired t-test: IQR: $t_{21} = 2.11$, $P = 0.0474$) even when we restricted our analysis to the first sequence renditions within a song bout (Fig. 3, B and C).

Changes to sequence durations could be caused by changes to the durations of syllables, gaps, or both. To analyze the contribution of gaps and syllables to age-dependent tempo changes, we computed the duration of each gap and syllable of the first instance of the focal sequence within a song bout for young and older adults. Thereafter, we computed the mean duration for each of the gaps and each of the syllables within the sequence and then summed the mean durations for all gaps and syllables of the sequence, respectively, as our measures of gap durations and syllable durations for the sequence. We found that the duration of gaps but not syllables significantly decreased over time (Fig. 3D; paired t-test: $t_{21} = 2.86$, $P = 0.0093$). We also computed the IQR of the durations of each gap and each syllable in the focal sequence, and averaged the IQRs across all gaps and syllables in the sequence as our measures of variability for gap durations and syllable durations, respectively. We found that the variability of gap but not syllable durations significantly decreased over time (Fig. 3E; paired t-test: $t_{21} = 2.26$, $P = 0.0349$). These data underscore the importance of silent gaps in the changes in song timing and support previous findings that gaps are more pliable than syllables (Cooper and Goller 2006; Glaze and Troyer 2006, 2013).

Because variance scales with measures of central tendencies, decreases in the IQR of sequence and gap durations could be driven by decreases in the median durations of sequence and gap. To test this, we used a repeated-measures analysis of covariance (RM-ANCOVA) with the median duration as a covariate. [The median is used here because the IQR is a measure of quantiles around the median and because there were significant age-dependent decreases in the median of sequence durations (paired t-test: $t_{21} = 2.99$, $P = 0.0070$) and gap durations (paired t-test: $t_{21} = 3.02$, $P = 0.0064$)]. Median durations were significantly related to the IQR of durations (H_0: slope = 0; $P < 0.015$ for sequences and gaps), and, moreover, controlling for age-dependent changes in median durations, the IQR of sequence and gap durations continued to be lower for older adults than for younger adults (sequence: $F_{1,21} = 3.29$, $P = 0.0839$; gaps: $F_{1,21} = 3.66$, $P = 0.0695$). This result is consistent with the finding that the magnitude of changes in the IQR of sequence and gap durations (10.6% and 8.6%, respectively) exceeds the magnitude of changes in the median duration of sequences and gaps (2.8% and 5.4%, respectively). This additional analysis supports the notion that syllable timing becomes more consistent as birds age.

Vocal motor changes across longer timescales: syllable structure

Age-dependent decreases in the variability of syllable structure are a hallmark of juvenile song development (Brainard and Doupe 2000; Kojima and Doupe 2011; Ölveczky et al. 2005, 2011; Ravbar et al. 2012; reviewed in Sakata and Vehrencamp 2012; Tchernichovski et al. 2001). For example, syllable variability has been found to decrease by ~50% over 1–2 mo in developing juvenile zebra finches (Ölveczky et al. 2005). Similar age-dependent changes to syllable structure have been noted in young adult zebra finches, albeit to a lesser degree (Brainard and Doupe 2001; Kao and Brainard 2006; Pytte et al. 2007). To assess age-dependent changes in adult Bengalese finches, we measured the variability of the FF of syllables with flat harmonic structure (n = 51 syllables; e.g., Brainard and Doupe 2001; Kao and Brainard 2006). In contrast to findings in adult zebra finches, we observed only a nonsignificant trend for decreasing variability of FF with age in adult Bengalese finches (Fig. 4; IQR: mixed-effects model: $F_{1,18.4} = 1.29$, $P = 0.2710$; paired t-test: $t_{50} = 1.72$, $P = 0.0920$). This was true regardless of whether we used the IQR or the coefficient of variation to quantify variability. The lack of significance is not likely to be due to insufficient statistical power since previous studies using fewer syllables and songs have found robust experimental effects on the variability of FF (e.g., Hampton et al. 2009; Kao and Brainard 2006; Sakata et al. 2008; Stepanek and Doupe 2010).

Spectral variability of syllables with flat harmonic structure has also been quantified using the spectral entropy (Stepanek and Doupe 2010). We assessed the degree to which this measure of spectral variability changed over time and, as in the previous analysis, found only a weak, nonsignificant effect of age (mixed-effects model: $F_{1,14.9} = 0.53$, $P = 0.4780$; paired t-test: $t_{50} = 1.73$, $P = 0.0904$).

Finally, to expand our analyses, we also computed other syllable features (mean frequency, density of the spectral entropy, syllable duration, loudness entropy, frequency slope, amplitude slope, and spectrotemporal entropy) for all syllables in the males' repertoires. Again, we found no significant change in the mean and variability of these measures over time (mixed-effects models: $P > 0.05$ for all measures; n = 548 syllables). Consequently, despite robust age-dependent changes in the temporal organization of song, we found little evidence for age-dependent changes in the spectral features of song.

Temporal window of adult vocal motor changes

The boundaries of sensitive periods are determined by analyzing the degree of behavioral and neural plasticity at different ages (Hensch 2005). Understanding the degree to which the vocal motor changes we observed in the present study persist throughout life or change as birds grow older will provide insight into the temporal dynamics of vocal motor and neurobiological changes. To this end, we obtained a third song recording when Bengalese finches were even older than the previously analyzed older adult recordings for a subset of birds (n = 7). The average ages of birds at the three recording periods were 6 mo (range: 5–9 mo), 20 mo (range: 14–34 mo), and 36 mo (range: 28–55 mo), and we analyzed the magnitude of age-dependent changes in syllable sequencing at branch points and in song tempo.

For these birds, we continued to find that transition entropy changed over time (mixed-effects model: $F_{2,11.6} = 4.67$, $P = 0.0326$; n = 24 branch points). Post hoc contrasts revealed that transition entropy was significant lower when birds were both “older adults” (entropy: 0.7 ± 0.1, mean ± SE) and “even older adults” (0.7 ± 0.1) than when they were “young adults” (1.1 ± 0.1; Student's t-test: $P < 0.025$ for both contrasts). However, sequence variability at branch points was not significantly different between “older adults” and “even older adults” (Student's t-test: $P = 0.9446$). These differences in the degree of vocal motor change were not due to differences in the amount of time between recordings; the interval between the young and older adult recordings (13.6 ± 2.2 mo) was not significantly different from the interval between the older adult and even older adult recordings (16.3 ± 1.8 mo; paired t-test: $t_6 = 1.10$, $P = 0.3132$). This suggests that the magnitude of age-dependent changes to syllable sequencing decreases over time.

To obtain a finer estimate of the temporal window of vocal motor change, we next analyzed how the magnitude of sequence changes across time varied as a function of age. In Fig. 5A, we plot the weighted change in transition entropy of individual birds at specific ages. Specifically, we computed the change in entropy across consecutive recordings (i.e., young → older and older → even older recordings) for each branch point within a bird and then computed a weighted change in entropy for each individual bird (i.e., change in entropy weighted by the prevalence of the branch point). Thereafter, we plotted the weighted change in entropy following the young adult recording (filled circles) and following the older adult recording (open circles) for each bird as a function of the bird's age. Consistent with our previous analyses, the weighted change in entropy from the young adult (range: 5–9 mo) to the older adult recording (range: 14–34 mo) was significantly less than zero (t-test: H_0: mean = 0; $t_6 = 2.45$, $P = 0.0495$). However, the weighted change in entropy from the older adult recording to the even older adult recording (range: 28–55 mo) was not significantly different from zero (t-test: H_0: mean = 0; $t_6 = 0.31$, $P = 0.7697$). Furthermore, the weighted change from the young to older adult recording was significantly more negative than the weighted change from the older adult to even older adult recording (paired t-test: $t_6 = 2.70$, $P = 0.0356$). These analyses suggests that sometime between 9 and 14 mo of age the degree of age-dependent changes to syllable sequencing significantly decreases (Fig. 5A).

We repeated this set of analyses for song tempo. As per previous analyses, we focused on the first rendition of a focal sequence in a song bout to calculate sequence and gap durations. Age-dependent changes in song tempo were, overall, less robust than age-dependent changes in syllable sequencing, and as a consequence, for this smaller subset of males, we did not find a significant change in sequence or gap durations over the three time periods (RM-ANOVA: sequence duration: $F_{2,12} = 1.50$, $P = 0.2612$; gap durations: $F_{2,12} = 0.93$, $P = 0.4205$). However, age-dependent changes in song tempo generally resembled those for syllable sequencing. For example, the average proportional change in sequence duration from the young adult to the older adult recording (Fig. 5B, filled circles) was more negative than the average proportional change from the older adult to the even older adult recording (open circles; paired t-test: $t_6 = 0.69$, $P = 0.5182$). Similarly, the average proportional change in gap durations from the young adult to the older adult recording was more negative than the average proportional change from the older adult to the even older adult recording (paired t-test: $t_6 = 1.19$, $P = 0.2782$).

In addition to mean changes to song tempo, the variability of sequence and gap durations decreased as birds aged (Fig. 3). With this limited subset of males, we did not obtain significant effects of age on the IQR of sequence and gap durations (RM-ANOVA: IQR of sequence durations: $F_{2,12} = 1.14$, $P = 0.3534$; gap durations: $F_{2,12} = 3.19$, $P = 0.0776$). However, just as for mean values, the average proportional change in the IQR of sequence durations from the young adult to the older adult recording (filled circles) was more negative than the average proportional change from the older adult to the even older adult recording (Fig. 5C, open circles; paired t-test: $t_6 = 0.72$ $P = 0.5003$), and the average proportional change in the IQR of gap durations from the young adult to the older adult recording was more negative than the average proportional change from the older adult to the even older adult recording (paired t-test: $t_6 = 1.82$, $P = 0.1188$).

Discussion

Vocal motor changes beyond the sensitive period for vocal plasticity.

Birdsong is a vocal motor behavior that, like speech, is learned during a sensitive period in development. During the "sensorimotor phase," juvenile songbirds "babble" and engage in vocal motor practice, which leads to rapid increases in vocal consistency across development (Brainard and Doupe 2000, 2001; Catchpole and Slater 2008; reviewed in Doupe and Kuhl 1999; Mooney 2009; Tchernichovski et al. 2001). Whereas juvenile song is widely considered plastic, adult song has historically been considered as "crystallized" because the gross structure of song remains stable following sexual maturation (3–4 mo) and because neural plasticity mechanisms are progressively lost over development (Brainard and Doupe 2013; Mooney 2009; Sakata and Vehrencamp 2012). However, plasticity mechanisms are not absent in adult songbirds, and adult vocal performance could continue to change (e.g., become more stereotyped) over time. Consistent with this notion, we found that the sequencing and timing of song syllables naturally became more stereotyped as young adult Bengalese finches (~6 mo) aged into older adults.

The manner in which song elements (syllables) are sequenced became more stereotyped with age (Figs. 1 and 2). This increase in sequence stereotypy was caused by a reduction in the number of distinct transitions produced at branch point sequences as well as an increase in the skew and predictability of transition probabilities. In the majority of branch point sequences, at least one transition was "pruned," and in a handful of cases (14%), all but one of the transitions were pruned such that sequences became completely stereotyped. Furthermore, for half of the branch point sequences, the most frequently produced transition of the branch point (i.e., the dominant transitions) became more prevalent, which also contributed to age-dependent increases in sequence stereotypy for those cases. As a consequence, our data indicate the presence of neural mechanisms that drive changes to syllable sequencing over months in adult Bengalese finches.

The evidence for neural mechanisms that drive changes to syllable sequencing seems to contrast with a previous study revealing active mechanisms to maintain transition probabilities in adult Bengalese finches. Using a targeted reinforcement paradigm, Warren et al. (2012) experimentally drove changes to transition probabilities and then observed that transitions probabilities returned to baseline levels within days following the termination of reinforcement. Consequently, the authors concluded that there exist brain mechanisms that rapidly and actively maintain transition probabilities in adult Bengalese finches. Although the ages of birds between our study and that of Warren et al. (2012) seem comparable, a noticeable difference between the studies is the timescale of observed vocal motor changes: we observed changes to transition probabilities over months to years, whereas Warren et al. (2012) observed the reestablishment of transition probabilities over days. Therefore, it is possible that there exist distinct mechanisms that act on shorter and longer timescales: one mechanism could act on shorter timescales to maintain transition probabilities, whereas another mechanism could act on longer timescales to slowly drive changes to transition probabilities.

Despite the presence of mechanisms to increase sequence stereotypy, in most cases sequence variability was not completely eliminated. Indeed, sequence variability was maintained in 86% of the branch points analyzed and was present even in our oldest birds (>5 yr). This suggests that there also could exist mechanisms to actively maintain sequence variability. Results from a previous study (Warren et al. 2012) suggest a similar phenomenon because sequence variability returned after experimentally driving increases in sequence stereotypy. Such a mechanism could be functionally significant since studies in the Bengalese finch and related species suggest that sequence variability is ethologically important (reviewed in Okanoya 2012). Taken together, when our observations are combined with those of Warren et al. (2012), the data suggest the existence of multiple mechanisms in the brain that act at different timescales and in different directions.

In addition to changes to syllable sequencing, we observed age-dependent changes to the timing of syllable sequences. In particular, we found that song tempo and the stereotypy of song tempo increased with age (Fig. 3). These changes were due to a decrease in the mean and variability of intersyllable gap durations but not to syllable durations. Such changes are analogous to developmental changes in syllable timing and suggest a shared mechanism underlying temporal changes in juvenile and adult songbirds (Glaze and Troyer 2013). For example, because intersyllable gaps coincide with short inspirations during song (Aronov et al. 2011; reviewed in Goller and Cooper 2004; Veit et al. 2011), the shortening of gaps could indicate that birds are more efficient at replenishing their air supply with shorter inspirations.

The robust changes to the temporal structure of adult Bengalese finch song contrast with the lack of significant changes to the spectral structure of song with age. Whereas the structure of song syllables (e.g., fundamental frequency) becomes more consistent over weeks during juvenile development (e.g., Okanoya 1997; Tchernichovski et al. 2001), we did not observe significant changes to the stereotypy of spectral features over months in adult Bengalese finches. This lack of age-dependent increases in syllable consistency contrasts with observations of age-dependent increases in syllable consistency in adult zebra finches (Kao and Brainard 2006; Pytte et al. 2007). One potential explanation for this difference is that previous studies examined age-dependent changes in younger adult songbirds. For example, Kao and Brainard (2006) observed changes in the stereotypy of fundamental frequency in 3-mo-old zebra finches, whereas the average age of our young adults was 6 mo (range: 4–11 mo). The difference in findings could also reflect a species difference in the nature of adult vocal motor change. Regardless of the underlying cause, our data indicate that the spectral structure of adult Bengalese finch song changes less over time than the temporal structure of song. This finding is consistent with models of song learning suggesting that birds first master the structure of syllables and then master the sequencing of syllables (e.g., Tchernichovski et al. 2000; Troyer and Doupe 2000a, 2000b).

A finer grained analysis of the temporal window of adult vocal motor change suggests that the magnitude of age-dependent changes to syllable sequencing and timing decrease after 1 yr of age (Fig. 5). Specifically, we found that the degree to which syllable sequencing and timing changed over time was reduced when birds were >14 mo old relative to when they were 5–9 mo old. This was especially conspicuous for syllable sequencing (Fig. 5A). These data suggest that the temporal organization of song stabilizes after 1 yr of age. However, these analyses do not rule out the possibility that birds >1 yr of age retain the capacity for change but that the rate of change is dramatically reduced after 1 yr. Nevertheless, this analysis highlights the dynamic nature of vocal motor change in adult songbirds and suggests age-dependent changes in the neural control of song in adult songbirds.

Potential mechanisms of vocal motor change

Birdsong is controlled by discrete and specialized neural circuits (Brainard and Doupe 2013; Mooney 2009). For example, the sequencing and timing of syllables is regulated by the intrinsic and recurrent circuitry within and between HVC and RA (robust nucleus of the arcopallium; Brainard and Doupe 2013; Jin 2009; Mooney 2009; Mooney and Prather 2005; Prather 2013). Neural activity in HVC encodes syllable sequencing (Fujimoto et al. 2011; Hahnloser et al. 2002; Schmidt 2003; Vu et al. 1994), and locally cooling HVC slows the intrinsic dynamics of HVC activity and decreases song tempo (Andalman et al. 2011; Long and Fee 2008; Long et al. 2010). Perturbations of activity in HVC and RA lead to significant changes to syllable sequencing and timing (Ashmore et al. 2005; Vu et al. 1994; Wang et al. 2008). As such, variability in the sequencing and timing of syllables are likely to be caused by variability in the patterns of activity in HVC and RA. Moreover, age-dependent changes in the stereotypy of syllable sequencing and timing are likely to be driven by age-dependent changes in the patterns of neural activity in HVC and RA.

The anterior forebrain pathway (AFP) is a forebrain-basal ganglia-thalamic loop that is critical for various forms of song plasticity (reviewed in Brainard and Doupe 2013). We do not, however, hypothesize that the AFP plays a major role in the age-dependent changes observed here. Whereas the AFP has repeatedly been found to regulate the control and plasticity of syllable structure (Andalman and Fee 2009; Charlesworth et al. 2012; Kao and Brainard 2006; Kao et al. 2005; Warren et al. 2011), the AFP seems to play a lesser role in the control and plasticity of syllable sequencing and timing (e.g., Ali et al. 2013; Hampton et al. 2009; Stepanek and Doupe 2010). For example, lesions of LMAN, the output of the AFP, affect the variability of syllable structure but not the variability of syllable sequencing or song tempo in adult Bengalese finches (Hampton et al. 2009). Lesions of LMAN and the basal ganglia nucleus Area X prevent feedback-driven plasticity in syllable structure but not in the timing of syllables (Ali et al. 2013). Although there exists some evidence that the AFP can influence syllable sequencing and timing (e.g., Brainard and Doupe 2001; Goldberg and Fee 2011; Hamaguchi and Mooney 2012; Ölveczky et al. 2005; Scharff and Nottebohm 1991; Thompson et al. 2011), we argue that the AFP is not the primary locus of adult vocal change in Bengalese finches because studies documenting AFP influences on syllable sequencing and timing also observe AFP influences on syllable structure. For example, infusions of bicuculline into LMAN increase the variability of syllable sequencing as well as the spectral entropy of individual song syllables in adult zebra finches (Hamaguchi and Mooney 2012), and inactivations of LMAN decrease sequence variability as well as the variability of syllable structure in juvenile zebra finches (Ölveczky et al. 2005). Because we did not observe significant changes to the variability of syllable structure in adult Bengalese finches, our findings are inconsistent with the model that the AFP drives age-dependent changes to Bengalese finch song.

Age-dependent vocal motor changes could also be caused by changes in circuits outside HVC, RA, and the AFP. For example, it has been hypothesized that circuitry involving the shell region of LMAN and dorsal arcopallium (Ad) could contribute to the plasticity of syllable sequencing (Bottjer and Altenau 2010). Subsequent studies should investigate how activity in the shell region of LMAN and Ad influences syllable sequencing and timing in the Bengalese finch.

Regardless of the loci of neural changes, changes to syllable sequencing and timing suggest the possibility that plasticity mechanisms are engaged during song production in adult

Bengalese finches. Indeed, the increases in the stereotypy of syllable sequencing and timing are consistent with the notion that adult song, like juvenile song, represents a form of vocal practice. Generally speaking, practice leads to increases in the consistency with which motor acts are performed (Costa 2011; Davids et al. 2006; Ericsson and Lehmann 1996; Glaze and Troyer 2013; Madison et al. 2013; Magill 2004; Müller and Sternad 2009; Vihman 1996). For example, with practice, singers become more adept at "hitting" desired notes, drummers improve their ability to produce beats at a consistent tempo, and piano players make fewer errors in the sequencing and timing of notes (e.g., Ericsson et al. 1993; Konczak et al. 2009, Sakai et al. 2004). Experimental studies in adult songbirds indicate that manipulations of auditory feedback can activate plasticity mechanisms to drive vocal motor change but relatively little is known about the degree to which adult birdsong naturally changes over time in a manner consistent with the practice framework (reviewed in Brainard and Doupe 2013; Sakata and Vehrencamp 2012). Our finding that adult Bengalese finch song becomes more stereotyped over time is consistent with this framework and supports the hypothesis that plasticity mechanisms could be naturally engaged during adult song production.

Hebbian mechanisms can explain various forms of behavioral and synaptic plasticity, including vocal motor plasticity in songbirds (Abbott and Nelson 2000; Dan and Poo 2004; Hanuschkin et al. 2013; Hebb 1949; Troyer and Doupe 2000a, 2000b). Models of song production hypothesize that syllable sequencing is regulated by functional connections between populations of neurons in vocal premotor regions (e.g., HVC) that encode song syllables (reviewed in Jin 2009; Katahira et al. 2011; Nishikawa and Okanoya 2006). Furthermore, these models propose that differences in transition probabilities at branch point sequences reflect differences in the strengths of functional connectivity between populations of neurons: infrequently produced transitions reflect weak functional connections and dominant transitions represent strong functional connections. Hebbian processes could lead to a weakening of weak connections and the strengthening of strong connections between populations of neurons encoding syllables. In an extreme form, such plasticity could resemble "winner-take-all" models of synaptic development in which a single connection remains after the pruning of weaker synapses (Abbott and Nelson 2000;

Troyer and Doupe 2000a, 2000b). In our data, both the pruning of infrequently produced transitions and the examples in which the transition probability of dominant transitions increase over time are consistent with Hebbian models of plasticity. Furthermore, branch point sequences that became completely stereotyped over time are intriguing examples of a winner-take-all phenomenon because in most of these cases the dominant transition became the sole transition. However, in contrast to the predictions of a Hebbian framework, we did not observe an overall trend for dominant transitions to become more prevalent over time. Indeed, dominant transitions were equally likely to increase or decrease in prevalence, and in some instances, previously nondominant transitions became dominant over time (e.g., Fig. 1D). The pruning of infrequent transitions and changes to dominant transitions were largely independent, suggesting that distinct mechanisms could regulate distinct types of age-dependent changes to syllable sequencing.

Age-dependent changes to syllable timing are also consistent with a Hebbian model of adult song plasticity. With repeated singing, Hebbian mechanisms could lead to increases in the strengths of the functional connections between populations that encode adjacent song syllables (Jun and Jin 2007; Nishikawa and Okanoya 2006). As functional connections between populations of neurons that encode adjacent syllables become stronger, the interval between syllables should decrease. We observed such decreases to intersyllable gaps (Fig. 3), suggesting that Hebbian mechanisms could also underlie age-dependent changes in the timing of song syllables.

Independent of Hebbian processes, age-dependent changes in syllable sequencing and timing could also reflect age-dependent changes in the sensory control of song. Both the sequencing and timing of adult song are critically dependent on auditory feedback (reviewed in Brainard and Doupe 2000; Mooney 2009; Prather 2013; Tschida and Mooney 2012). In the Bengalese finch, deafening leads to changes in syllable sequencing within days, and acute perturbations of auditory feedback lead to real-time changes in the sequencing and timing of syllables (Okanoya and Yamaguchi 1997; Sakata and Brainard, 2006; Woolley and Rubel 1997). The rapidity with which deafening alters song has been found to decrease with age in a related songbird species, the zebra finch, suggesting that adult song becomes less dependent on auditory feedback as adult birds continue to master the performance of song (Brainard and Doupe 2001; Lombardino and Nottebohm 2000). As a consequence, age-dependent increases in sequence stereotypy and song tempo could similarly reflect a decreased dependence on auditory feedback in Bengalese finches. Consistent with this notion is the finding that perturbations of auditory feedback cause smaller changes to song when Bengalese finches produce the more stereotyped and faster courtship song than when they produce the more variable and slower non-courtship song (Sakata and Brainard 2009). Because syllable sequencing and timing tend to be more stable in adult Bengalese finches >1 yr of age, it would be interesting to assess whether perturbations of auditory feedback have more rapid and dramatic effects on the temporal organization of song in Bengalese finches <1 yr of age than in birds >1 yr old.

Acknowledgements

We thank L. E. Matheson for assistance with data collection and analysis and S. C. Woolley, M. S. Brainard, D. Mets, K. Bouchard, and D. Ostry for critical input on analysis, interpretation, and presentation.

References

Aamodt SM, Kozlowski MR, Nordeen EJ, Nordeen KW. Distribution and developmental change in [3H]MK-801 binding within zebra finch song nuclei. *J Neurobiol* 23: 997–1005, 1992.

Abbott LF, Nelson SB. Synaptic plasticity: taming the beast. *Nature Neurosci* 3 Suppl: 1178–1183, 2000.

Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80: 494–506, 2013.

Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci USA* 106: 12518–12523, 2009.

Andalman AS, Foerster JN, Fee MS. Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature. *PLoS One* 6: e25461, 2011.

Aronov D, Veit L, Goldberg JH, Fee MS. Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. *J Neurosci* 45: 16353–316368, 2011.

Ashmore RC, Wild JM, Schmidt MF. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 25: 8543–8554, 2005.

Balmer TS, Carels VM, Frisch JL, Nick TA. Modulation of perineuronal nets and parvalbumin with developmental song learning. *J Neurosci* 29:12878–12885, 2009.

Bottjer SW, Altenau B. Parallel pathways for vocal learning in basal ganglia of songbirds. *Nat Neurosci* 13: 153–155, 2010.

Brainard MS, Doupe AJ. Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci* 1: 31–40, 2000.

Brainard MS, Doupe AJ. Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J Neurosci* 21: 2501–2517, 2001.

Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36:489–517, 2013.

Catchpole CK, Slater PJB. *Bird song biological themes and variations, 2nd edn*. Cambridge: Cambridge University Press, 2008.

Charlesworth JD, Tumer EC, Warren TL, Brainard MS. Learning the microstructure of successful behavior. *Nat Neurosci* 14: 373–380, 2011.

Charlesworth JD, Warren TL, Brainard MS. Covert skill learning in a cortical-basal ganglia circuit. *Nature* 486: 251–255, 2012.

de Cheveigné A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. *J Acoust Soc Am* 111: 1917–1930, 2002.

Chi Z, Margoliash D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32: 899–910, 2001.

Cooper BG, Goller F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J Neurophysiol* 95: 3798–3809, 2006.

Cooper BG, Méndez JM, Saar S, Whetstone AG, Meyers R, Goller F. Age-related changes in the Bengalese finch song motor program. *Neurobiol Aging* 33: 564–568, 2012.

Costa RM. A selectionist account of de novo action learning. *Curr Opin Neurobiol* 21: 579–586, 2011.

Creutzfeldt OD, Heggelund P. Neural plasticity in visual cortex of adult cats after exposure to visual patterns. *Science* 188: 1025–1027, 1975.

Dan Y, Poo MM. Spike timing-dependent plasticity of neural circuits. *Neuron* 44: 23–30, 2004.

Davids K, Bennett S, Newell K. *Movement system variability*. Champaign, IL: Human Kinetics. 2006.

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O. How sleep affects the developmental learning of bird song. *Nature* 433: 710–716, 2005.

Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Pyschol Rev* 100: 363–406, 1993.

Ericsson KA, Lehmann AC. Expert and exceptional performance: evidence of maximal adaption to task constraints. *Annu Rev Psychol*, 47: 363–406, 1996.

Erzurumlu RS, Gaspar P. Development and critical period plasticity of the barrel cortex. *Eur J Neurosci* 35: 1540–1553, 2012.

Fujimoto H, Hasegawa T, Watanabe D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* 31: 10023–10033, 2011.

Gentner TQ, Hulse SH. Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Anim Behav* 56: 579–594, 1998.

Gil D, Slater PJB. Song organization and singing patterns of the willow warbler, *Phylloscopus trochilus. Behaviour* 137: 759–782, 2000.

Glaze CM, Troyer TW. Temporal structure in zebra finch song: implications for motor coding. *J Neurosci* 26: 991–1005, 2006.

Glaze CM, Troyer TW. Development of temporal structure in zebra finch song. *J Neurophysiol* 109: 1025–1035, 2013.

Goldberg JH, Fee MS. Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. *J Neurophysiol* 105: 2729–2739, 2011.

Goller F, Cooper BG. Peripheral motor dynamics of song production in the zebra finch. *Ann N Y Acad Sci* 1016: 130–152, 2004.

Grafton ST, Hazeltine E, Ivry RB. Motor sequence learning with the nondominant left hand. *Exp Brain Res* 146: 369–378, 2002.

Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C. FoxP2 expression in avian vocal learners and non-learners. *J Neurosci* 24: 3164–3175, 2004.

Hahnloser RH, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419: 65–70, 2002.

Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci* 32: 11671–11687, 2012.

Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J Neurophysiol* 101: 3235–3245, 2009.

Hanuschkin A, Ganguli S, Hahnloser RH. A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models. *Front Neural Circuits* 7: 106, 2013.

Hebb DO. *The organization of behavior: a neuropsychological theory*. New York: Wiley, 1949.

Hensch TK. Critical period plasticity in local cortical circuits. *Nat Rev Neurosci* 6: 877–888, 2005.

Immelmann K. Song development in the zebra finch and other estrilid finches. In: *Bird Vocalizations*, edited by Hinde RA. New York: Cambridge Univ. Press, 1969, 61–74.

Jin DZ. Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Phys Rev E Stat Nonlin Soft Matter Phys* 80: 051902, 2009.

Johnson F, Soderstrom K, Whitney O. Quantifying song bout production during zebra finch sensory-motor learning suggests a sensitive period for vocal practice. *Behav Brain Res* 131: 57–65, 2002.

Jun JK, Jin DZ. Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. *PLoS ONE* 2: e723, 2007.

Kao MH, Doupe AJ, Brainard MS. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433: 638–643, 2005.

Kao MH, Brainard MS. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96: 1441–1455, 2006.

Katahira K, Suzuki K, Okanoya K, Okada M. Complex sequencing rules of birdsong can be explained by simple hidden Markov processes. *PLoS One* 6: e24516, 2011.

Kato M, Okanoya K. Molecular characterization of the song control nucleus HVC in Bengalese finch brain. *Brain Res* 1360: 56–76, 2010.

Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci USA* 108: 1687–1692, 2011.

Konczak J, van der Velden H, Jaeger L. Learning to play the violin: motor control by freezing, not freeing degrees of freedom. *J Motor Behav* 41: 243–252, 2009.

Kuhl PK. Brain mechanisms in early language acquisition. *Neuron* 67: 713–727.

Leblois A, Wendel BJ, Perkel DJ. Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *J Neurosci* 30: 5730–5743, 2010.

Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013.

Lombardino AJ, Nottebohm F. Age at deafening affects the stability of learned song in adult male zebra finches. *J Neurosci* 20: 5054–5064, 2000.

Long MA, Jin DZ, Fee MS. Support for a synaptic chain model of neuronal sequence generation. *Nature* 468: 394–399, 2010.

Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456: 189–194, 2008.

Madison G, Karampela O, Ullén F, Holm L. Effects of practice on variability in an isochronous serial interval production task: Asymptotical levels of tapping variability after training are similar to those of musicians. *Acta Psychologica* 143: 119–128, 2013.

Magill RA. *Motor learning and control: concepts and applications* (7th ed.). New York: McGraw-Hill, 2004.

Mooney R. Neural mechanisms for learned birdsong. *Learn Mem* 16: 655–669, 2009.

Mooney R, Prather JF. The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *J Neurosci* 8: 1952–1964, 2005.

Müller H, Sternad D. Motor learning: changes in the structure of variability in a redundant task In: *Progress in Motor Control*: Springer US, 2009, 439–456.

Nishikawa J, Okanoya K. Dynamical neural representation of song syntax in Bengalese Finch: a model study. *Ornitholog Science* 5: 95–103, 2006.

Nixdorf-Bergweiler BE, Wallhäusser-Franke E, DeVoogd TJ. Regressive development in neuronal structure during song learning in birds. *J Neurobiol* 27: 204–15, 1995.

Okanoya K. The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann N Y Acad Sci* 1016: 724–735, 2004.

Okanoya K. Behavioural Factors Governing Song Complexity in Bengalese Finches. *Intl J Comp Psychol* 25: 44–59, 2012.

Okanoya K, Yamaguchi A. Adult Bengalese finches (*Lonchura striata* var. *domestica*) require real time auditory feedback to produce normal song syntax. *J Neurobiol* 33: 343–356, 1997.

Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3: e153, 2005.

Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol* 106:386–397, 2011.

Osmanski MS, Dooling RJ. The effect of altered auditory feedback on control of vocal production in budgerigars (*Melopsittacus undulatus*). *J Acoust Soc Am* 126: 911–919, 2009.

Podos J, Lahti DC, Moseley DL. Vocal performance and sensorimotor learning in songbirds. *Adv Study Behav* 40: 159–195, 2009.

Prather JF. Auditory signal processing in communication: perception and performance of vocal sounds. *Hear Res* 305: 144–155, 2013.

Price PH. Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93: 260–277, 1979.

Pytte CL, Gerson M, Miller J, Kirn JR. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Ravbar P, Lipkind D, Parra LC, Tchernichovski O. Vocal exploration is locally regulated during song learning. *J Neurosci* 32: 3422–3432, 2012.

Reed JL, Kaas JH. Statistical analysis of large-scale neuronal recording data. *Neural Netw* 23: 673–684, 2010.

Roberts TF, Tschida KA, Klein ME, Mooney R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* 463: 948–952, 2010.

Sakai K, Hikosaka O, Nakamura K. Emergence of rhythm during motor learning. *Trends Cogn Sci* 8: 547–553, 2004.

Sakata JT, Brainard MS. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006.

Sakata JT, Brainard MS. Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *J Neurophysiol* 102: 2485–2497, 2009.

Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99: 1700–1711, 2008.

Sakata JT, Vehrencamp SL. Integrating perspectives on vocal performance and consistency. *J Exp Biol* 215: 201–209, 2012.

Samuels M, Witmer J. *Statistics for the Life Sciences.* 3rd edition. New Jersey: Prentice Hall. 2002.

Scharff C, Nottebohm F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11: 2896–2913, 1991.

Schmidt MF. Pattern of interhemispheric synchronization in HVc during singing correlates with key transitions in the song pattern. *J Neurophysiol* 90: 3931–3949, 2003.

Sober SJ, Brainard MS. Adult birdsong is actively maintained by error correction. *Nat Neurosci* 12: 927–931, 2009.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F. Dual pre-motor contribution to songbird syllable variation. *J Neurosci* 31: 322–330, 2011.

Trainor LJ. Are there critical periods for musical development? *Dev Psychobiol* 46: 262–278, 2005.

Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J Neurophysiol* 84: 1204–1223, 2000a.

Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning II. Temporal hierarchies and the learning of song sequence. *J Neurophysiol* 3: 1224–1239, 2000b.

Tschida K, Mooney R. The role of auditory feedback in vocal learning and maintenance. *Curr Opin Neurobiol* 22: 320–327, 2012.

Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* 450: 1240–1244, 2007.

Veit L, Aronov D, Fee MS. Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds. *J Neurophysiol* 106: 1747–1765, 2011.

Vihman MM. *Phonological development: the origins of language in the child. Applied language studies.* Oxford: Blackwell Publishing, 1996.

Vu ET, Mazurek ME, Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14: 6924–6934, 1994.

Wang CZ, Herbst JA, Keller GB, Hahnloser RH. Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biol* 6: e250, 2008.

Warren TL, Charlesworth JD, Tumer EC, Brainard MS. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32: 15414–15425, 2012.

Warren TL, Tumer EC, Charlesworth JD, Brainard MS. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J Neurophysiol* 106: 1806–1821, 2011.

Wohlgemuth MJ, Sober SJ, Brainard MS. Linked control of syllable sequence and phonology in birdsong. *J Neurosci* 30: 12936–12949, 2010.

Woolley SM, Rubel EW. Bengalese finches *Lonchura Striata domestica* depend upon auditory feedback for the maintenance of adult song *J Neurosci* 17: 6380–6390, 1997.

Figures

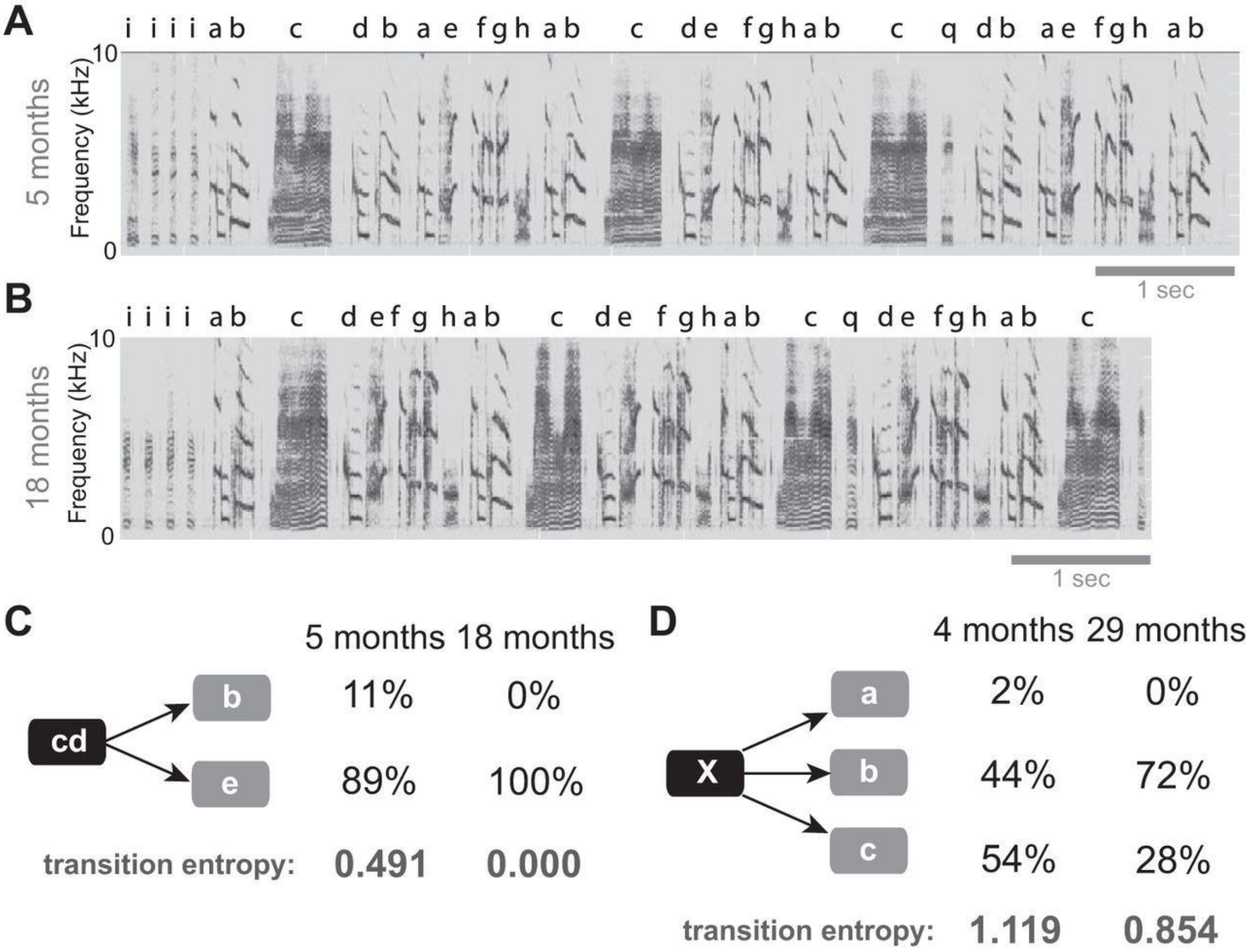


Figure 1. Structure of adult Bengalese finch song and sequence plasticity. (a). Spectrogram (time on the x-axis, frequency on the y-axis, brightness as amplitude) of a 5-month old Bengalese finch ('young adult'). Above the song are labels for each syllable of the song, and this labeling scheme is used for offline song analysis. As in other Bengalese finches, the bird's song consists of stereotyped and variable sequences of syllables. For example, the sequence 'fghabc' is stereotyped, and the sequences 'abc' and 'cd' are 'branch point sequences' that are followed by variable syllable transitions. The sequence 'abc' can be followed by 'd' or 'q', whereas the sequence 'cd' can be followed by a 'b' or 'e'. (b). The song of the same male 13 months later ('older adult'). Generally speaking, the gross structure and organization of the bird's song remains the same. For example, the sequence 'fghabc' remains stereotyped, and the 'abc' remains a branch point sequence followed by a 'd' or 'q'. However, in contrast to this bird's song at 5 months of age, at 18 months he only

produces ‘e’ after the sequence ‘cd’. (c). Transition probabilities and transition entropy of the branch point sequence ‘cd’ for the bird depicted in (a) and (b). (d). Transition probabilities and transition entropy of a branch point sequence from another bird. In this example, the syllable ‘X’ was followed by ‘a’, ‘b’, or ‘c’. Across the 25 months between recordings, the ‘a’ transition was pruned, the ‘b’ transition became more prevalent, and the ‘c’ decreased in prevalence; these changes in transition probabilities led to a decrease in transition entropy. Age-dependent pruning of infrequently produced transitions as depicted in (c) and (d) were common across birds.

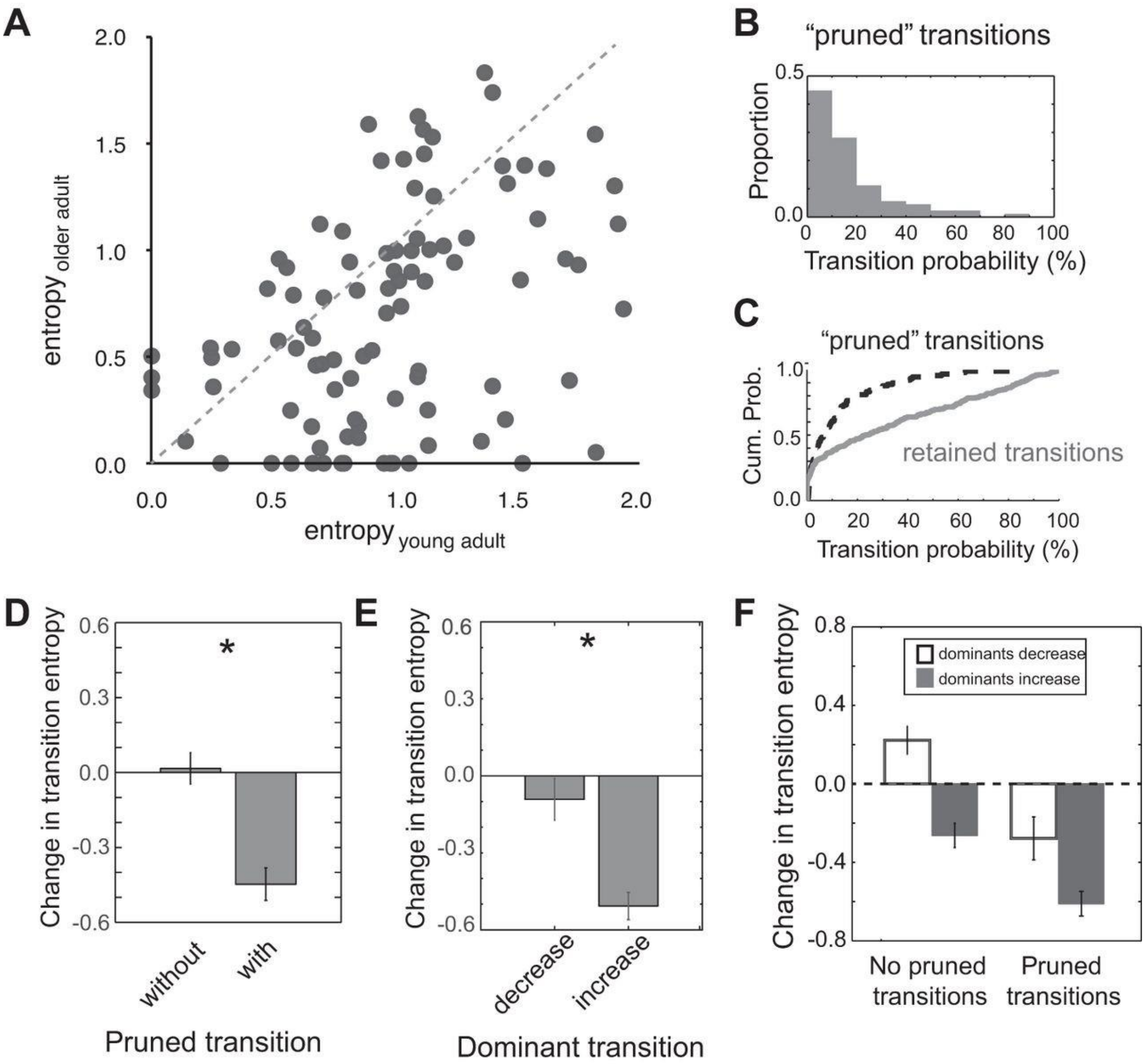


Figure 2. Age-dependent sequence plasticity at branch point sequences. A: transition entropy significantly decreases over time in adult Bengalese finches. Plotted are transition entropies, a measure of sequence variability, of individual branch point sequences when birds were young adults (x-axis) and older adults (y-axis). Most of the points lie below the line of unity (dashed line), indicating that transition entropy was lower when birds were older adults than when they were younger adults ($P < 0.001$). B: many transitions were "pruned" (i.e., observed in the younger adult recordings but not in the older adult recordings), and most of these transitions were infrequently produced when birds were younger adults. Plotted is a distribution of transitions probabilities (as young adults) of transitions that were pruned with age. C: pruned transitions were produced less frequently

than transitions that were retained over time. Plotted is a cumulative distribution of transition probabilities (Cum. Prob.; as young adults) of transitions that were pruned (dashed line) and retained (solid line) over time. D: the decrease in transition entropy was significantly greater for branch point sequences with at least 1 pruned transition than for branch point sequences without a pruned transition (*$P < 0.05$). The change in entropy was significantly different from 0 only for branch points with at least 1 pruned transition. E: the decrease in transition entropy was significantly greater for branch point sequences in which dominant transitions increased in prevalence than for branch point sequences in which dominant transitions decreased in prevalence (*$P < 0.05$). The change in entropy was significantly different from 0 only for branch points in which dominant transitions increased. F: the pruning of infrequently produced transitions and increases in the transition probabilities of dominant transitions independently contributed to the decrease in transition entropy with age. In a 2-way ANOVA with the presence of a pruned transition (yes or no) and the direction of change of the dominant transition (up or down) as independent factors, both transition pruning and increases in dominant transitions independently and additively contributed to a decrease in transition entropy over time.

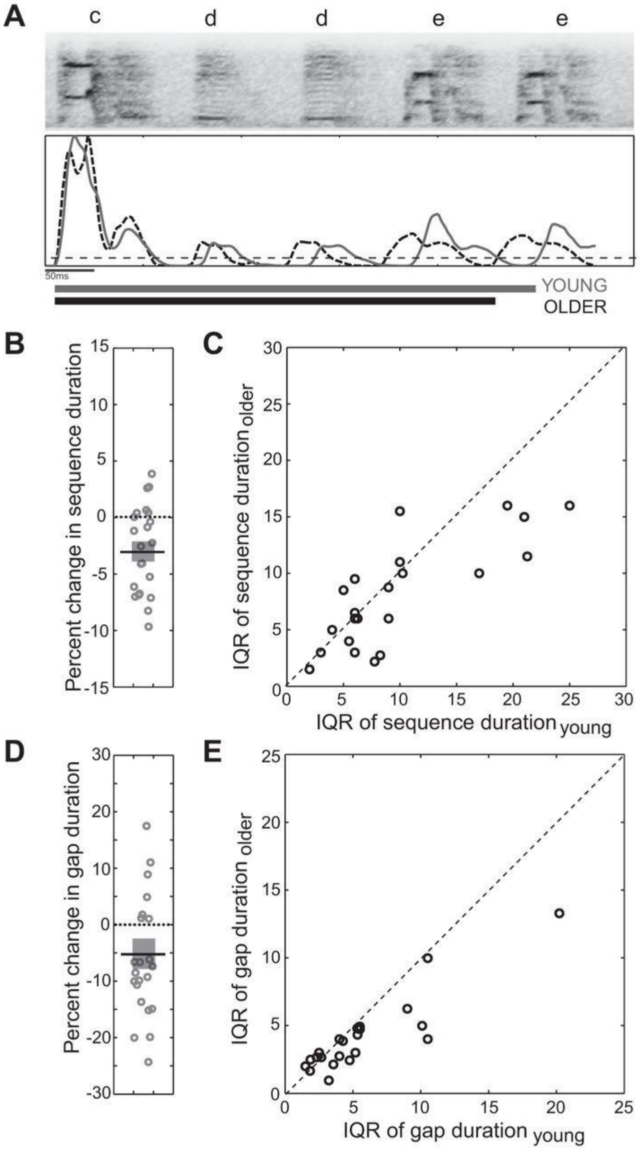


Figure 3. Age-dependent increase in song tempo and the stereotypy of song tempo. A: spectrogram (top) and mean rectified amplitude waveform (bottom) of the sequence "cddee" from a Bengalese finch. We measured sequence duration of "cddee" from the onset of "c" to the onset of the second "e". The waveform of the sequence produced as an older adult (dashed black line) was significantly shifted to the left relative to the waveform of the bird as a younger adult (solid gray line), indicating that sequence durations were shorter when the bird was older. The dashed horizontal line indicates the thresholds used to compute sequence duration, and the bars (bottom) indicate the mean sequence durations for this sequence at both ages. B: percent changes in sequence duration (first rendition of the sequence within a song bout) for each bird as birds aged. The average change was

negative and significantly different from 0, indicating that sequence durations decreased and song tempo increased. Solid line and shaded box represent mean change ± SE. C: temporal variability with which sequences were produced [i.e., interquartile range (IQR) of sequence duration for the first sequence rendition] significantly decreased with age ($P < 0.05$). D: decreases in sequence duration were driven by decreases in the duration of intersyllable gaps. Plotted are the percent changes in mean gap duration over time. The average change was negative and significantly different from 0. The duration of syllables did not change over time. Solid line and shaded box represent mean change ± SE. E: variability of gap durations also decreased with age ($P < 0.05$).

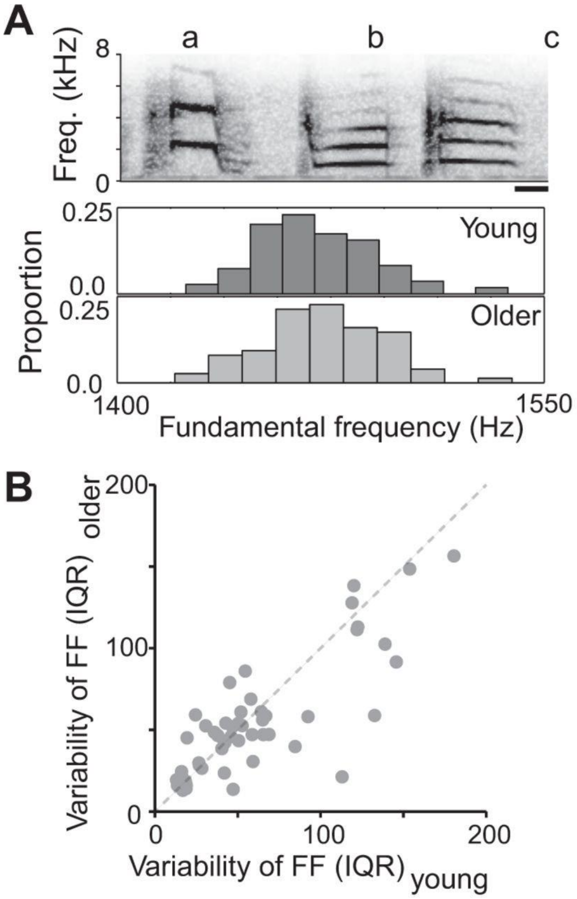


Figure 4. Lack of age-dependent changes in syllable structure. A: we measured the fundamental frequency (FF) of syllables with flat, harmonic structure (e.g., the syllable "b" in the spectrogram, top). For each syllable, we measured the FF on each syllable rendition and then plotted the distribution of FF measurements for each age. We computed the IQR of the distribution of FF as the measure of variability. In this example, there was little change in the variability of FF between young (middle) and older adult song (bottom). Scale bar, 50 ms. B: there was no significant change in the variability of syllable structure as adult Bengalese finches grew older, although there was a trend for the IQR of FF to decrease with age.

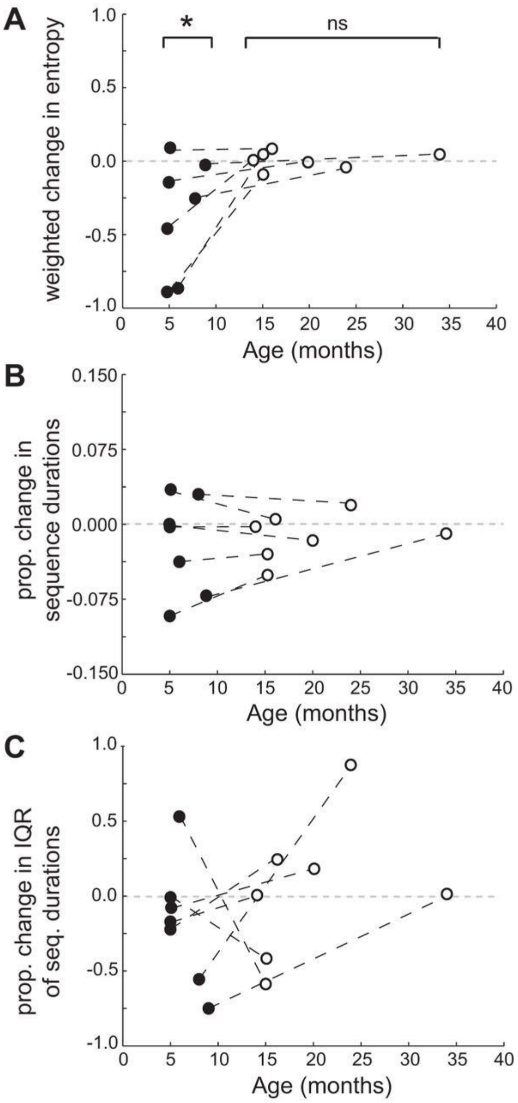


Figure 5. Temporal window of vocal motor changes. We analyzed the degree to which the variability of syllable sequencing (A), duration of sequences (B), and variability of sequence durations (C) changed in young adult and older birds (n = 7 birds). Plotted are the magnitudes of vocal motor changes from the young adult recording (mean = 6 mo) to the older adult recording (mean = 20 mo) as a function of the age of the bird at the young adult recording (filled circles; range: 5–9 mo). Also plotted are the magnitudes of vocal motor changes from the older adult to the even older adult recordings (mean = 36 mo) as a function of the age of the bird at the older adult recording (open circles; range: 14–34 mo). This allows us to visualize the degree of vocal motor change as a function of the age of the bird, and across A–C, there is a trend for the magnitude of vocal motor changes to approach zero in older adult birds. Data are the weighted change in transition entropy (A), the

proportional change in sequence durations (B), and the proportional change in the IQR of sequence durations (C) across 2 consecutive time points (young → older and older → even older). *P < 0.05, distribution is significantly less than 0.

Linking statement from chapter 1 to chapter 2

In the first chapter we identified age-dependent changes to Bengalese finch song sequencing and structure. From past literature, we noticed that many of the specific changes we found (e.g., more stereotyped sequencing, faster tempo) were similar to changes that have been observed when Bengalese finches sing towards females. Therefore, we sought to understand whether such social context-dependent changes to song could be linked to age-dependent changes. To this end, we analyzed these changes within individuals and asked whether context-dependent changes in young adult birds could predict age-dependent changes in those same birds.

Chapter 2

Predicting plasticity: acute context-dependent changes to vocal performance predict long-term age-dependent changes

Logan S. James and Jon T. Sakata

Published: *Journal of Neurophysiology* 2015

Abstract

Understanding the factors that predict and guide variation in behavioral change can lend insight into mechanisms of motor plasticity and individual differences in behavior. The performance of adult birdsong changes with age in a manner that is similar to rapid context-dependent changes to song. To reveal mechanisms of vocal plasticity, we analyzed the degree to which variation in the direction and magnitude of age-dependent changes to Bengalese finch song could be predicted by variation in context-dependent changes. Using a repeated-measures design, we found that variation in age-dependent changes to the timing, sequencing, and structure of vocal elements ("syllables") was significantly predicted by variation in context-dependent changes. In particular, the degree to which the duration of intersyllable gaps, syllable sequencing at branch points, and fundamental frequency of syllables within spontaneous [undirected (UD)] songs changed over time was correlated with the degree to which these features changed from UD song to female-directed (FD) song in young-adult finches (FDyoung). As such, the structure of some temporal features of UD songs converged over time onto the structure of FDyoung songs. This convergence suggested that the FDyoung song could serve as a stable target for vocal motor plasticity. Consequently, we analyzed the stability of FD song and found that the temporal structure of FD song changed significantly over time in a manner similar to UD song. Because FD song is considered a state of heightened performance, these data suggest that age-dependent changes could reflect practice-related improvements in vocal motor performance.

Introduction

Motor plasticity and learning are characterized by increases in the speed, consistency, and accuracy of motor performance and are driven by reinforcement, practice, and feedback-dependent mechanisms (Adams 1987; Doya 2000; Drake and Palmer 2000; Ericsson et al. 1993; Hickok and Poeppel 2007; Sanes and Donoghue 2000; Schmidt and Lee 1988; Shadmehr and Krakauer 2008; Ungerleider et al. 2002; Willingham 1998; Wolpert et al. 2003). Empirical and computational models of motor plasticity indicate that reinforcement and error signals based on sensory targets shape the performance of a variety of behaviors (Dayan and Cohen 2011; Magill 2004; Willingham 1998) and that as performance improves, the neural coding of these behaviors can become more efficient [e.g., Karni et al. (1995, 1998); Petersen et al. (1998); Poldrack (2000); Zatorre et al. (2007)]. These mechanisms continue operating to maintain or improve motor performance, even after individuals acquire expertise. For example, baseline and peak performances of trained musicians and singers continue to improve as they engage in practice (Drake and Palmer 1999; Ericsson 2008; Keith and Ericsson 2007). However, despite our understanding of fundamental mechanisms of motor learning and plasticity, relatively little is known about the factors that predict and regulate individual variation in motor plasticity [e.g., Golestani et al. (2002, 2007); Herholz and Zatorre (2012); Landi et al. (2011); Tomassini et al. (2011)]. Furthermore, our understanding of neural mechanisms of motor plasticity is based primarily on the performance of relatively simple behaviors, and little is known about mechanisms underlying plasticity in the performance of more complex and natural behaviors.

Birdsong is an important model system for understanding mechanisms that guide the plasticity of evolutionarily important behaviors, as well as individual variation in motor plasticity. Like other forms of motor learning and plasticity, the learning and development of birdsong involve the acquisition of a sensory target, motor practice, and changes in the speed and consistency of motor performance (Brainard and Doupe 2000, 2013; Doupe and Kuhl 1999; Mooney 2009). During a sensitive period in development, juvenile songbirds memorize the songs of adult tutors, and these memorized songs serve as sensory templates or targets that guide song development. Thereafter, juveniles engage in extensive vocal motor practice to refine their initially “noisy” and variable vocalizations into structured and stereotyped vocal elements (“syllables”) that resemble the sensory targets [e.g., Kelly and Sober (2014); Ölveczky et al. (2005); Tchernichovski et al. (2001)]. In addition to changes to syllable structure, song becomes faster, and syllable sequencing becomes more consistent over the course of juvenile song development (Doupe and Kuhl 1999; Glaze and Troyer 2006, 2013; Okanoya 1997; Troyer and Doupe 2000). By the time individual songbirds reach sexual maturity, their songs are relatively stable and similar to the songs memorized during development. Whereas such developmental changes are generally thought to be mediated by target-based plasticity mechanisms, changes to the speed and consistency of song are also observed in birds deprived of the opportunity to acquire a sensory target (Kojima and Doupe 2007; Livingston et al. 2000; Morrison and Nottebohm 2003). As such, there could be additional mechanisms that act in concert with target-based mechanisms to sculpt the development of song.

Despite the overall convergence of juvenile songs onto the structure of tutor songs, there exists substantial individual variation in song development (Catchpole and Slater 2008; Doupe and Kuhl 1999). The structure, timing, and sequencing of song elements can vary significantly among pupils that share a tutor, as well as between tutors and pupils, highlighting the importance of understanding individual variation in vocal motor plasticity. It has recently been found that the developmental trajectory of a juvenile's song could be predicted by analyzing acute context-dependent changes to song performance (Kojima and Doupe 2011). In particular, it was found that juvenile songbirds significantly change the structure of their songs when singing to females [female-directed (FD) songs] compared with when producing spontaneous songs in isolation [undirected (UD) songs] and moreover, that the structure of the juvenile's FD song was highly similar to the UD song that the individual ultimately produced as an adult. These data indicate that social context not only reveals a heightened level of vocal performance but also provides predictive insight into individual variation in vocal plasticity. Additionally, these data suggest that the understanding of mechanisms of social influences on vocal performance could reveal mechanisms of vocal plasticity.

Despite the relative stability of adult song structure, adult song continues to change in a manner that resembles vocal motor plasticity during development. Specifically, the spontaneous UD songs of adult Bengalese finches continue to become faster and more stereotyped in sequencing over time (James and Sakata 2014). The similarity in the nature of developmental and adult vocal motor plasticity suggests that similar mechanisms could guide vocal motor change in juvenile and adult songbirds (James and Sakata 2014; Kao and Brainard 2006; Kojima and Doupe 2011; Pytte et al. 2007; Sakata and Vehrencamp 2012). For example, just as social context predicts developmental song plasticity, context-dependent changes to adult song could similarly predict adult vocal plasticity. Indeed, just as UD song becomes faster and more stereotyped in sequencing over time, song acutely becomes faster and more stereotyped in sequencing when adult Bengalese finches produce FD song, indicating broad similarity in context- and age-dependent changes to song (Dunning et al. 2014; Hampton et al. 2009; Heinig et al. 2014; Sakata et al. 2008).

Here, we conducted a series of analyses to reveal potential mechanisms underlying adult vocal plasticity. We first used a repeated-measures design to analyze the degree to which variation in long-term age-dependent changes to adult Bengalese finch song (i.e., changes to UD song over time) could be predicted by understanding variation in acute context-dependent changes to song when birds were young adults (i.e., changes from UD_{young} to FD_{young} song). Thereafter, we proposed two models to explain the nature of age-dependent changes to UD song and tested these models by independently analyzing how age affected the structure of FD song.

Materials and methods

Animals and data collection

Bengalese finches (2–3 mo old) were purchased from vendors (Exotic Wings & Pet Things, Ontario, Canada) and shipped to McGill University. All birds were housed on a 14-h light:10-h dark cycle with other birds and provided food and water ad libitum. All procedures were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

For all song recordings, birds were housed individually in sound-attenuating chambers (TRA Acoustics, Ontario, Canada). Song was recorded using an omnidirectional microphone (Countryman Associates, Menlo Park, CA) positioned above the male's cage. Computerized, song-activated recording systems were used to detect and digitize song (Sound Analysis Pro, v 1.04, digitized at 44.1 kHz; http://ofer.sci.ccny.cuny.edu/sound_analysis_pro). Recorded songs were digitally filtered (0.3–8 kHz) for offline analysis using software custom written in the Matlab programming language (MathWorks, Natick, MA).

To assess context-dependent changes to song, we collected renditions of UD and FD song of adult Bengalese finches following a protocol similar to Sakata et al. (2008). UD songs are spontaneously produced when birds are alone, whereas FD songs are produced during courtship interactions with females. Males were moved into a sound-attenuating chamber at least one night before collecting UD and FD song. To collect FD songs, we placed a cage with a female next to the experimental male's cage and monitored his behavior via video. FD songs are readily distinguishable from UD songs because they are produced after a male approaches or faces a female, accompanied by a courtship dance (e.g., pivoting body from side to side) and associated with the fluffing of the male's plumage (Morris 1954; Zann 1996). Only songs that were accompanied by at least two of the above behaviors were categorized as FD songs. FD songs were almost always produced soon after the introduction of a female, and we removed females after 30–60 s to ensure that FD songs were produced at a short latency following exposure to a female. The modal interval between exposures to females was 4–5 min, which allowed for the collection of UD song between female exposures. However, not all birds produced renditions of UD song between female presentations; therefore, UD songs were also recorded for 30 min before and after the testing session. The number of UD [28.8 ± 1.7 (mean ± SE)] and FD (12.6 ± 0.8) songs collected was comparable with previous studies (Hampton et al. 2009; Heinig et al. 2014; James and Sakata 2014; Matheson et al. 2015; Sakata and Brainard 2009; Sakata et al. 2008).

We collected the UD and FD songs of young-adult Bengalese finches (UD_{young} and FD_{young} songs; n = 14 birds; 4–7 mo). We then returned males to their group cage (n = 10) or housed them with females to serve as breeders (n = 4) for 10–39 mo (21.0 ± 2.7 mo). Thereafter, we recorded the UD songs of these males as older adults (UD_{older} songs). This design allowed us to assess the degree to which context-dependent changes to song in young-adult Bengalese finches (i.e., from UD_{young} to FD_{young}) predicted age-dependent changes to UD song (from UD_{young} to UD_{older}) (James and Sakata 2014). Some of the older-adult recordings (n = 7) were included in a previous study (James and Sakata 2014). The magnitudes of context- and age-dependent changes to song were not significantly different between group-housed and breeding males; therefore, both were combined in the analysis.

Results from the analysis of context- and age-dependent changes suggested two models of vocal motor change, and the differentiation of these models required the analysis of the FD songs of older birds (FD_{older}). We were able to record the FD songs of a subset (n = 7) of Bengalese finches when they were older adults, allowing us to analyze subsequently how social context and age interacted to affect song performance using a factorial design.

Song analysis

Bengalese finch song consists of distinct acoustic elements arranged in both stereotyped and variable sequences. An example of a Bengalese finch song is provided in Fig. 1. For purposes of description and analysis, we use the term "syllable" to refer to individual acoustic elements that are separated from each other by >5 ms of silence (James and Sakata 2014; Okanoya and Yamaguchi 1997; Warren et al. 2012). Identical to previous studies, we manually labeled syllables based on visual inspection of spectrograms following amplitude-based syllable segmentation in Matlab [e.g., Heinig et al. (2014); James and Sakata (2014); Long and Fee (2008); Matheson et al. (2015); Sakata et al. (2008); Stepanek and Doupe (2010); Warren et al. (2012)].

Nodes in song with variable sequencing are called "branch points," and we analyzed context- and age-dependent changes to syllable sequencing at such branch points (n = 46 branch points in 14 males). For example, in Fig. 1, the branch point "cd" can be seen transitioning to both "b" and "e." Sequence variability at branch points is not simply biological noise but reflects a controlled aspect of song that is stable over days and weeks and modulated by social context (Hampton et al. 2009; Heinig et al. 2014; James and Sakata 2014; Matheson and Sakata 2015; Okanoya and Yamaguchi 1997; Sakata et al. 2008; Warren et al. 2012). Stereotyped and branch points were identified during manual labeling and confirmed using bigram plots [e.g., Fujimoto et al. (2011); Heinig et al. (2014); Kakishita et al. (2009); Matheson et al. (2015); Okanoya and Yamaguchi (1997)]. We analyzed the probability of different syllable transitions (typically two to five distinct transitions per branch point) immediately following a specific sequence of syllables, paying close attention to longer-range statistics in sequencing (Fujimoto et al. 2011; James and Sakata 2014; Matheson et al. 2015; Warren et al. 2012). Sequences were considered to be branch points if transition probabilities for all transitions were <95% under any experimental condition. We computed the transition entropy, a measure of variability, of each branch point using the following formula

$$\text{transition entropy} = \Sigma - p_i \times \log_2(p_i)$$

where the sum is over all transitions produced at the branch point, and pi is the probability of the ith transition across all renditions of the branch point. For our intents and purposes, the transition entropy indicates the number of bits of information required to summarize the extent of variation in syllable transitions. Branch points with transitions that are more variable (i.e., closer to uniform probability) have higher transition entropy scores. Entropy scores in our dataset ranged from 0 to 1.83 bits (sequences with entropy scores of 0 are included in the analysis because some branch points become completely stereotyped over time or across contexts). Only branch points that occurred at least 15 times during each recording session were analyzed (mean ± SE: 73.0 ± 4.4 renditions; range: 15–342). Instances in which song was terminated immediately following the branch point were not included in the calculation of entropy.

Changes to song tempo were analyzed using methods similar to previous studies (Cooper et al. 2012; James and Sakata 2014). Specifically, we identified a single, commonly produced sequence in an individual's songs (e.g., “mnop” in Fig. 1) and measured the duration of all syllables and intersyllable gaps within the sequence. Because the songs of young- and older-adult birds could have been acquired under different recording conditions, we compared normalized acoustic envelopes of syllable sequences to assess changes to song tempo. For this, we extracted the waveform of the sequence and then rectified, smoothed (5 ms square window), and resampled (1 kHz) the waveform. Thereafter, we normalized the envelope between zero and one using the minimum and peak values. After this normalization, the amplitude traces across ages and contexts were comparable, and we applied a common threshold on these normalized traces to find syllable onsets and offsets for each recording. This normalization allowed us to analyze age- and context-dependent changes to syllable and intersyllable gap durations, which can independently contribute to changes in sequence durations. We computed the mean duration of each syllable (range: 10–189 ms) and of each intersyllable gap within the sequence (range: 16–146 ms). We examined context- and age-dependent changes to song tempo by analyzing the mean duration of each individual syllable and gap, as well as the sum of all syllables and of all gaps within a sequence (i.e., total duration of syllables or total duration of gaps within a sequence). We also summed all mean syllable and gap durations together to compute sequence durations (i.e., interval from the onset of the first syllable to the onset of the last syllable of the sequence; range: 188–549 ms). Because sequence durations increase as the song progresses (Chi and Margoliash 2001; Cooper and Goller 2006; Glaze and Troyer 2006) and because song durations can change as a function of social context and age (James and Sakata 2014; Kao and Brainard 2006), we restricted our analysis to the first occurrence of the sequence in each song. Given the range of syllable, gap, and sequence durations, we also calculated the percent change in durations across social contexts and ages for data visualization and analysis.

To analyze changes to syllable structure, we analyzed the fundamental frequency (FF) of syllables that had distinct and stable harmonic structure (n = 33 syllables from 14 males; e.g., syllables “c” and “d” in Fig. 1). The FF of such syllables is tightly regulated by the nervous system and represents an important metric for song development and performance (Brainard and Doupe 2013; Kao and Brainard 2006; Sakata et al. 2008; Sakata and Vehrencamp 2012). To compute the FF, we calculated the autocorrelation of a segment of the sound waveform and defined the FF as the distance, in Hz, between the zero-offset peak and the highest peak in the autocorrelation function. To improve the resolution and accuracy of frequency estimates, we performed a parabolic interpolation of the peak of the autocorrelation function (de Cheveigné and Kawahara 2002). Each rendition of a syllable was visually screened to ensure that we analyzed only examples devoid of sound artifacts that could affect FF calculations [e.g., sound of movement, female calls in background; 105.1 ± 6.7 renditions (range: 18–317)]. For each syllable, we computed the mean and variability of FF across renditions, two aspects of song that change over development and across social contexts [e.g., Hampton et al. (2009); Kelly and Sober (2014); Sakata et al. (2008)]. The mean FF of syllables that we measured ranged from 0.4 to 4.5 kHz. Because of this wide range of values, we also calculated the percent change in mean FF across social contexts and ages for data visualization and analysis. We characterized the variability of FF across renditions using the interquartile region (IQR; distance between the 25th and 75th percentiles) divided by the median (50th percentile). This normalized IQR (normIQR; range: 0.0059–0.0784) is analogous to the coefficient of variation (SD/mean) in that it normalizes for differences in central tendencies (i.e., median or mean) but is more robust to outliers than other measures of variability (Samuels and Witmer 2002). Results were comparable regardless of whether we used the normIQR or coefficient of variation to characterize variability.

Statistical analyses

A central objective of this study was to assess the relationship between acute context-dependent changes (from UD_{young} to FD_{young} song) (Dunning et al. 2014; Hampton et al. 2009; Heinig et al. 2014; Sakata et al. 2008) and longer-term plasticity in the structure of UD song (UD_{young} to UD_{older}) (James and Sakata 2014). Therefore, we first analyzed variation in song structure among the UD_{young}, FD_{young}, and UD_{older} songs of individual birds. Because we measured multiple distinct examples of branch points and syllables (e.g., transition entropies of multiple branch points, the duration of multiple gaps within a sequence, the duration and FF of multiple syllables) within the songs of individual Bengalese finches across different social contexts and ages, we used mixed-effects models to analyze changes to transition entropy, syllable and gap durations, and FF. For these models, independent variables were GROUP (UD_{young}, FD_{young}, and UD_{older} songs); BIRD ID; and BRANCH POINT ID, GAP ID, or SYLLABLE ID, nested within BIRD ID. BIRD ID, SYLLABLE ID, GAP ID, and BRANCH POINT ID were random variables. Because only one sequence was measured per bird, we used a repeated-measures ANOVA to analyze context- and age-dependent changes to sequence durations. Tukey's honestly significant difference was used for post hoc comparisons. We used mixed-effects models (transition entropy, transition probabilities, syllable and gap durations, normIQR of FF) and Pearson's product moment correlation (sequence durations) to analyze the relationship between the magnitudes of context- and age-dependent changes to song. To analyze how social context and age interacted to influence song control for the subset of males in which UD and FD songs were collected when birds were young and older adults, we used a factorial design with CONTEXT, AGE, and CONTEXT × AGE as independent variables and BIRD ID and BRANCH POINT ID, GAP ID, or SYLLABLE ID, nested within BIRD ID, as random variables. Analyses were conducted using JMP 10 (SAS Institute, Cary, NC) and Matlab, and $\alpha = 0.05$ for all tests.

Results

Relationship between context- and age-dependent changes to syllable sequencing

The variability with which syllables are sequenced within spontaneously produced UD song generally decreases over time (James and Sakata 2014). However, there is considerable individual variation in the direction and magnitude of age-dependent changes to syllable sequencing. Age-dependent changes to syllable sequencing generally resemble context-dependent changes: sequence variability decreases significantly when adult Bengalese finches produce FD song relative to when they produce UD song (Hampton et al. 2009; Heinig et al. 2014; Matheson et al. 2015; Sakata et al. 2008). This suggests the possibility that context-dependent changes (UD_{young} to FD_{young}) could predict the direction and magnitude in which the sequencing of syllables within UD song changes over time (UD_{young} to UD_{older}). As such, we compared the UD_{young}, FD_{young}, and UD_{older} songs of individual Bengalese finches using a repeated-measures design.

We first compared the magnitude of age- and context-dependent changes to the transition entropy of branch points ($n = 46$ branch points from 14 birds). Transition entropy was significantly different among the UD_{young}, FD_{young}, and UD_{older} songs of individual birds (per branch point; $F_{2,90} = 8.4$, $P = 0.0005$). In particular, transition entropy was significantly lower for FD_{young} and UD_{older} songs than for UD_{young} song ($P < 0.025$ for both) and not significantly different between FD_{young} and UD_{older} songs (Fig. 2A; $P = 0.3896$). Consequently, the magnitude of context-dependent changes to the transition entropy of individual branch points (i.e., differences between UD_{young} and FD_{young} songs) was not significantly different than the magnitude of age-dependent changes (i.e., differences between UD_{young} and UD_{older} songs; per branch point; $F_{1,45} = 1.4$, $P = 0.2487$).

Similarities in the extent of context- and age-dependent changes to transition entropy suggest the possibility that individual variation in context-dependent changes to syllable sequencing could predict individual variation in age-dependent changes to syllable sequencing. To this end, we first analyzed the relationship between context- and age-dependent changes to transition entropy and found that the magnitude of acute context-dependent changes was significantly and positively correlated with the magnitude of long-term age-dependent changes (Fig. 2B; per branch point; $F_{1,43.2} = 4.3$, $P = 0.0430$). Branch points in which transition entropy changed more across social contexts also changed more over time.

The correlation between context- and age-dependent changes to transition entropy suggested the possibility that individual transition probabilities could change in similar manners across social context and age. However, because distinct types of changes to syllable sequencing could drive similar changes to entropy, it remained possible that context- and age-dependent changes to specific sequences could be independent. Therefore, we assessed the relationship between context- and age-dependent changes to the transition probabilities of all transitions across all branch points. At one branch point (Fig. 2C), social context and age had similar effects on transition probabilities: the probability of transitions from cd to "q" was virtually eliminated across context (from UD_{young} to FD_{young}) and age (from UD_{young} to UD_{older}), and the transition probabilities from cd to b and e increased across both context and age. For another branch point (Fig. 2D), the transition probabilities from "fa" to "g" and b were approximately equal when the bird produced UD_{young} song, and the transition probability to g increased across both social context and age. Such examples were common, and consequently, across all transitions in all branch points, there was a significant and positive relationship between the changes in transition probabilities caused by social context and by age (Fig. 2E; per transition; $F_{1,134.1} = 119.7$, $P < 0.0001$). These analyses highlight that the direction and magnitude of context-dependent changes to syllable sequencing significantly predicted the direction and magnitude of age-dependent changes.

Relationship between context- and age-dependent changes to syllable timing

Previous experiments have found that the FD songs of adult Bengalese finches are faster than their UD songs and that UD songs become faster with age (Dunning et al. 2014; Hampton et al. 2009; James and Sakata 2014; Matheson et al. 2015; Sakata and Brainard 2009; Sakata et al. 2008). We found similar patterns of change to song tempo, as depicted in the example in Fig. 3A. The average duration of the sequence "abcd" for an individual bird decreased from UD_{young} song [497.0 ± 1.2 (mean ± SE) ms] to FD_{young} song (495.1 ± 1.4 ms) and to UD_{older} song (483.6 ± 1.0 ms). Such changes were prevalent across birds, and consequently, we observed significant differences in sequence durations among UD_{young}, FD_{young}, and UD_{older} songs (per bird; $F_{2,26} = 5.3$, $P = 0.0114$). Sequence durations were shorter for UD_{older} than for UD_{young} songs ($P = 0.0083$) but not significantly different between UD_{young} and FD_{young} songs ($P = 0.2672$) or between FD_{young} and UD_{older} songs ($P = 0.2330$). The lack of significant difference between UD_{young} and FD_{young} songs observed in the post hoc contrasts was not consistent with previous studies. Therefore, we also computed the percent changes in sequence durations from UD_{young} to FD_{young} song for each bird and assessed whether the average change across social contexts was significantly different than zero (t-test; H_0: mean = 0). In contrast to the above analysis but consistent with previous studies (Dunning et al. 2014; Hampton et al. 2009; Matheson et al. 2015; Sakata and Brainard 2009; Sakata et al. 2008), we observed that the average change in sequence durations from UD_{young} to FD_{young} song was significantly less than zero, confirming that young-adult Bengalese finches produce faster songs when singing to females (−2.0 ± 0.8% change; per bird; $t_{13} = 2.6$, $P = 0.0201$). Analyses of the percent changes from UD_{young} to UD_{older} song and from FD_{young} to UD_{older} song were consistent with the previous analysis: the percent change in sequence durations from UD_{young} to UD_{older} was significantly less than zero (per bird; $P = 0.0044$), indicating that UD song became faster over time, and the percent change from FD_{young} to UD_{older} was not significantly different from zero (per bird; $P = 0.2883$), supporting the previous analysis that song tempo was similar between FD_{young} and UD_{older} songs. Given these patterns, it is not surprising that the magnitudes of context- and age-dependent changes to sequence durations were not statistically different (Fig. 3B; per bird; $F_{1,13} = 1.3$, $P = 0.2790$).

The duration of a sequence consists of the durations of multiple syllables and of multiple intersyllable gaps, and syllables and gaps could be differentially affected by social context and age [e.g., Cooper and Goller (2006); Glaze and Troyer (2006); James and Sakata (2014); Thompson et al. (2011)]. In Fig. 3A, we plot an example from an individual bird, and the total duration of gaps within this sequence decreased from UD_{young} song (113.6 ± 1.4 ms) to FD_{young} song (97.1 ± 2.1 ms) and to UD_{older} song (96.5 ± 1.6 ms). This pattern was consistent across birds, and consequently, we found that the total duration of gaps within individual sequences was significantly different among UD_{young}, FD_{young}, and UD_{older} songs (per bird; $F_{2,26} = 10.8$, $P = 0.0004$). Total gap durations were shorter for FD_{young} and UD_{older} songs than for UD_{young} song ($P < 0.0070$ for both), indicating a significant effect of age and context on gap durations, and not significantly different between UD_{older} and FD_{young} songs ($P = 0.5131$). Similarly, when we analyzed changes to each individual intersyllable gap, we found a significant difference among UD_{young}, FD_{young}, and UD_{older} songs (per gap; $F_{2,70} = 10.2$, $P < 0.0001$) that was driven by a significant decrease in FD_{young} and UD_{older} songs relative to UD_{young} song ($P < 0.01$ for both contrasts). As such, the magnitudes of context- and age-dependent changes to gap durations were not statistically different (Fig. 3B; per bird; $F_{1,13} = 0.5$, $P = 0.4907$).

In contrast to gap durations, the total duration of syllables within a sequence was not significantly different among UD_{young}, FD_{young}, and UD_{older} songs (per bird; $F_{2,26} = 1.7$, $P = 0.2075$). The variation in syllable duration was also not significant when each syllable was examined individually (per syllable; $F_{2,71} = 1.15$, $P = 0.3205$). Taken together, these analyses indicate that changes to sequence durations are driven primarily by changes to gap durations, not syllable durations.

We then analyzed the degree to which variation in context-dependent changes to gap, syllable, and sequence durations predicted variation in age-dependent changes. The magnitudes of context- and age-dependent changes to the total duration of gaps within a sequence were significantly and positively correlated (Fig. 3C; per bird; $r = 0.62$, $n = 14$, $P = 0.0174$): young adults that demonstrated larger context-dependent decreases in the total duration of intersyllable gaps within a sequence also demonstrated larger decreases in gap durations with age. This positive relationship was also observed when we analyzed each individual gap within a sequence, but the relationship was not statistically significant (per gap; $F_{1,20.4} = 2.04$, $P = 0.1686$). The magnitudes of context- and age-dependent changes to the total duration of syllables within a sequence (per bird; $r = 0.18$, $n = 14$, $P = 0.5492$), the duration of each syllable within a sequence (per syllable; $F_{1,31.7}$, $P = 0.2661$), and sequence durations were not significantly correlated (Fig. 3D; per bird; $r = -0.07$, $n = 14$, $P = 0.8141$). The lack of correlation at the sequence level, despite the correlation at the gap level, is likely due to the lack of correlation between context and age effects on syllable durations and the fact that syllables contribute more to sequence durations than gaps.

Taken together, these analyses highlight the importance of changes to gap durations for rapid context-dependent and longer-term age-dependent changes to song tempo. Gap durations (but not syllable durations) were affected by social context and by age, and the magnitude of context-dependent changes to gap durations (but not to syllable durations) was correlated with the magnitude of age-dependent changes.

Relationship between context- and age-dependent changes to syllable structure.

The FF of syllables with flat, harmonic structure is higher and less variable when male Bengalese finches produce FD song than when they produce UD song (Hampton et al. 2009; Matheson et al. 2015; Sakata et al. 2008). In contrast, the mean and variability of the FF of syllables for UD song do not change significantly with age (James and Sakata 2014). We observed similar context- but not age-dependent changes to syllable structure ($n = 33$ syllables in 14 males). The mean FF of syllables with flat, harmonic structure was significantly different among the UD_{young}, FD_{young}, and UD_{older} songs of individual birds (per syllable; $F_{2,64} = 3.7$, $P = 0.0304$). Mean FF was higher during FD_{young} song than during both UD_{young} and UD_{older} song, but the difference was only significant between FD_{young} and UD_{older}

song (P = 0.0228). The variability of FF (normIQR) was also significantly different among $\mathrm{UD}_{\mathrm{young}}$, $\mathrm{FD}_{\mathrm{young}}$, and $\mathrm{UD}_{\mathrm{older}}$ songs (Fig. 4A; per syllable; $F_{2,64} = 11.9$, $P < 0.0001$). Variability was significantly lower during $\mathrm{FD}_{\mathrm{young}}$ song than during $\mathrm{UD}_{\mathrm{young}}$ or $\mathrm{UD}_{\mathrm{older}}$ song ($P < 0.011$ for both) but not different between $\mathrm{UD}_{\mathrm{young}}$ and $\mathrm{UD}_{\mathrm{older}}$ songs (P = 0.1711).

Despite that social context and age differentially affected the mean and variability of FF, we found that variation in context-dependent changes to syllable structure tended to predict variation in age-dependent changes. For example, for one syllable, mean FF increased from $\mathrm{UD}_{\mathrm{young}}$ song (1,736 ± 5 Hz) to $\mathrm{FD}_{\mathrm{young}}$ song (1,748 ± 5 Hz) and to $\mathrm{UD}_{\mathrm{older}}$ song (1,782 ± 3 Hz), which amounted to a 0.7% increase across social contexts and a 2.6% increase with age (Fig. 4B). For another syllable, mean FF decreased from $\mathrm{UD}_{\mathrm{young}}$ song (1,185 ± 3 Hz) to $\mathrm{FD}_{\mathrm{young}}$ song (1,169 ± 4 Hz) and to $\mathrm{UD}_{\mathrm{older}}$ song (1,112 ± 3 Hz), amounting to a 1.4% decrease across social contexts and a 6.2% decrease with age (Fig. 4C). For both examples, the direction in which mean FF changed was similar across social contexts and age. As such, there was a significant and positive relationship between the magnitudes of context- and age-dependent changes to mean FF (Fig. 4D; per syllable; $F_{1,27.0} = 9.4$, $P = 0.0049$).

There was generally a linear relationship between context- and age-dependent changes to the variability of FF. For example, for the syllable summarized in Fig. 4B, the variability of FF decreased from $\mathrm{UD}_{\mathrm{young}}$ song (0.052) to $\mathrm{FD}_{\mathrm{young}}$ song (0.027) and to $\mathrm{UD}_{\mathrm{older}}$ song (0.028), which corresponded to a 48% and 46% decrease in the variability of FF, respectively, across social contexts and age. For the syllable summarized in Fig. 4C, the variability of FF decreased from $\mathrm{UD}_{\mathrm{young}}$ song (0.036) to $\mathrm{FD}_{\mathrm{young}}$ song (0.031) but increased from $\mathrm{UD}_{\mathrm{young}}$ song to $\mathrm{UD}_{\mathrm{older}}$ song (0.045), corresponding to a 15% decrease and 26% increase across social context and age, respectively. When examining all syllables, the magnitudes of context- and age-dependent changes to the variability of FF were not related significantly (Fig. 4E; per syllable; $F_{1,30.9} = 0.4$, $P = 0.5308$). However, there was a single outlying value, and this relationship was significant and positive when this outlier was removed from the analysis ($F_{1,22.5} = 5.1$, $P = 0.0342$). Generally speaking, large context-dependent decreases in the variability of FF were related to decreases in the variability of FF with age, whereas smaller context-dependent decreases or context-dependent increases were related to increases in the variability of FF with age.

Testing models of adult vocal motor change

The preceding analyses confirm our hypothesis that context-dependent vocal motor changes in young-adult Bengalese finches (UD_{young} to FD_{young}) lend predictive insight into age-dependent vocal motor changes to UD song (UD_{young} to UD_{older}). In addition, we observed that some temporal features of adult UD song converge over time onto the structure of FD_{young} song. Specifically, whereas UD_{young} song differed in many ways from FD_{young} song, there was no significant difference in sequence variability and syllable timing between UD_{older} and FD_{young} songs that these birds produced. We propose two models—a “target model” and a “performance model”—that could account for this convergence of UD song onto the FD song of young adults. The target model postulates that the FD_{young} song represents a stable “target” for age-dependent changes to UD song. The performance model proposes that context- and age-dependent changes reflect changes in vocal performance. It has been hypothesized that FD song represents a state of heightened performance, possibly indicative of an individual's current “best” performance (Byers et al. 2010; Dunning et al. 2014; Kao and Brainard 2006; Podos et al. 2009; Sakata et al. 2008; Sakata and Vehrencamp 2012; Woolley and Doupe 2008). Because the UD_{young} songs change toward the structure of FD_{young} songs over time, the performance model interprets age-dependent changes to song as increases in vocal performance, possibly due to practice. Whereas the two models are consistent in their predictions about age-dependent changes to UD song, the models differ in their predictions regarding age-dependent changes to FD song. The performance model predicts that the age-dependent changes to song are the result of an overall increase in vocal performance. Thus the age-dependent changes observed for UD songs should similarly be observed for FD songs: gap durations and sequence variability of FD songs should decrease over time (Fig. 5A). In contrast, the target model predicts that the FD_{young} song represents a stable target for age-dependent change, and thus FD song should not change across time (Fig. 5A). As such, the target but not performance model predicts a significant interaction between age and social context on transition entropy and gap durations, whereas the performance but not target model predicts significant and independent effects of age and context.

We were able to collect renditions of UD and FD songs from seven Bengalese finches when they were both young and older adults to test these models of vocal motor change. When simultaneously examining context- and age-dependent changes to sequence variability (n = 21 branch points in seven birds), we found a significant effect of age (per branch point; $F_{1,60} = 24.3$, $P < 0.0001$), a trend for social context ($F_{1,60} = 3.3$, $P = 0.0751$), and no significant interaction between age and social context on the transition entropy of individual branch points ($F_{1,60} = 0.6$, $P = 0.4372$). Importantly, post hoc contrasts revealed that the transition entropy of FD songs was significantly lower in older birds than in younger birds ($P = 0.0241$; Fig. 5B).

Similar effects of age and social context were observed for gap durations. Age and social context independently affected the duration of individual gaps within a sequence (per gap; age: $F_{1,51} = 34.4$, $P < 0.0001$; context: $F_{1,51} = 10.8$, $P = 0.0018$; age × context: $F_{1,18} = 1.3$, $P = 0.2564$) and the total duration of gaps within a sequence (per bird; age: $F_{1,18} = 13.8$, $P = 0.0016$; context: $F_{1,18} = 4.3$, $P = 0.0519$; age × context: $F_{1,18} = 0.5$, $P = 0.4765$). Importantly, post hoc contrasts revealed that the duration of individual gaps during FD song were significantly shorter in older adults than in younger adults (per gap; $P = 0.0085$; Fig. 5C).

Social context but not age or the interaction between age and context significantly affected the variability of FF (normIQR) of individual syllables (per syllable; age: $F_{1,45} = 0.3$, $P = 0.5793$; context: $F_{1,45} = 25.9$, $P < 0.0001$; age × context: $F_{1,45} = 0.7$, $P = 0.4216$). Within this subset of birds, the mean FF of individual syllables tended to increase from UD to FD song, but the effects of context, age, or the interaction between age and context were not significant (per syllable; age: $F_{1,45} = 0.2$, $P = 0.6448$; context: $F_{1,45} = 2.0$, $P = 0.1689$; age × context: $F_{1,45} = 0.0$, $P = 0.9597$). Consistent with the lack of age-dependent changes to the mean and variability of FF for UD songs (Fig. 4), post hoc contrasts reveal no difference in the variability of FF (per syllable; $P = 0.7680$; Fig. 5D) or mean FF (per syllable; $P = 0.9833$; Fig. 5E) between FD_{young} and FD_{older} songs.

Taken together, these data indicate that age and social context did not significantly interact to affect any song feature measured but that age and context independently affected sequence variability and gap durations. Of particular importance was that syllable sequencing and timing were significantly different between FD_{young} and FD_{older} songs. As such, temporal features changed in the same manner over time for FD song as for UD song, lending greater support for the performance than the target model of vocal motor change.

Discussion

The structure of evolutionarily important behaviors is not static but can change considerably over time (Catchpole and Slater 2008; Colonnese et al. 1996; Doupe and Kuhl 1999; Fentress 1992; Glaze and Troyer 2013; James and Sakata 2014; Sakata and Vehrencamp 2012). The understanding of the processes that guide and predict such behavioral changes can lend insight into mechanisms of motor plasticity and control as well as processes mediating individual differences in behavior [e.g., Kojima and Doupe (2011); Sakata and Crews (2003)]. Here, we demonstrate that understanding individual variation in the acute effects of social context on syllable sequencing, song tempo, and syllable structure lends predictive insight into variation in the direction and magnitude of age-dependent changes to syllable sequencing, timing, and structure. For example, when comparing the effect of social context (UD_{young} to FD_{young}) with the effect of age on UD song in these same individuals (UD_{young} to UD_{older} songs), we found that the direction and magnitude of context-dependent changes to the duration of silent gaps between syllables predicted the direction and magnitude of age-dependent changes to gap durations. Consequently, by examining individual variation in social context-dependent changes to adult Bengalese finch song, we can predict individual trajectories and magnitudes of age-dependent changes to some important features of song.

The correlations in the magnitude of age- and context-dependent changes to song suggest that neurophysiological changes that drive context-dependent changes to syllable sequencing, timing, and structure could resemble the neurophysiological changes that underlie age-dependent variation in syllable sequencing, timing, and structure. Birdsong is controlled primarily by two forebrain circuits. Neurons in the vocal motor pathway (VMP), which includes the HVC (used as proper name) and robust nucleus of the arcopallium (RA), encode the motor commands for song and are functionally analogous to neurons in the mammalian premotor, supplementary, and primary motor cortical areas (Doupe and Kuhl 1999; Fee and Scharff 2010). The activity of neurons in HVC affects and encodes information about syllable sequencing and timing (Andalman et al. 2011; Ashmore et al. 2005; Basista et al. 2014; Fujimoto et al. 2011; Hahnloser et al. 2002; Kosche et al. 2015; Long and Fee 2008; Long et al. 2010; Prather et al. 2008; Sakata and Brainard 2008; Schmidt 2003; Vu et al. 1994; Wang et al. 2008; Yu and Margoliash 1996). Therefore, we propose that context- and age-dependent changes to syllable sequencing at branch points and to gap durations are mediated by neurophysiological changes in HVC.

The control and plasticity of song are also regulated by activity in the anterior forebrain pathway (AFP), a circuit that is homologous to cortical-basal ganglia-thalamic loops in mammals and includes the basal ganglia nucleus Area X, the dorsal lateral nucleus of the anterior thalamus, and the lateral magnocellular nucleus of the anterior nidopallium (LMAN) (Brainard and Doupe 2013; Doupe et al. 2005; Reiner et al. 2004; Woolley and Kao 2015). For example, lesions of LMAN, the primary interface between the AFP and VMP, consistently affect the variability of syllable structure [reviewed in Woolley and Kao (2015)]. As such, social context and age could affect LMAN activity in similar ways to shape syllable structure. Manipulations of LMAN activity have also been found to affect song tempo persistently and to affect syllable sequencing transiently and persistently [e.g., Brainard and Doupe (2001); Hamaguchi and Mooney (2012); Kao and Brainard (2006); Kobayashi et al. (2001); Kubikova et al. (2014); Thompson et al. (2011); Williams and Mehta (1999)]. However, a number of studies do not observe significant AFP contributions to the sequencing and timing of song elements [e.g., Ali et al. (2013); Hampton et al. (2009); Leblois and Perkel (2012); Stepanek and Doupe (2010)]. Of particular relevance is that lesions of LMAN do not affect the social modulation of syllable sequencing and song tempo in Bengalese finches (Hampton et al. 2009). Consequently, we hypothesize that the AFP could mediate the covariation in age- and context-dependent changes to syllable structure but not the correlated changes to syllable sequence and timing for adult Bengalese finch song.

Given the similarity in context- and age-dependent changes to syllable timing and sequencing, we hypothesize that the pattern of neural activity across song control nuclei should be similar when older Bengalese finches produce UD song and when young-adult Bengalese finches produce FD song. Fewer neurons in both the VMP and AFP express the immediate early gene early growth response protein 1 (EGR-1) when adult Bengalese finches produce the faster and less variable FD song than when they produce the slower and more variable UD song (Matheson et al. 2015). As such, we propose that fewer neurons in song control circuitry, in particular, the VMP, will express EGR-1 when older-adult Bengalese finches produce song than when young adults produce song. Such neural changes would be consistent with models linking motor learning and performance to increased neural efficiency (Poldrack 2000). For example, the supplementary motor area and premotor cortex are less activated by finger tapping in expert piano players than in nonmusicians, and such reductions have been hypothesized to reflect more efficient neural processing for motor performance (Hund-Georgiadis and von Cramon 1999; Jäncke et al. 2000; Krings et al. 2000; Lotze et al. 2003). Similar decreases have also been observed in premotor cortex as individuals learn how to play melodies (Chen et al. 2012), and in language-related areas as individuals become more proficient with languages [e.g., Briellmann et al. (2004); Perani et al. (2003); Vingerhoets et al. (2003)]. In songbirds, developmental changes to EGR-1 mRNA expression that are consistent with this hypothesis have been observed in RA: singing causes a smaller increase in EGR-1 mRNA expression in RA when adult zebra finches produce their stereotyped songs than when juvenile zebra finches produce their variable songs (Jarvis et al. 1998; Jin and Clayton 1997; Whitney et al. 2000).

Regardless of the nature of social context and age effects on neural activity, these data provide some insight into the process of adult vocal motor change. Our data demonstrate that the temporal patterning of UD songs converge over time onto the temporal structure of FD_{young} songs (Figs. 2 and 3). By examining how both the UD and FD songs of individual birds changed over time, we tested the hypotheses that FD_{young} songs served as stable targets for song plasticity (target model) or that age-dependent changes simply reflect a change in vocal motor performance (performance model), focusing on how FD song changed over time to distinguish between these models. We found greater support for the performance model of vocal motor change because the stereotypy of syllable sequencing and tempo of an individual's FD song increased significantly over time (Fig. 5). It has been proposed that FD song reveals the physiological limit of a bird's performance (e.g., FD song is the bird's fastest possible song) (Podos et al. 2009), and our results indicate that the best (FD) rendition of song can change as baseline (UD) vocal performance changes.

Despite the evidence against the notion that the FD song of young birds represents a target for adult song plasticity, it remains possible that sensory targets could guide adult vocal plasticity. In particular, it is possible that tutor songs memorized during the first month of development serve as the target for adult vocal motor change. Indeed, some studies suggest that neurons in the adult songbird brain continue to represent the structure of the tutor's song [e.g., Gobes and Bolhuis (2007); Phan et al. (2006); van der Kant et al. (2013)], although the representation of tutor song seems to diminish over development (Achiro and Bottjer 2013; Nick and Konishi 2005; Solis and Doupe 2000). We were unable to test this hypothesis here because of the lack of tutor song recordings for the birds in this study (all birds were purchased from outside vendors), but this study motivates targeted analyses of similarities in the temporal patterning of tutor and pupil songs in Bengalese finches.

Variation in context-dependent changes to syllable structure were also significantly related to age-dependent variation, although the relationship was more complicated than that for temporal features of song. Social context and age affected syllable sequencing and timing to similar degrees, and individual variation in the direction and magnitude of context-dependent changes predicted variation in age-dependent changes. In contrast, social context but not age significantly affected the mean and variability of FF. Despite the lack of significant changes to the mean and variability of FF over time, variation in age-dependent changes to these spectral features of song were correlated with variation in context-dependent changes. For example, syllables that increased more in FF from UD_{young} to

FD_{young} song tended to increase in FF from UD_{young} to UD_{older} song, whereas syllables that increased less or decreased in FF across social context tended to decrease in FF with age (Fig. 4). Despite the lack of overall change in FF over time, these data highlight the potential for context-dependent changes to predict variation in the nature of changes to the spectral structure over time.

In summary, these data demonstrate that by examining variation in the degree to which social context affects the control of an evolutionarily important behavior, one can gain insight into individual differences in the trajectory of behavioral change over time. Our data are consistent with the notion that individual variation in motor ability predicts variation in the magnitude of motor learning [e.g., Adams (1987); Magill (2004)] and suggest that age-dependent changes could reflect improvements to vocal performance (e.g., shorter gap durations). Furthermore, these data suggest the possibility that mechanisms that acutely regulate vocal performance could shape vocal motor plasticity.

Acknowledgements

The authors thank S. C. Woolley, L. E. Matheson, D. M. Toccalino, A. Bernard, and R. Krahe for important contributions to data collection, analysis, interpretation, and presentation.

References

Achiro JM, Bottjer SW. Neural representation of a target auditory memory in a cortico-basal ganglia pathway. *J Neurosci* 33: 14475–14488, 2013.

Adams JA. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin* 101: 41–71, 1987.

Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80: 494–506, 2013.

Andalman AS, Foerster JN, Fee MS. Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature. *PLoS One* 6: e25461, 2011.

Ashmore RC, Wild JM, Schmidt MF. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 25: 8543–8554, 2005.

Basista MJ, Elliott KC, Wu W, Hyson RL, Bertram R, Johnson F. Independent premotor encoding of the sequence and structure of birdsong in avian cortex. *J Neurosci* 34: 16821–16834, 2014.

Brainard MS, Doupe AJ. Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci* 1: 31–40, 2000.

Brainard MS, Doupe AJ. Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J Neurosci* 21: 2501–2517, 2001.

Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36: 489–517, 2013.

Briellmann RS, Saling MM, Connell AB, Waites AB, Abbott DF, Jackson GD. A high-field functional MRI study of quadri-lingual subjects. *Brain Lang* 89: 531–542, 2004.

Byers J, Hebets E, Podos J. Female mate choice based upon male motor performance. *Animal Behaviour* 79: 771–778, 2010.

Catchpole CK, Slater PJB. *Bird song biological themes and variations, 2nd edn.* Cambridge: Cambridge University Press, 2008.

Chen JL, Rae C, Watkins KE. Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage* 59: 1200–1208, 2012.

de Cheveigné A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. *J Acoust Soc Am* 111: 1917–1930, 2002.

Chi Z, Margoliash D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32: 899–910, 2001.

Colonnese MT, Stallman EL, Berridge KC. Ontogeny of action syntax in altricial and precocial rodents: Grooming sequences of rat and guinea pig pups. *Behaviour* 133: 1165–1195, 1996.

Cooper BG, Goller F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J Neurophysiol* 95: 3798–3809, 2006.

Cooper BG, Méndez JM, Saar S, Whetstone AG, Meyers R, Goller F. Age-related changes in the Bengalese finch song motor program. *Neurobiol Aging* 33: 564–568, 2012.

Dayan E, Cohen LG. Neuroplasticity subserving motor skill learning. *Neuron* 72: 443–454, 2011.

Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999

Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. *Trends Neurosci* 28: 353–363, 2005.

Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr Opin Neurobiol* 10: 732–739, 2000.

Drake C, Palmer C. Skill acquisition in music performance: relations between planning and temporal control. *Cognition* 74: 1–32, 2000.

Dunning JL, Pant S, Bass A, Coburn Z, Prather JF. Mate choice in adult female Bengalese finches: females express consistent preferences for individual males and prefer female-directed song performances. *PLoS One* 9: e89438, 2014.

Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychological review* 100: 363, 1993.

Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 15: 988–994, 2008.

Fee MS, Scharff C. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J* 51: 362–377, 2010.

Fentress JC. Emergence of pattern in the development of mammalian movement sequences. *J Neurobiol* 23: 1529–1556, 1992.

Fujimoto H, Hasegawa T, Watanabe D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* 31: 10023–10033, 2011.

Glaze CM, Troyer TW. Temporal structure in zebra finch song: implications for motor coding. *J Neurosci* 26: 991–1005, 2006

Glaze CM, Troyer TW. Development of temporal structure in zebra finch song. *J Neurophysiol* 109: 1025–1035, 2013.

Gobes SM, Bolhuis JJ. Birdsong memory: a neural dissociation between song recognition and production. *Curr Biol* 17: 789–793, 2007.

Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C. Brain structure predicts the learning of foreign speech sounds. *Cereb Cortex* 17: 575–582, 2007.

Golestani N, Paus T, Zatorre RJ. Anatomical correlates of learning novel speech sounds. *Neuron* 35: 997–1010, 2002.

Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419: 65–70, 2002.

Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci* 32: 11671–11687, 2012.

Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J Neurophysiol* 101: 3235–3245, 2009.

Heinig A, Pant S, Dunning J, Bass A, Coburn Z, Prather JF. Male mate preferences in mutual mate choice: finches modulate their songs across and within male-female interactions. *Anim Behav* 97: 1–12, 2014.

Herholz SC, Zatorre RJ. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76: 486–502, 2012.

Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci* 8: 393–402, 2007.

Hund-Georgiadis M, von Cramon DY. Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res* 125: 417–425, 1999.

James LS, Sakata JT. Vocal motor changes beyond the sensitive period for song plasticity. *J Neurophysiol* 112: 2040–2052, 2014

Jäncke L, Shah NJ, Peters M. Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. Brain *Res Cogn Brain Res* 10: 177–183, 2000.

Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F. For whom the bird sings: context-dependent gene expression. *Neuron* 21: 775–788, 1998.

Jin H, Clayton DF. Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. *Neuron* 19: 1049–1059, 1997.

Kakishita Y, Sasahara K, Nishino T, Takahasi M, Okanoya K. Ethological data mining: an automata-based approach to extract behavioral units and rules. *Data Mining and Knowledge Discovery* 18: 446–471, 2009.

Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377: 155–158, 1995.

Karni A, Meyer G, Rey-Hipolito C, Jezzard P, Adams MM, Turner R, Ungerleider LG. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci USA* 95: 861–868, 1998.

Kao MH, Brainard MS. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96: 1441–1455, 2006.

Keith N, Ericsson KA. A deliberate practice account of typing proficiency in everyday typists. *J Exp Psychol Appl* 13: 135–145, 2007.

Kelly AM, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 15: 1089–1102, 2005.

Kelly CW, Sober SJ. A simple computational principle predicts vocal adaptation dynamics across age and error size. *Front Integr Neurosci* 8: 75, 2014.

Kobayashi K, Uno H, Okanoya K. Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. *Neuroreport* 12: 353–358, 2001.

Kojima S, Doupe AJ. Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J Neurophysiol* 98: 2099–2109, 2007.

Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci USA* 108: 1687–1692, 2011.

Kosche G, Vallentin D, Long MA. Interplay of inhibition and excitation shapes a premotor neural sequence. *J Neurosci* 35: 1217–1227, 2015.

Krings T, Töpper R, Foltys H, Erberich S, Sparing R, Willmes K, Thron A. Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett* 278: 189–193, 2000.

Kubikova L, Bosikova E, Cvikova M, Lukacova K, Scharff C, Jarvis ED. Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. *Sci Rep* 4: 6590, 2014.

Landi SM, Baguear F, Della-Maggiore V. One week of motor adaptation induces structural changes in primary motor cortex that predict long-term memory one year later. *J Neurosci* 31: 11808–11813, 2011.

Leblois A, Perkel DJ. Striatal dopamine modulates song spectral but not temporal features through D1 receptors. *Eur J Neurosci* 35: 1771–1781, 2012.

Livingston FS, White SA, Mooney R. Slow NMDA-EPSCs at synapses critical for song development are not required for song learning in zebra finches. *Nat Neurosci* 3: 482–488, 2000.

Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456: 189–194, 2008.

Long MA, Jin DZ, Fee MS. Support for a synaptic chain model of neuronal sequence generation. *Nature* 468: 394–399, 2010.

Lotze M, Scheler G, Tan H-RM, Braun C, Birbaumer N. The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20: 1817–1829, 2003.

Magill RA. *Motor Learning and Control: Concepts and Applications* (7th ed.). New York: McGraw-Hill, 2004.

Matheson LE, Sakata JT. Catecholaminergic contributions to vocal communication signals. *Eur J Neurosci* 14: 1180–1194, 2015.

Matheson LE, Sun H, Sakata JT. Forebrain circuits underlying the social modulation of vocal communication signals. *Dev Neurobiol* 76: 47–63, 2016.

Mooney R. Neurobiology of song learning. *Curr Opin Neurobiol* 19: 654–660, 2009.

Morris D. The reproductive behaviour of the zebra finch, with special reference to pseudofemale behaviour and displacement activities. *Behaviour* 6: 271–322, 1954.

Morrison RG, Nottebohm F. Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *J Neurobiol* 24: 1045–1064, 1993.

Nick TA, Konishi M. Neural song preference during vocal learning in the zebra finch depends on age and state. *J Neurobiol* 62: 231–242, 2005.

Okanoya K. Voco-auditory behavior in the Bengalese finch: A comparison with the zebra finch. *Biomed Res* 18: 53–70, 1997

Okanoya K, Yamaguchi A. Adult Bengalese finches (*Lonchura striata* var. *domestica*) require real time auditory feedback to produce normal song syntax. *J Neurobiol* 33: 343–356, 1997.

Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3: e153, 2005.

Petersen SE, van Mier H, Fiez JA, Raichle ME. The effects of practice on the functional anatomy of task performance. *Proc Natl Acad Sci USA* 95: 853–860, 1998.

Perani D, Abutalebi J, Paulesu E, Brambati S, Scifo P, Cappa SF, Fazio F. The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum Brain Mapp* 19: 170–182, 2003.

Phan ML, Pytte CL, Vicario DS. Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proc Natl Acad Sci USA* 103: 1088–1093, 2006.

Podos J, Lahti DC, Moseley DL. Vocal performance and sensorimotor learning in songbirds. *Adv Study Behav* 40: 159–195, 2009.

Poldrack RA. Imaging brain plasticity: conceptual and methodological issues--a theoretical review. *Neuroimage* 12: 1–13, 2000.

Prather JF, Peters S, Nowicki S, Mooney R. Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature* 451: 305–310, 2008.

Pytte CL, Gerson M, Miller J, Kirn JR. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Medina L, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Güntürkün O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, Smulders TV, Wada K, Dugas-Ford J, Husband S, Yamamoto K, Yu J, Siang C, and Jarvis ED. Revised nomenclature for avian telencephalon and some related brainstem nuclei. *J Comp Neurol* 473: 377– 414, 2004.

Sakata JT, Crews D. Embryonic temperature shapes behavioural change following social experience in male leopard geckos, *Eublepharis macularius*. *Anim Behav* 66: 839–846, 2003.

Sakata JT, Brainard MS. Online contributions of auditory feedback to neural activity in avian song control circuitry. *J Neurosci* 28: 11378–11390, 2008.

Sakata JT, Brainard MS. Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *J Neurophysiol* 102: 2485–2497, 2009.

Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99: 1700–1711, 2008.

Sakata JT, Vehrencamp SL. Integrating perspectives on vocal performance and consistency. *J Exp Biol* 215: 201–209, 2012.

Samuels M, Witmer J. *Statistics for the Life Sciences. 3rd edition*. New Jersey: Prentice Hall, 2002.

Sanes JN, Donoghue JP. Plasticity and primary motor cortex. *Annu Rev Neurosci* 23: 393–415, 2000.

Schmidt MF. Pattern of interhemispheric synchronization in HVc during singing correlates with key transitions in the song pattern. *J Neurophysiol* 90: 3931–3949, 2003.

Schmidt RA, Lee T. *Motor control and learning*. Champaign, IL: Human kinetics, 1998.

Shadmehr R, Krakauer JW. A computational neuroanatomy for motor control. *Exp Brain Res* 185: 359–381, 2008.

Solis MM, Doupe AJ. Compromised neural selectivity for song in birds with impaired sensorimotor learning. *Neuron* 25: 109–121, 2000.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F. Dual pre-motor contribution to songbird syllable variation. *J Neurosci* 31: 322–330, 2011.

Tomassini V, Jbabdi S, Kincses ZT, Bosnell R, Douaud G, Pozzilli C, Matthews PM, Johansen-Berg H. Structural and functional bases for individual differences in motor learning. *Hum Brain Mapp* 32: 494–508, 2011.

Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning II. Temporal hierarchies and the learning of song sequence. *J Neurophysiol* 84: 1224–1239, 2000.

Ungerleider LG, Doyon J, Karni A. Imaging brain plasticity during motor skill learning. *Neurobiol Learn Mem* 78: 553–564, 2002.

van der Kant A, Derégnaucourt S, Gahr M, Van der Linden A, Poirier C. Representation of early sensory experience in the adult auditory midbrain: implications for vocal learning. *PLoS One* 8: e61764, 2013.

Vingerhoets G, Van Borsel J, Tesink C, van den Noort M, Deblaere K, Seurinck R, Vandemaele P, Achten E. Multilingualism: an fMRI study. *Neuroimage* 20: 2181–2196, 2003.

Vu ET, Mazurek ME, Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14: 6924–6934, 1994.

Wang CZ, Herbst JA, Keller GB, Hahnloser RH. Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biol* 14: e250, 2008.

Warren TL, Charlesworth JD, Tumer EC, Brainard MS. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32: 15414–15425, 2012.

Whitney O, Soderstrom K, Johnson F. Post-transcriptional regulation of zenk expression associated with zebra finch vocal development. *Brain Res Mol Brain Res* 80: 279–290, 2000.

Williams H, Mehta N. Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *J Neurobiol* 39: 14–28, 1999.

Willingham DB. A neuropsychological theory of motor skill learning. *Psychol Rev* 105: 558–584, 1998.

Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci* 358: 593–602, 2003.

Woolley SC, Doupe AJ. Social context-induced song variation affects female behavior and gene expression. *PLoS Biol* 6: e62, 2008.

Woolley SC, Kao MH. Variability in action: Contributions of a songbird cortical-basal ganglia circuit to vocal motor learning and control. *Neuroscience* 296: 39–47, 2015.

Yu AC, Margoliash D. Temporal hierarchical control of singing in birds. *Science* 273: 1871–1875, 1996.

Zann RA. *The zebra finch: a synthesis of field and laboratory studies*. New York: Oxford University Press, 1996.

Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 8: 547–558, 2007.

Zatorre RJ, Delhommeau K, Zarate JM. Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front Psychol* 3: 544, 2012.

Figures

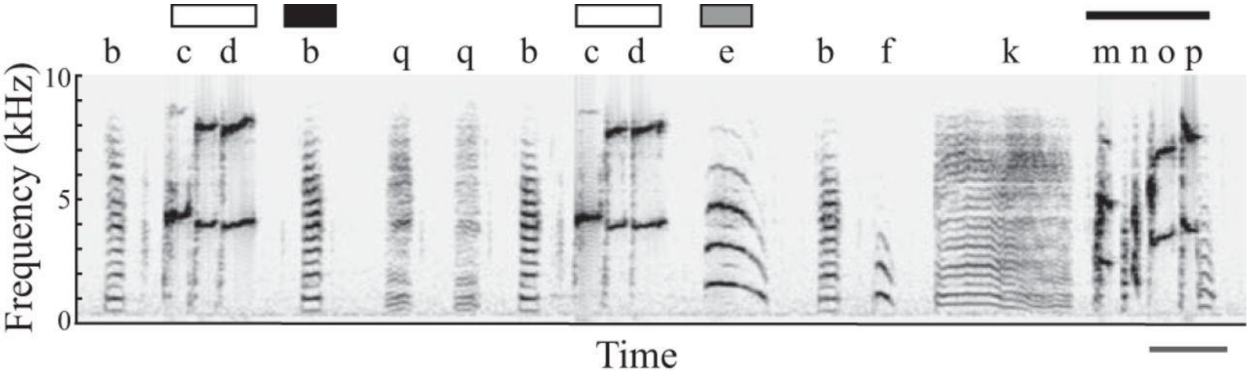


Figure 1. Adult Bengalese finch song. A spectrogram (time on the x-axis, frequency on the y-axis, darkness as amplitude) of a rendition of Bengalese finch song. Labels for syllables are located above the spectrogram. As with other Bengalese finches, this bird produces both variable (branch point) and stereotyped sequences within his song. The sequence “cd” (white bars) is an example of a branch point because it can be followed by a “b” (black bar), “e” (gray bar), or “q” (transition to q is not observed in this rendition). In contrast, the sequence “mnop” (black line) is a stereotyped sequence in which syllables are always produced in this order. Scale bar, 200 ms.

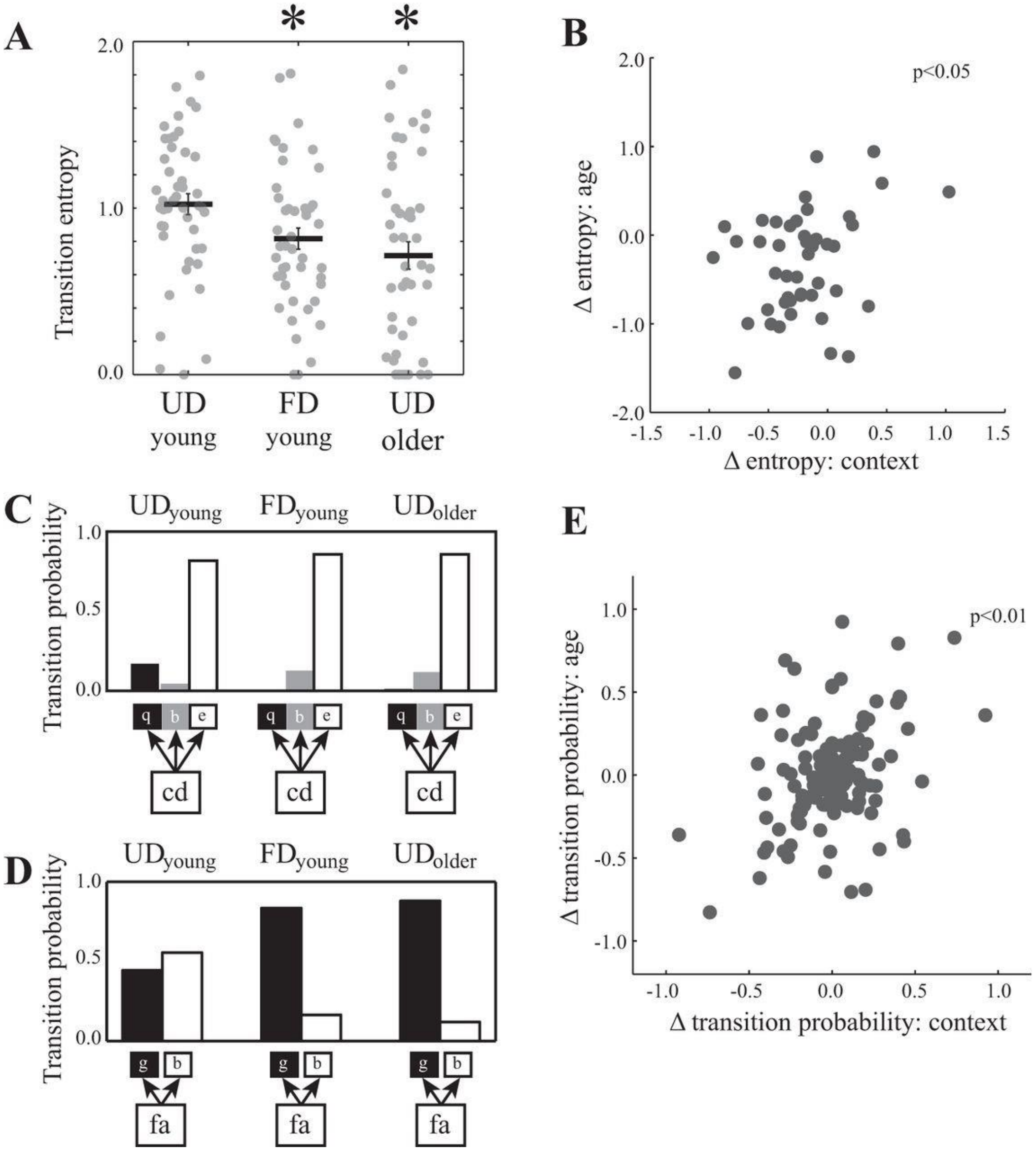


Figure 2. Context-dependent changes to syllable sequencing predicted age-dependent changes to syllable sequencing. A: transition entropy, a measure of sequence variability, decreased significantly across social context and age. Plotted is the mean (±SE) transition entropy of branch points for the undirected and female-directed young-adult (UD_{young} and FD_{young}, respectively) and UD older-adult (UD_{older}) songs of individual birds (n = 46 branch points), with raw data values plotted in gray. Transition entropy is significantly lower for FD_{young} and UD_{older} song than for UD_{young} song and not different between FD_{young} and UD_{older} song. *$P < 0.05$, significantly different than UD_{young} song. B: there is a significant, positive

relationship between the magnitude of change (Δ) in entropy caused by social context and by age (P = 0.0430). Plotted is the difference in transition entropy across context ($FD_{young} - UD_{young}$) and the difference across age ($UD_{older} - UD_{young}$). C: an example of context- and age-dependent changes to syllable sequencing at branch points. The bird can produce the syllables q (black bar), b (gray bars), and e (white bars), following the branch-point sequence cd (same bird depicted in Fig. 1). Social context and age led to a virtual elimination of the transition to q and increases in transitions to b and e. D: another example of context- and age-dependent changes to syllable sequencing at branch points. The bird produced the syllables "g" (black bars) and b (white bars) following the branch point "fa," and the changes to both transitions across context and age were comparable. E: context-dependent changes to transition probabilities for all branch points were significantly correlated with age-dependent changes to transition probabilities ($P < 0.0001$). Plotted are the differences in transition probabilities across context ($FD_{young} - UD_{young}$) and across age ($UD_{older} - UD_{young}$).

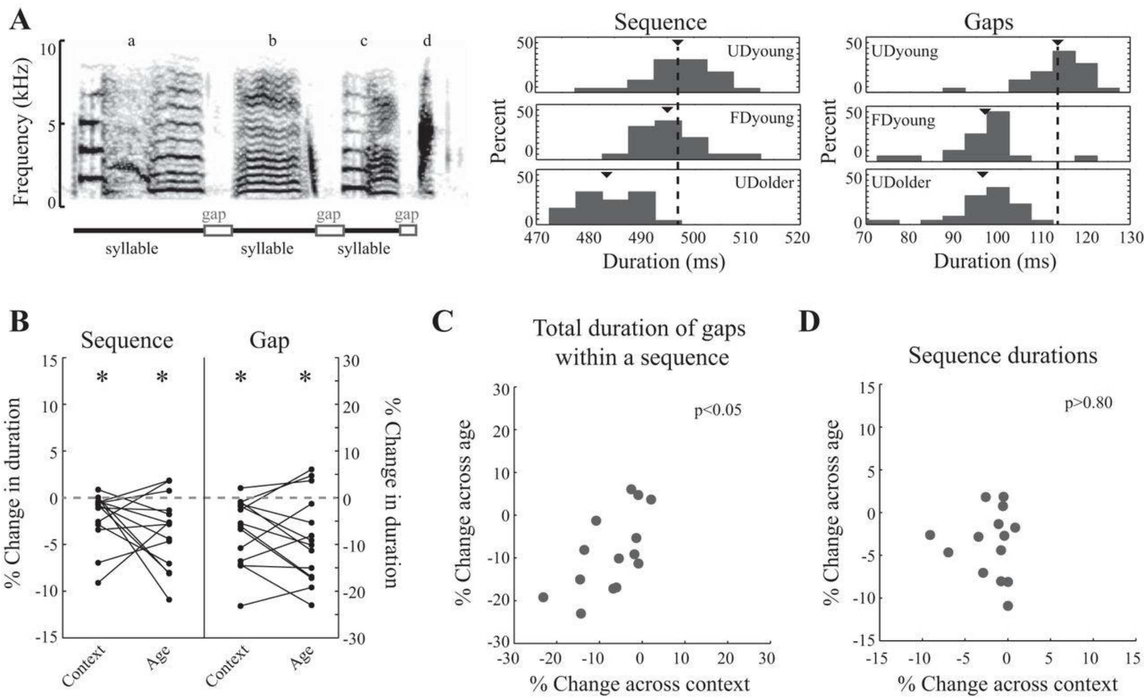


Figure 3. Context-dependent changes to syllable timing predicted age-dependent changes to syllable timing. A: spectrogram of a stereotyped sequence used to analyze context- and age-dependent changes to song tempo. Labels for syllables are located above the spectrogram, and syllables (lines) and intersyllable gaps (empty boxes) are highlighted under the spectrogram. Histograms display sequence durations (middle) and the total duration of gaps within the sequence (right) across renditions, and triangles indicate the mean duration for each condition. Dashed lines correspond to the mean duration for UD_{young} song. In this example, both sequence and gap durations were shorter for both FD_{young} and UD_{older} song than for UD_{young} song. B: sequence (left) and gap (right) durations change significantly across social context and age, and the magnitudes of context- and age-dependent changes to sequence and gap durations were not significantly different (n = 14 sequences). Plotted are the percent changes to sequence and gap durations across context (UD_{young} to FD_{young}) and age (UD_{young} to UD_{older}). *P < 0.05 indicates that the mean percent change was significantly different than 0 (t-test). C: context-dependent changes to gap durations (percent change) are correlated with age-dependent changes (P = 0.0174). Plotted are the percent changes to gap durations across context (UD_{young} to FD_{young}) and across age (UD_{young} to UD_{older}). D: context-dependent changes to sequence durations

(percent change) were not correlated with age-dependent changes (P = 0.8141). This lack of correlation at the sequence level (despite the correlation at the gap level) is likely due to the lack of correlation in syllable durations (P = 0.5492) as well as the fact that syllables contribute more to sequence durations than gaps.

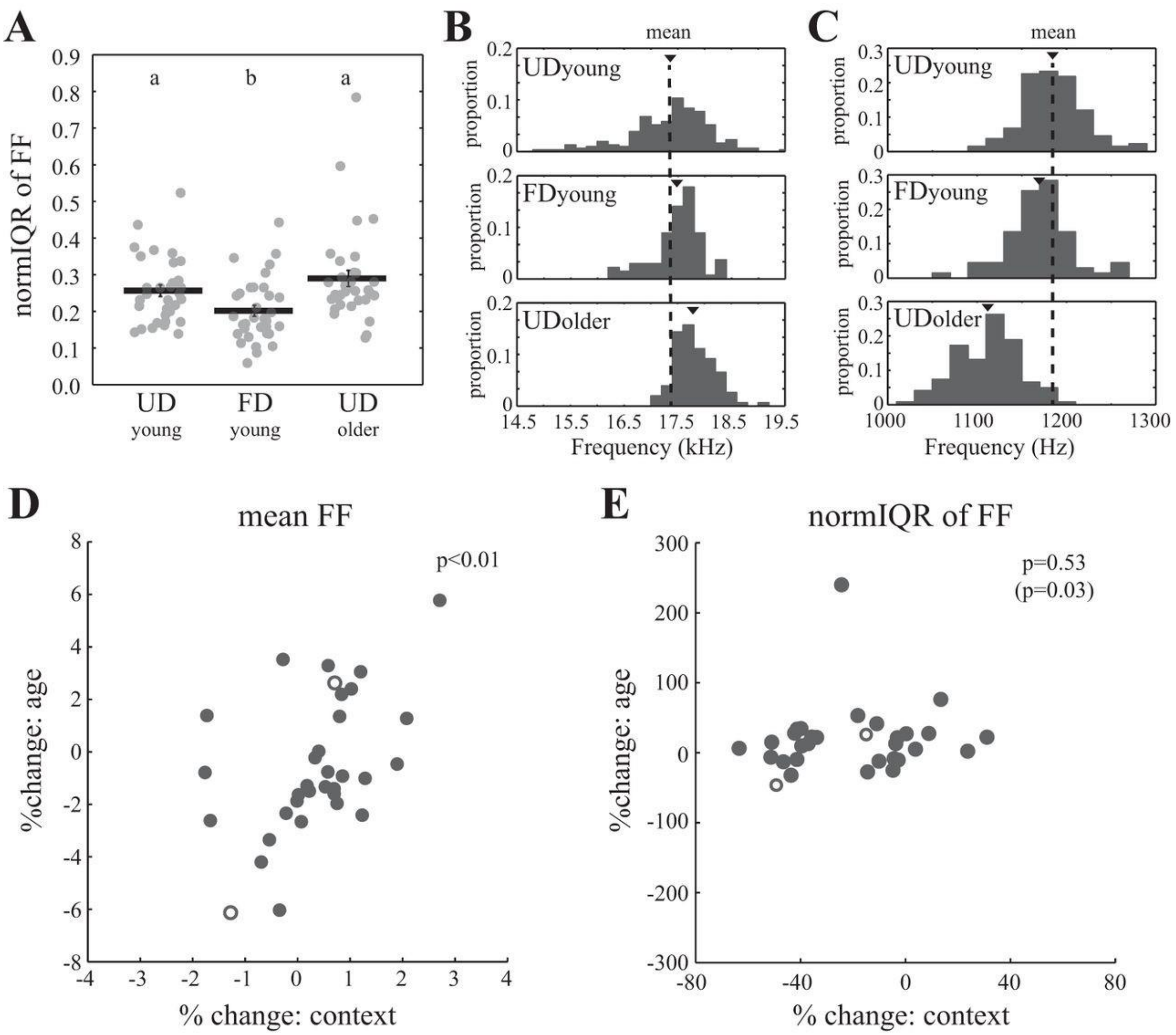


Figure 4. Context-dependent changes to fundamental frequency (FF) predicted age-related variation in FF (n = 33 syllables in 14 males). A: social context but not age affected the variability of FF. Plotted are normalized interquartile region (normIQR) values for each individual syllable, with lines indicating the mean ± SE for each condition. Groups with different letters indicate groups that are significantly different. B: distributions of FF across renditions for an individual syllable with flat, harmonic structure. Triangles indicate the mean FF for each condition, and the dashed line corresponds to the mean FF for UD_{young} song to help visualize context- and age-dependent changes. The mean FF of this syllable increased across both social contexts (from UD_{young} to FD_{young}) and age (from UD_{young} to UD_{older}). The variability of FF for this syllable decreased across both social contexts and age. C: distributions of FF across renditions for another syllable. The mean FF of this syllable

decreased across both context and age. The variability of FF for this syllable decreased slightly across context and increased with age. D: variation in the magnitude of change in the mean FF across social context (percent change from UD_{young} to FD_{young}) predicts variation in the magnitude of change across age (percent change from UD_{young} to UD_{older}; $P = 0.0049$). Open circles represent example syllables depicted in B and C. E: overall, there was no significant relationship between the magnitude of context-dependent changes (percent change from UD_{young} to FD_{young}) to the variability (normIQR) of FF and the magnitude of age-dependent changes (percent change from UD_{young} to UD_{older}; $P = 0.5308$). However, after the removal of the value for a single outlying syllable, the relationship between context- and age-dependent changes to normIQR was significant ($P = 0.0342$). Open circles represent syllables summarized in B and C.

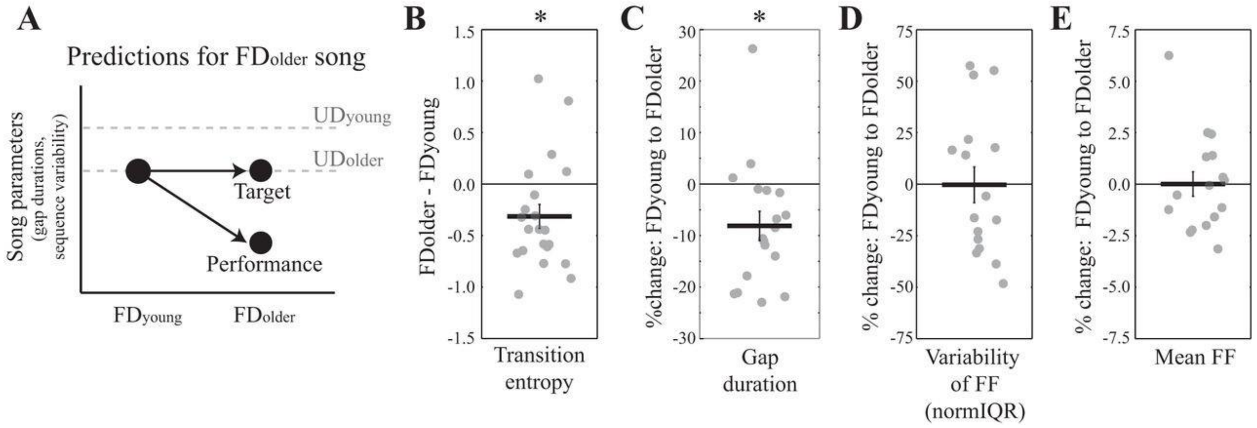


Figure 5. The testing of the “target” and “performance” models of vocal motor change. A: predictions of the target and performance models. The performance but not the target model predicts that features that changed significantly over time for UD song (syllable sequencing and tempo) should similarly change for FD song. As such, the performance model predicts that gap durations and sequence variability of FD song (black circles) should decrease over time. The target model predicts that these features of FD song should remain the same over time. Dashed gray lines represent UD$_{\text{young}}$ and UD$_{\text{older}}$ songs and highlight the decrease for UD song over time. B: the transition entropy of branch points decreased from FD$_{\text{young}}$ to FD$_{\text{older}}$ songs [Tukey's honestly significant difference (HSD); $P = 0.0241$], indicating age-dependent changes to syllable sequencing in FD song. Whereas analyses were performed using a factorial design (see results), depicted in this panel (as well as subsequent panels of this figure) are changes to song features for FD song with age. We depict the FD$_{\text{young}}$-FD$_{\text{older}}$ contrast because this contrast is central for the comparison of target and performance models and plot changes instead of mean values for each condition, as differences in mean values are difficult to see for some features because of the large range of values (see materials and methods). C: gap durations decreased from FD$_{\text{young}}$ to FD$_{\text{older}}$ song (Tukey's HSD; $P = 0.0085$), demonstrating an age-dependent increase to the tempo of FD song. D: the variability (normIQR) of FF does not change significantly from FD$_{\text{young}}$ to FD$_{\text{older}}$ song (Tukey's HSD; $P = 0.7680$). E: mean FF does not change significantly from FD$_{\text{young}}$ to FD$_{\text{older}}$ song (Tukey's HSD; $P = 0.9833$). *$P < 0.05$ indicates features of FD song that change significantly over time.

Linking statement from chapter 2 to chapter 3

In chapters 1 and 2 we documented age-dependent changes in Bengalese finches and found that context-dependent modulations to song could predict those age-dependent changes. In the next chapter, we repeat these two analyses in a different species of songbird, the closely related zebra finch. One powerful aspect of songbirds as a model system is the ability to compare species and highlight which aspects of song learning and plasticity are fundamental to many species, compared to other aspects that may be more species-specific. We also added an additional analysis in this manuscript by including a group of birds that were non-socially housed during development. With this analysis, we aimed to determine whether social interactions with a singing male are important for the age-dependent changes to song and the link between age- and context-dependent changes.

Chapter 3

Developmental modulation and predictability of age-dependent vocal plasticity in adult zebra finches

Logan S. James & Jon T. Sakata

Resubmitted to: *Brain Research*

Abstract

Predicting the nature of behavioral plasticity can provide insight into mechanisms of behavioral expression and control. Songbirds like the zebra finch rely on vocal signals for communication, and the performance of these signals demonstrate considerable plasticity over development. Traditionally, these signals were thought to be fixed in adulthood, but recent studies have revealed significant age-dependent changes to spectral and temporal features of song in adult songbirds. A number of age-dependent changes to song resemble acute changes to adult song performance across social contexts (e.g., when an adult male sings to a female relative to when he sings in isolation). The ability of variation in social context-dependent changes to predict variation in age-dependent plasticity would suggest shared mechanisms, but little is known about this predictability. In addition, although developmental experiences can shape adult plasticity, little is known about the extent to which social interactions during development affect age-dependent change to adult song. To this end, we systematically analyzed age- and context-dependent changes to adult zebra finch song, and then examined the degree to which age-dependent changes varied across birds that were social or non-socially tutored birds and to which social context-dependent changes predicted age-dependent changes. Non-socially tutored birds showed more dramatic changes to the broad structure of their motif over time than socially tutored birds, but non-socially and socially tutored birds did not differ in the extent of changes to various spectral and temporal features of song. Overall, we found that adult zebra finches produced longer and more spectrally stereotyped songs when they were older than when they were younger. Moreover, regardless of developmental tutoring, individual variation in age-dependent changes to song bout duration and syllable repetition were predicted by variation in social context-dependent changes to these features. These data indicate that social experiences during development can shape some aspects of adult plasticity and that acute context-dependent and long-term age-dependent to some song features could be mediated by modifications within similar neural substrates.

1. Introduction

The performance of various behaviors, particularly those that are learned, can undergo significant changes across an individual's lifetime. Long-term changes to behavioral performance are often characterized by increases in speed, accuracy, and consistency, and these changes are thought to be driven by reinforcement, practice, and feedback-dependent mechanisms (Costa et al., 2004; Dhawale et al., 2017; Doya, 2000; Drake and Palmer, 2000; Ericsson et al., 1993; Makino et al., 2016; Schmidt, 1988; Ungerleider et al., 2002; Woolley, 2019). Even after achieving expertise in a behavior, these mechanisms continue to operate and modulate performance; for example, even trained musicians and singers continue to improve their performance with practice (Dhawale et al., 2017; Drake and Palmer, 2000; Ericsson, 2008; Keith and Ericsson, 2007). However, we know relatively little about the factors that can regulate and predict the extent of such behavioral plasticity, especially regarding the plasticity of complex and natural motor behaviors (Golestani et al., 2007, 2002; Herholz and Zatorre, 2012; Landi et al., 2011).

Songbirds offer a powerful opportunity to understand the factors and mechanisms that guide the performance and plasticity of complex, learned behaviors. Songbirds such as zebra finches learn how to produce their vocal communication signals ("songs") through sensory learning and sensorimotor practice during development (Brainard and Doupe, 2000, 2013; Mooney, 2009; Tchernichovski et al., 2001). Juvenile songbirds memorize the sound of the song of an adult tutor, and this auditory memory serves as a sensory template to guide vocal motor learning throughout development (Bolhuis et al., 2010; Brainard and Doupe, 2002; Chen et al., 2016; Prather et al., 2017). Young songbirds initially produce "noisy", poorly-structured vocalizations which are refined into highly stereotyped sequences of vocal elements ("syllables") in a process that involves changes to the speed, sequencing and consistency of motor actions (Brainard and Doupe, 2002; Kelly and Sober, 2014; Kojima and Doupe, 2007; Livingston et al., 2000; Morrison and Nottebohm, 1993; Murphy et al., 2017; Ölveczky et al., 2005).

At sexual maturity, some species of songbirds, including zebra finches, are no longer able to learn a new song (Beecher and Brenowitz, 2005; Doupe and Kuhl, 1999). Despite this, individuals in many of these species display long-term age-dependent changes in song performance. For example, relative to young adults, older adult zebra finches produce faster and more stable songs with lower variability in syllable structure (Brainard and

Doupe, 2001; Kao and Brainard, 2006; Lombardino and Nottebohm, 2000). Similarly, older Bengalese finches produce faster songs that are more consistent in syllable structure and sequencing than young adult Bengalese finches (James and Sakata, 2014). Analagous changes have also been observed in other songbird species (Ota and Soma, 2014; Podos et al., 2009; Rivera-Gutierrez et al., 2012; Sakata and Vehrencamp, 2012).

Behavioral plasticity in adulthood can be shaped by experiences during development (Sakata and Crews, 2004). For example, embryonic environments can shape the degree to which social experiences with females lead to changes in social behavior (Sakata and Crews, 2003), and developmental experiences also shape the manner in which adult rodents respond to stressors and various types of learning paradigms (Dalla and Shors, 2009; Kosten et al., 2012; Shors and Miesegaes, 2002). The degree to which plasticity in various aspects of social and communicative behaviors is shaped by developmental experiences remains largely unexplored.

In addition to the possibility that developmental experiences could shape adult song plasticity, it is possible that the nature of adult song change could be predicted by analyzing other forms of song modulation. In particular, adult songbirds acutely modify their songs in response to different social contexts in a manner that resembles long-term age-dependent changes to song. For example, zebra finches and Bengalese finches sing faster and more stereotyped songs when singing towards a female compared to when singing alone (Aronov and Fee, 2012; James et al., 2018; James and Sakata, 2015; Kao and Brainard, 2006; Matheson et al., 2016; Sakata et al., 2008; Scharff and Nottebohm, 1991; Sossinka and Böhner, 1980; Woolley and Doupe, 2008). Given the similarities in the nature of song changes, it is possible that the direction and extent of context-dependent changes to song in young adult finches could predict the direction and extent of age-dependent changes to song. Consistent with this notion, social context-dependent changes to the sequencing, timing, and acoustic structure of songs predict long-term age-dependent changes to Bengalese finch song (James and Sakata, 2015). However, the degree to which age-dependent changes are similarly predictable across other species remains largely unexplored.

Here, we provide a longitudinal study documenting a wide array of age-dependent changes to song in adult zebra finches as well as the extent to which these changes can be predicted by context-dependent changes. In addition, we compared the extent and predictability of age-dependent changes between birds with different social experiences during development (i.e., adult zebra finches that were socially tutored during development or were tutored in a non-social manner throughout development). Little is known about the factors that regulate age-dependent changes to adult song, and significant insights can be gained by complementing the discovery of features that change in a correlated manner across social context and age with existing knowledge about neural mechanisms underlying context-dependent changes in those features.

2. Results

We examined age- and context-dependent changes to song in zebra finches tutored and raised in a social environment (n=8) or tutored in a non-social environment (tutored via operantly-triggered song playback) and raised in isolation until they were sexually mature (4 months of age; n=9). At 4 months of age ("young adult"), birds were individually housed in a sound-attenuating chamber, and we analyzed how both groups of birds changed their songs across social contexts by collecting interleaved renditions of undirected (UD_{young}) and female-directed (FD_{young}) song. After these recordings, all birds were socially housed until they were individually recorded as older adults (8-39 months of age; all UD song; UD_{older}). [We did not detect any relationship between the number of months between recordings and the magnitude of change in song features (see Experimental Procedure); therefore all UDolder songs are considered together, regardless of the duration between recordings.] This experiment was designed to evaluate the degree to which context-dependent changes to song (UD_{young} to FD_{young}) compared to and were able to predict age-dependent changes to song (UD_{young} to UD_{older}; see James and Sakata, 2015). In addition, this study aimed to assess the extent to which developmental song tutoring could affect the nature and predictability of age-dependent changes to song.

2.1 Age- and context-dependent changes to song motifs

Zebra finch songs consist of highly structured vocal elements (“syllables”) separated by silent intervals (“gaps”; Figure 1; see Experimental Procedure for additional details on song definitions). Songs are produced in “bouts,” which represent contiguous epoch of vocalizations in which syllables are separated by <500 ms. The fundamental unit of zebra finch song is the “motif”: a stereotyped sequence of syllables that is repeated multiple times within a song bout (e.g., “abcde” in Figure 1). Motifs can be readily identified because of the stereotypy of syllable sequencing within the motif, because gap durations between syllables within the motif are generally shorter than those between syllables outside the motif, and because adult zebra finches usually terminate their song bout at the end of a motif (Figure 1).

Normally-reared zebra finches generally produce the same motif throughout their life (Funabiki and Funabiki, 2008; Kao and Brainard, 2006; Pytte et al., 2007; Zann, 1996). We confirmed this to be the case for all of our socially tutored birds; they repeated the same sequence of syllables across time points (UD_{young} and UD_{older}). For example, the socially tutored bird in Figure 1 produced the motif “abcde” both when he was a young adult (4 months old) and again 30 months later. A number of non-socially tutored adults in this study also produced the same motif across age. However, roughly half of the non-socially tutored birds (5 of the 9 birds) exhibited salient changes to the acoustic structure of their motifs over time (Figure 2; Supplementary Figure 1). Overall, we observed a number of examples in which the structure of a syllable was substantially modified over time as well as examples of syllable additions to or deletions from the motif. In the example provided in Figure 2A, all syllables retained across UD_{young} and UD_{older} songs exhibited noticeable acoustic changes. In addition, the “g” syllable was dropped from the bird’s motif, and a syllable (“h”) was added to the end of his motif. Including this example, we observed three syllables that were dropped from the motif and six syllables that were added to the motif (Supplementary Figure 1). We also observed two examples in which a single syllable “split” into two distinct syllables over time (Supplementary Figure 1; see also Okubo et al., 2015).

Overall, these broad observations suggest that the spectral structure of motifs changed more over time for non-socially tutored birds than for socially tutored birds. In addition, these changes to the songs of non-socially tutored birds could reflect a convergence toward their tutor's song. To examine these possibilities, we used an established multivariate algorithm (see Experimental Procedure; Tchernichovski et al., 2000) to compute the acoustic similarity between songs. We first calculated the acoustic similarity of motifs produced when birds were young adults and when birds were older adults (UD_{young} and UD_{older}) and then compared the degree of similarity between socially and non-socially tutored birds. Lower similarity scores between UD_{young} and UD_{older} would represent larger changes in acoustic structure over time. Consistent with our visual comparisons of spectrograms, we found that the degree of acoustic similarity between UD_{young} and UD_{older} songs was significantly lower for non-socially tutored birds than for socially tutored birds (t-test; $t_{16}=2.3$, $p=0.0337$). We also found that song changes in non-socially tutored birds did not reflect a convergence toward their tutor's song since the similarity to the tutor song was not significantly different between their UD_{young} and UD_{older} songs ($F_{1,8}=0.2$, $p=0.6343$).

Songbirds rapidly change some aspects of the acoustic structure of their songs in ways that resemble age-dependent changes (James and Sakata, 2015, 2014; Kojima and Doupe, 2011). Therefore, we investigated the extent to which the acoustic structure of motifs changed across social context in young adult birds (UD_{young} to FD_{young}). All birds in the study, regardless of tutoring, produced the same motif across social contexts (Figure 2A & 2B). For example, the acoustic structure of syllables in the song of the non-socially tutored bird depicted in Figure 2A did not seem to substantially change across social contexts (but did change with age). Thus, our qualitative inspection of spectrograms suggested that song motifs were acoustically more similar across social contexts than across age and that the extent of context-dependent changes to motifs was not different between socially and non-socially tutored birds.

To test these observations, we computed acoustic similarity scores between UD_{young} and FD_{young} songs and compared these values to similarity scores across age (see above). We used a full-factorial mixed effect model with "Method of Tutoring" (social vs. non-social) and “Similarity Comparison” [context (UD_{young}:FD_{young}) vs. age (UD_{young}:UD_{older})] as the independent factors, acoustic similarity as the dependent factor, and BirdID as a random factor (see Experimental Procedure). We observed a significant main effect of “Similarity Comparison”, with acoustic similarity being significantly higher for context comparisons than for age comparisons ($F_{1,15}$=37.2, p<0.0001; Figure 2C). In addition, there was a significant interaction between “Method of Tutoring” and “Similarity Comparison” ($F_{1,15}$=12.1, p=0.0034), which was driven by the fact that acoustic similarity across age was different between socially and non-socially tutored birds (p=0.0356) but acoustic similarity across context was not different between socially and non-socially tutored birds (p=0.5422).

2.2 Age- and context-dependent changes to syllable structure, timing, and sequencing

We next examined experimental changes to various features of song performance known to change across social contexts (e.g., Cooper and Goller, 2006; James et al., 2018; Kao and Brainard, 2006; Matheson et al., 2016; Sakata et al., 2008; Sossinka and Böhner, 1980; Teramitsu and White, 2006) and investigated the extent to which similar types of changes were observed across age. Specifically, we calculated the mean as well as the variability (coefficient of variation, or CV) of acoustic features within UD_{young}, FD_{young}, and UD_{older} conditions, and then compared these features across conditions using mixed effects models (see Experimental Procedure). Overall, we found no interaction between condition (UD_{young}, FD_{young}, and UD_{older}) and the method of tutoring (social vs. non-social; p>0.10 for all); in other words, socially and non-socially tutored birds exhibited the same type and magnitude of changes across conditions. Consequently, for ease of presentation, we only report the main effect of condition here (see Supplementary Table 1 for full models).

2.2.1 Syllable structure: Mean and variability of the fundamental frequency

Across various studies, the mean and variability (CV) of FF of syllables with flat, harmonic structure (see syllable “e” in Figure 1) have been found to change across social context and age (e.g., James et al., 2018; Kao and Brainard, 2006; Leblois et al., 2010; Sakata et al., 2008; Teramitsu and White, 2006). The mean FF of harmonic syllables significantly differed across conditions ($F_{2,48}$=3.2, p=0.0477; Figure 3A), with mean FF being significantly higher during FD_{young} song than during the UD_{older} song (p=0.0456).

There was also significant variation in the CV of FF across experimental conditions ($F_{2,48}$=28.9, p<0.0001; Figure 3B). Post-hoc tests indicated significant differences across all three conditions (p<0.015 for all). In particular, the CV of FF was lower for FD_{young} and UD_{older} songs compared to UD_{young} songs, indicating that syllables became more stereotyped in pitch when singing to females and over time. In addition, the CV of FF was lower during FD_{young} song than during UD_{older} song, suggesting that social context had a more substantial effect on syllable stereotypy than age.

2.2.2 Syllable timing: song and sequence durations

Mean song bout durations significantly varied across conditions ($F_{2,30}$=14.7, p<0.0001; Figure 4A). Overall, FD_{young} song bouts were longer than UD_{young} and UD_{older} song bouts (p<0.02 for all), and UD_{older} song bouts tended to be longer than UD_{young} song bouts (p=0.0678). In other words, song bout durations increased across social context and over time.

The CV of song bout durations also significantly varied across conditions ($F_{2,30}$=43.9, p<0.0001; Figure 4B), with song bout durations being more variable during FD_{young} song than during UD_{young} and UD_{older} songs (p<0.0001 for both).

Motifs are the principal components of song bouts, and changes to song bout durations could be due to changes in the number of times a motif is repeated within a song as well as changes in the speed at which motifs were produced. (Changes in the repetition of glue syllables that are produced between motifs could also contribute to changes in song bout duration, and these are discussed later). There was significant variation in the number of times motifs were repeated within a song bout across conditions ($F_{2,30}$=11.0, p=0.0003; Figure 4C), with birds producing more motifs per song bout during both FD_{young} and UD_{older}

songs than during UD_{young} song (p<0.002 for both). There was no difference between FD_{young} and UD_{older} songs in the number of motifs per song bout (p=0.9581).

Similar to the CV of song bout durations, the CV of the motifs per song bout varied across experimental conditions ($F_{2,30}$=21.3, p<0.0001; Figure 4D), with CV being higher for FD_{young} song than for UD_{young} and UD_{older} songs (p<0.0001 for both).

We next compared sequence (motif) durations, a proxy for song tempo, across age and context (e.g., the motif "abcde" in Figure 1). For the few birds that changed their motifs across age, we computed the duration of the sequence of syllables that was common across recordings. We found significant variation in sequence durations across conditions ($F_{2,30}$=4.0, p=0.0291; Figure 4E), with sequence durations being shorter for UD_{older} song (p=0.0381) and FD_{young} song (p=0.0749) than for UD_{young} song. Sequence durations were not different between FD_{young} and UD_{older} songs. In other words, song tempo increased across social context and age. To try to isolate the nature of changes to sequence durations, we also analyzed the total duration of syllables and of inter-syllable intervals ("gaps") within the motifs. Although the trends for syllable and gap durations were the same as for sequence durations (i.e., durations were shorter for UD_{older} and FD_{young} song than UD_{young} song), we did not find a significant effect of condition on the duration of syllables ($F_{2,30}$=1.9, p=0.1667) or gaps ($F_{2,30}$=1.0, p=0.3842).

There was no significant variation across experimental conditions in the CV of sequence durations ($F_{2,30}$=0.4, p=0.6917; Figure 4F).

2.2.3 Syllable sequencing

Variation in the repetition of vocal elements as well as sequencing patterns of non-repeating elements have been examined to assess experimental changes to syllable sequencing. The onset of song is preceded by the repetition of simple vocalizations ("introductory notes"; "i" in Figure 1), and similar types of vocalizations ("glue syllables") can be repeated between motifs (e.g., "x" in Figure 1; Bruno and Tchernichovski, 2017;

Horita et al., 2008; Kubikova et al., 2014; Sossinka and Böhner, 1980; Zann, 1996). There was significant variation in the mean number of introductory notes across conditions ($F_{2,30}$=13.5, p<0.0001), with more introductory notes produced before FD_{young} song than before UD_{young} and UD_{older} songs (p<0.0008 for both; Figure 5A).

Similarly, there was significant variation in the CV of introductory note repetition across conditions ($F_{2,30}$=11.8, p=0.0002; Figure 5B), with introductory note repetition being more variable during FD_{young} song than during UD_{young} and UD_{older} songs (p<0.004 for both).

Like introductory notes, there was significant variation in the mean number of times glue syllables were repeated ($F_{2,30}$=18.2, p<0.0001; Figure 5C), with glue syllable repetition being higher during FD_{young} song than during UD_{young} and UD_{older} songs (p<0.005 for both).

The CV of glue syllable repetition also significantly varied across condition ($F_{2,30}$=29.2, p<0.0001; Figure 5D). Post-hoc tests indicated significant differences between all three conditions, with the CV being highest in FD_{young} song, followed by UD_{young} song, and then UD_{older} song (p<0.03 for all). Thus, the variability in glue repeat number increased across context but decreased across age.

In addition to examining the repetition of introductory notes and glue syllables, three related measures of syllable sequencing are used to analyze experimental changes to zebra finch song: sequence linearity, sequence consistency and sequence entropy (see Experimental Procedure; syllable repetitions were excluded from this analysis). Overall, zebra finches produce motifs that are very stereotyped in sequencing, but there can exist variation in the timing of song terminations within the motif and whether birds transition to glue syllables after the motif, immediately transition to another motif, or terminate the song bout (Figure 6A). For example, in Figure 6A, the bird terminated his motif (by transitioning to the glue syllable “x” or ending the song) after the “a”, “b”, or “c” syllables, albeit with different probabilities. Such sequence variation can be captured, to different degrees, using these three measures of sequencing. While there was no effect of condition on sequence linearity ($F_{2,30}$=0.7, p=0.4840; Figure 6B), sequence consistency did vary across conditions ($F_{2,30}$=3.7, p=0.0368; Figure 6C), with sequence consistency being significantly higher during UD_{older} song than during FD_{young} song (p=0.0348). In addition, there was significant variation in sequence entropy across conditions ($F_{2,30}$=3.9, p=0.0314; Figure 6D), with sequence entropy tending to be lower during UD_{older} song than during both UD_{young} (p=0.0522) and FD_{young} (p=0.0593) songs.

The preceding analyses of sequence transitions throughout the song bout suggest that syllable sequencing could become more stereotyped over time and highlight the transitions to glue syllables as sources of sequence variability. To assess the extent to which sequence variability within the motif contributed to these patterns, we conducted additional sets of analyses in which we only included transitions between syllables within the motif (i.e., we excluded introductory notes, glue syllables and transitions from the last syllable in the motif but included song interruptions). We no longer observed any significant differences across conditions (linearity: $F_{2,32}$=0.8, p=0.4619; consistency: $F_{2,32}$=0.5, p=0.6015; weighted entropy: $F_{2,32}$=2.9, p=0.0712; see Supplementary Figure 2).

2.3 Predicting age-dependent changes to syllable structure, timing, and sequencing

We next investigated the degree to which individual variation in age-dependent changes to song could be predicted by individual variation in context-dependent changes. For the analysis of acoustic similarity (SAP; see Section 2.1), we analyzed the extent to which acoustic similarity scores across contexts (UD_{young} to FD_{young}) correlated with acoustic similarity scores across age (UD_{young} to UD_{older}; Kojima and Doupe, 2011). We did not find that individual variation in acoustic similarity across context could significantly predict individual variation in acoustic similarity over time.

To analyze the predictability of changes to syllable structure, timing, and sequencing, we computed the percent change or difference across social context and correlated this change with the percent change or difference across age. Similar to the analyses presented in Section 2.2, we generally did not observe that the relationship between context- and age-dependent changes varied significantly across tutoring methods (see Experimental

Procedure); consequently, we collapsed birds across methods of tutoring for analysis (n=17; see Supplementary Table 2 for full models).

Variation in the magnitude of age-dependent changes to five song features could be significantly predicted by variation in the extent of context-dependent changes (Figure 7; Spearman's ρ). On a bird-by-bird basis, age-dependent changes to the CV of FF (average percent change across syllables within each bird's song; ρ=0.49, p=0.0470), mean song bout durations (ρ=0.72, p=0.0017), the mean number of motifs per song bout (ρ=0.58, p=0.0174), the mean number of glue syllable repetitions (ρ=0.67, p=0.0041), and the CV of glue syllable repetitions (ρ=0.52, p=0.0406) were significantly and positively correlated with context-dependent changes. For four of these five features, (all but the mean number of glue syllable repetitions), there appeared to be an outlying value that could have influenced the statistical significance of the relationship. Therefore, despite the robustness of Spearman's ρ to outliers, we also analyzed the relationships without the outlying value. The relationships remained significant even after the removal of an outlying value for mean song bout durations (ρ=0.66, p=0.0055) and the mean number of motifs per song bout (ρ=0.50, p=0.0471). However, the relationships became non-significant after the removal of the outlying value for the CV of FF (ρ=0.39, p=0.1405) and the CV of glue syllable repetition (ρ=0.41, p=0.1268).

3. Discussion

The performance of many important behaviors can change over time, and understanding the processes that predict such changes can lend insight into potential mechanisms of behavioral plasticity and control as well as individual differences in behavior. Here, we analyzed age-dependent changes to song in adult zebra finches, compared the extent of this change between birds with different social experiences during development, and assessed the degree to which age-dependent song changes could be predicted by context-dependent song changes. We found that song bout duration and tempo, the mean and variability of the fundamental frequency (FF) of syllables with harmonic structure, and the variability of syllable sequencing changed over time in adult zebra finches and that the extent of these changes did not differ between socially and non-socially tutored birds. Non-socially tutored birds (i.e., birds individually raised and operantly tutored) did, however, demonstrate larger changes to motif composition than socially tutored birds (i.e., birds reared under semi-natural breeding contexts). Further, we discovered that, regardless of development tutoring, individual variation in the nature and extent of age-dependent changes to song bout duration and glue syllable repetition were predicted by context-dependent changes to these features when birds were young adults. This predictability suggests that activity in neural circuits controlling these features in older adult zebra finches may more closely resemble that of young adults singing female-directed song than that of young adults singing undirected song (James and Sakata, 2015).

Broadly speaking, zebra finch song (as well as the songs of other songbirds) is regulated by activity in two canonical brain pathways: the vocal motor pathway (VMP) and the anterior forebrain pathway (AFP; reviewed in Brainard and Doupe, 2013). The VMP is a circuit connecting forebrain areas like HVC (used as a proper name) and the robust nucleus of the arcopallium (RA) with hindbrain and midbrain areas that contain vocal motor and respiratory neurons. The AFP is a basal ganglia-thalamic-cortical circuit that includes the basal ganglia nucleus Area X, the medial portion of the dorsolateral thalamic nucleus (DLM), and the lateral magnocellular nucleus of the anterior nidopallium (LMAN). These pathways have been the focus of research aimed at understanding the neural substrates of song control and plasticity, and many of the brain areas in the VMP and AFP have been implicated in developmental and context-dependent changes to song (Aronov and Fee, 2012; Brainard and Doupe, 2013; Mooney, 2009; Murphy et al., 2017; Sakata and Vehrencamp, 2012; Woolley and Kao, 2015). Therefore, it is important to consider how age- and context-dependent changes to song features, in particular song features that demonstrate correlated changes across age and context, could be mediated by brain areas within the VMP and AFP.

Song bout durations changed in the same direction across social context and with age: adult zebra finches produced longer song bouts when they sang to females ($FD_{young}>UD_{young}$) and when they sang undirected song as older adults ($UD_{older}>UD_{young}$).

However, age- and context-dependent changes in song bout duration were driven by slightly different vocal motor changes. Age- and context-dependent increases in song bout duration were both driven by increases in the number of motifs per song bout, however, context-dependent increases (but not age-dependent increases) were further driven by increases in glue syllable repetition. This suggests that age- and context-dependent changes in song bout durations were mediated by both shared and distinct neural mechanisms. It should be noted that while sequence durations decreased across both age and context, this had little effect on song bout durations since the magnitude of these changes (in milliseconds) were much smaller than the changes to song bout durations (in seconds).

Despite the differences in the magnitude of age- and context-dependent changes to song bout durations, there was a significant relationship between the magnitudes of age- and context-dependent changes: birds that increased their song bout durations more when they sang to females increased their song bout durations more as they grew older. While context- and age-dependent changes to song bout duration have been previously reported in adult zebra finches (e.g., Arnold, 1975; Cooper and Goller, 2006; Johnson et al., 2002; Kao and Brainard, 2006; Scharff and Nottebohm, 1991), little is known about how song bout duration is regulated by activity in song control circuitry. Manipulations of neuromodulators in the brain, including catecholamines and vasotocin, have been found to affect bout durations in various songbird species, and it is possible that context- and age-dependent changes in zebra finches involve similar changes to neuromodulatory systems (Salvante et al., 2009; Voorhuis et al., 1991). Androgens have also been found to affect song bout durations (e.g., Alward et al., 2013; Riters et al., 2000; Smith et al., 1997; but see Arnold, 1975). However, context-dependent changes to song are too rapid to be explained by changes in circulating androgen levels (e.g., Sakata et al., 2008), and it is not known whether circulating androgen concentrations change with age in adult zebra finches. Neurosteroid synthesis of androgens within the song system could contribute to short-term, context-dependent changes and to longer-term, age-dependent changes to bout durations, but little is known about how social context and age alter the expression and activity of steroidogenic enzymes in the adult zebra finch brain.

The relationship between age- and context-dependent changes to song bout durations seem to be mediated by the relationships between age- and context-dependent changes in the number of motifs per bout ("motif repetition") and in glue syllable repetition. In this respect, it is important to think about mechanisms underlying motif and glue syllable repetition in zebra finches. Many studies that document age and social context effects on song bout durations in adult zebra finches acknowledge that changes to song bout durations can be readily explained by changes to motif repetition (Arnold, 1975; Cooper and Goller, 2006; Johnson et al., 2002; Kao and Brainard, 2006; Leblois and Perkel, 2012; Scharff and Nottebohm, 1991). However, little is known the factors that affect the repetition of motifs, especially with regard to age- and context-dependent changes. For example, while lesions or inactivations of LMAN as well as antagonism of D1 receptors in Area X affect the social modulation of spectral features of zebra finch song, these manipulations do not affect motif repetition or the social modulation of motif repetition (Kao and Brainard, 2006; Leblois and Perkel, 2012; Stepanek and Doupe, 2010). It is possible that manipulations of activity or dopamine receptors in areas outside of the AFP could affect the social modulation of song bout duration and motif repetition.

Because age- and context-dependent changes to the repetition of glue syllables are correlated, changes to glue syllable repetition could also contribute to the relationship between age- and context-dependent changes to song bout durations. However, the relationship between context- and age-dependent changes in glue syllable repetition is more complicated than that for song bout durations or motif repetition. This is because, while individual-by-individual variation in context-dependent changes was significantly related to individual-by-individual variation in age-dependent changes, glue syllable repetition increased across social contexts (from UD_{young} song to FD_{young} song) but did not significantly change with age (from UD_{young} song to UD_{older} song). Overall, birds that demonstrated relatively large increases in glue syllable repetition during FD song increased glue syllable repetition with age whereas birds that demonstrated relatively small increases or decreases in glue syllable repetition during FD song decreased glue syllable repetition with age. This suggests that glue syllable repetition does not reliably change in one direction over time but that individual variation in the extent of change across social context can predict the direction of change over time. From a mechanistic perspective, neural activity and dopaminergic tone in the vocal basal ganglia nucleus Area X have been implicated in syllable repetition (Kobayashi et al., 2001; Kubikova et al., 2014; Tanaka et al., 2016), and it is possible that basal ganglia circuitry might be important for both the modulation and plasticity of glue syllable repetition.

Interestingly, whereas age-dependent changes to the repetition of glue syllables were predicted by social context-dependent changes, age-dependent changes to the repetition of introductory notes, which often acoustically resemble glue syllables, was not predicted by social context-dependent changes. This suggests that the mechanisms underlying vocal repetition before song initiation could be distinct from those underlying vocal repetition during song. Consistent with this notion, experimental expression of a mutant form of the Huntingtin gene in Area X leads to increases in glue syllable repetition but not introductory note repetition (Tanaka et al., 2016). In addition, introductory notes are repeated more than glue syllables (Rajan and Doupe, 2013), and there tends to be greater variability in the repetition of glue syllables than of introductory notes (Figure 5).

Although some song features change in similar ways across social context and age, other age-dependent changes were not accurately predicted by examining context-dependent changes. For example, the mean and variability of introductory note repetition and the variability in song bout durations increased across social context but tended to remain the same or decrease across age. The mean FF of syllables with flat, harmonic structure and the variability of syllable sequencing decreased with age but did not change with social context. Not surprisingly, individual variation in the social modulation of these features did not correlate with individual variation in age-dependent changes. In addition, while context-dependent changes to overall acoustic structure of motifs can predict age-dependent changes in developing birds (Kojima and Doupe, 2011), we did not find the same effect in adult birds.

Broadly speaking, context-dependent changes were more likely to predict age-dependent changes in temporal features of zebra finch song than spectral features. A number of studies have reported distinct neural contributions to temporal and spectral features of song (Ali et al., 2013; Aronov and Fee, 2012; Hampton et al., 2009; Long and Fee, 2008; Murphy et al., 2017; Stepanek and Doupe, 2010). For example, temperature manipulations in the vocal motor nucleus HVC have a more substantive impact on song tempo than on the spectral structure of syllables (Long and Fee, 2008). Conversely, lesions and inactivations of LMAN have more reliable and robust effects on the spectral structure of syllables than on song tempo or syllable sequencing (Hampton et al., 2009; Stepanek and Doupe, 2010; Ali et al. 2013). Our data suggest that context-dependent changes to activity in circuits that preferentially regulate temporal song features (e.g., the VMP) are more predictive of age-dependent changes than context-dependent changes to activity in circuits that preferentially regulate spectral features (e.g., the AFP). Further investigations into how social context differentially affects these neural circuits will be particularly useful in providing insight into the mechanisms of age-dependent changes.

Consistent with previous findings reporting that birds that have impoverished social experiences during development demonstrate more substantive song plasticity in adulthood (see also Derégnaucourt et al., 2013; Jones et al., 1996), we found that non-socially tutored birds demonstrated larger changes to the acoustic structure of motifs over time than socially tutored birds. For example, non-socially tutored birds dropped syllables from their motif and integrated “novel” syllables into their motif over time, phenomena that were not observed in our socially tutored birds (Supplementary Figure 1). It is difficult to identify precisely how novel syllables were integrated into the song, but some of these syllables may have originated from glue syllables that became regularized in structure and sequencing and, thus, integrated into the motif (Lipkind et al., 2017).

Regardless of the mechanism underlying motif changes in non-socially tutored birds, one possibility is that the adult song changes of these birds reflect a delayed maturational process. Birds that have been deprived of social interactions during development display an elongated ‘sensitive’ period’ and slower vocal motor development (Derégnaucourt et al., 2013; Eales, 1985); consequently, changes to the motifs of non-socially tutored birds could reflect a convergence toward the tutor song. However, we did not find that the motifs of non-socially tutored adults became more similar to the tutor song over time. Furthermore, the variability of syllable structure and sequencing, song features that serve as benchmarks for song development and maturation (Kelly and Sober, 2014; Scharff and Nottebohm, 1991; Tchernichovski et al., 2001), were not different between socially or non-socially tutored birds in this study. As such, these results suggest that young adult zebra finches analyzed here underwent song changes that seem to be distinct from developmental song plasticity during juvenile development. Relatedly, while the FD songs of juvenile zebra finches have been found to predict maturational changes in motif structure (Kojima and Doupe, 2011), we did not observe such predictability for our non-socially tutored adults, supporting the notion that the birds in our study were not in an extended developmental state.

These observations in the zebra finch provide the opportunity for comparisons of vocal plasticity and predictability with the closely related Bengalese finch (James and Sakata, 2015). With regard to age-dependent changes, the variability of syllable sequencing decreased and song tempo increased over time in both species. However, the variability of syllable structure (fundamental frequency) decreased with age in the zebra finch but not in the Bengalese finch. Age-dependent changes to the variability of syllable sequencing could be predicted by social context-dependent changes in both species, but age-dependent changes in syllable timing and the variability of syllable structure were predicted by context-dependent changes in Bengalese finches but not in zebra finches. Together, this comparison highlights that the predictability of some song features is consistent across species but also reveals a number of species differences in the predictability of song plasticity.

Such species variation could be rooted in species variation in song organization as well as the neural control of song. One striking difference in song organization between the species is that syllable sequencing is more variable in Bengalese finches than in zebra finches (Fujimoto et al., 2011; Matheson et al., 2016; Murphy et al., 2017; Okanoya, 2004;

Sakata and Brainard, 2006; Warren et al., 2012), a behavioral difference that is likely linked to variation in the structure and function of focal neural circuits. These species also differ in the duration of silent intervals between syllables, with Bengalese finch song consisting of longer gaps than zebra finch song (Araki et al., 2016). While there are many similarities in the neural control of song between zebra and Bengalese finches, species differences in predictability could be related to species differences in the neural control of song tempo or syllable structure. For example, LMAN lesions seem to have larger impacts on song tempo in zebra finches than Bengalese finches (Hampton et al., 2009; Kao and Brainard, 2006). In addition, the species differ in how neurons in the sensorimotor nucleus HVC (used as proper name) respond to auditory stimuli, to perturbations of auditory feedback, and to changes in the social context of song (Hamaguchi et al., 2014; Hampton et al., 2009; Matheson et al., 2016; Murphy et al., 2017). Consequently, species differences in the AFP and VMP could underlie variation in the predictability of age-dependent changes to song.

Overall, these results demonstrate that performance features in adult zebra finch song change over time, that changes in these performance features were not affected by the nature of tutoring experiences during development, and that a subset of these changes could be predicted by analyzing context-dependent changes. The increases in song bout duration and consistency suggest that song performance improves with age (Podos et al., 2009; Sakata and Vehrencamp, 2012; Woolley, 2019). Furthermore, the predictability of age-dependent changes in song bout duration and in motif and glue syllable repetition by social context-dependent changes suggests that neural activity during the production of female-directed courtship song could provide predictive insight into the nature of change in neural activity associated with age and vocal practice (James and Sakata, 2015; Sakata and Vehrencamp, 2012). Further, because the variability of a number of song features changes over time and across contexts and because variability is the substrate upon with plasticity mechanisms operate (Brainard and Doupe, 2013; Dhawale et al., 2017), subsequent investigations could relate the extent of context- and age-dependent changes to song to experimentally-driven plasticity in song features (Ali et al., 2013; Andalman and Fee, 2009; Hisey et al., 2018; Roberts et al., 2017).

4. Experimental procedure

4.1 Animals and recording

Zebra finches (*Taeniopygia guttata*; n=17) were bred at McGill University. Eight birds were socially tutored by an adult in our breeding colony, whereas nine birds were individually raised and tutored via song played from a speaker (non-socially tutored; see James et al., 2018; James and Sakata, 2017). Socially tutored birds were raised with and tutored by their parents in a semi-natural breeding colony until 60 days post-hatch (dph), and thereafter housed in same-sex cages. Non-socially tutored birds in this study represent a subset of birds from a previous study in which biological predisposition in vocal learning were documented (James and Sakata, 2017). Briefly, non-socially tutored birds were raised by both parents until <7 days of age in sound-attenuating chambers ("soundbox"; TRA Acoustics, Ontario, Canada), after which fathers were removed from the nest. When juvenile males were nutritionally independent (30-40 days post-hatch), they were housed individually in a soundbox and operantly tutored with synthesized songs for >30 days. Song stimuli consisted of five zebra finch song elements (termed "syllables", contiguous epochs of sound separated by short silent gaps (>5 ms)) that are commonly observed in zebra finch song. Syllables in song stimuli were bandpass-filtered (0.3–10 kHz), normalized by their maximum amplitude, ramped to avoid sharp onsets and offsets in amplitude, and presented in randomized sequences (James and Sakata, 2017). The same set of syllables was used to tutor all non-socially tutored birds in this study. Non-socially tutored birds remained housed individually until they were 4 months old, at which songs were recorded for analysis (see below).

All birds were housed on a 14-h light:10-h dark cycle and provided food and water ad libitum. All procedures were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

For all song recordings, birds were housed individually in sound-attenuating chambers (TRA Acoustics, Ontario, Canada). Song was recorded using an omnidirectional microphone (Countryman Associates, Menlo Park, CA) positioned above the male's cage. Computerized, song-activated recording systems were used to detect and digitize song (Sound Analysis

Pro (SAP), v 1.04, digitized at 44.1 kHz; http://ofer.sci.ccny.cuny.edu/sound_analysis_pro). Recorded songs were digitally filtered (0.3–8 kHz) for offline analysis using software custom written in the Matlab programming language (MathWorks, Natick, MA).

We investigated context-dependent changes to song in all birds at 4 months of age ("young adults"). To assess context-dependent changes to song, we collected renditions of undirected (UD) and female-directed (FD) song following a protocol identical to James et al (2018). Males were moved into a sound-attenuating chamber at least one night before collecting UD and FD song. UD songs are spontaneously produced when birds are alone, whereas FD songs are produced during courtship interactions with females. To collect FD songs, we briefly placed a cage with a female next to the experimental male's cage and monitored his behavior via video. FD songs are readily distinguishable from UD songs because they are produced after a male approaches or faces a female, accompanied by a courtship dance (e.g., pivoting body from side to side) and associated with the fluffing of the male's plumage (Morris, 1954; Zann, 1996). Only songs that were accompanied by at least two of the above behaviors were categorized as FD songs. FD songs were almost always produced soon after the introduction of a female, and we removed females after 30–60 s to ensure that FD songs were produced at a short latency following exposure to a female. The interval between exposures to females was 4–5 min, which allowed for the collection of UD song between female exposures. However, not all birds produced renditions of UD song between female presentations; therefore, UD songs were also recorded for 30 min before and after the testing session. The number of UD (36.3 ± 3.3) and FD (28.8 ± 3.3) song bouts collected was comparable with previous studies (Hampton et al., 2009; Heinig et al., 2014; James et al., 2018; James and Sakata, 2015, 2014; Matheson et al., 2016; Sakata et al., 2008; Sakata and Brainard, 2009). After collecting UD and FD songs, all birds were group-housed in same-sex cages in the bird colony.

To assess age-dependent changes to song, each bird was re-recorded at least 4.5 months following the collection of UD and FD songs (19.7 ± 2.3 months). Only UD songs were recorded at this time (UD_{older} songs; 38.2 ± 2.0 song bouts).

4.2 Song analysis

An example of a zebra finch song is provided in Fig. 1. For purposes of description and analysis, we use the term "syllable" to refer to individual acoustic elements that are separated from each other by >5 ms of silence (James and Sakata, 2015, 2014; Warren et al., 2012). Identical to previous studies, we manually labeled syllables based on visual inspection of spectrograms following amplitude-based syllable segmentation in Matlab (e.g., Heinig et al., 2014; James and Sakata, 2015, 2014; Long and Fee, 2008; Matheson et al., 2016; Sakata et al., 2008; Stepanek and Doupe, 2010; Warren et al., 2012).

Zebra finch songs are organized into bouts, epochs of song vocalizations containing at least one complete motif and in which syllables are maximally separated by 500 ms of silence. Song bouts are generally preceded by the repetition of a spectrally simple element termed an "introductory note" and consist of the repetition of a stereotyped syllable sequence (termed the "motif"; Sossinka and Böhner, 1980; Zann, 1996). Each zebra finch's motif is readily identifiable because sequencing is relatively stereotyped within the motif and because inter-syllable durations (i.e., "gaps") within a motif are generally shorter than gaps outside the motif. All birds in this analysis produced songs with conspicuous motif structure. Between motifs, individuals will occasionally produce one or more "glue syllables", which are often, but not always, spectrally similar to introductory notes (Bruno and Tchernichovski, 2017; Horita et al., 2008; Kubikova et al., 2014; Yu and Margoliash, 1996; Zann, 1996). Glue syllables are not considered part of the motif because their production is more variable than the syllables within the motif (i.e., glue syllables may or may not be produced between motifs).

To measure context- and age-dependent changes in the repetition of introductory notes and glue syllables (James et al., 2018; Kao and Brainard, 2006; Sossinka and Böhner, 1980), we computed the mean number of times introductory notes or glue syllables were repeated across each instance of repetition (within each experimental context: UD_{young}, FD_{young}, and UD_{older}), as well as the coefficient of variation (CV = standard deviation/mean) in the number of times introductory notes or glue syllables were repeated across instances of repetition.

To analyze changes in overall motif acoustic structure, we used the similarity comparison feature in SAP to compute the spectral similarity between motifs (Tchernichovski et al., 2000). Specifically, we cropped 10 randomly selected motifs per bird per condition (UD_{young}, FD_{young} and UD_{older}). For one set of analyses, we conducted all pairwise comparisons between UD_{young} and FD_{young} motifs (n=100 comparisons) to assess similarity across context, and all pairwise comparisons between UD_{young} and UD_{older} to assess similarity across age. The similarity calculation uses one motif as the "template" and asks how much spectral content in the template is contained in the comparison song (referred to as an "asymmetric" comparison). To calculate overall similarity across contexts and age, we ran comparisons with UD_{young} motifs as the template for both the age and context comparisons. We averaged all the similarity scores to yield an overall similarity score for each bird's motif across age and context. Furthermore, to understand whether the substantial changes to motifs in non-socially tutored birds resulted in motifs that were more acoustically similar to the tutor song stimulus, we ran SAP similarity analyses comparing each of the UD_{young} and UD_{older} motifs for each bird to a motif of the tutor stimuli (tutor song as the template). We averaged all the similarity scores to yield an overall similarity score for each bird's UD_{young} and UD_{older} motifs to the tutor song.

To analyze experimental changes to song bout durations, we measured the duration from the onset of the first syllable of the first motif to the offset of the last syllable in the song bout. We also counted the number of complete motifs produced in each song bout. We computed the average song bout duration and number of motifs per song bout in each experimental period (UD_{young}, FD_{young}, and UD_{older}).

To examine changes to syllable structure, we analyzed the fundamental frequency (FF) of syllables that had harmonic structure (e.g., syllable "e" in Figure 1; n=26 syllables in 17 birds). The FF of such syllables is tightly regulated by the nervous system and represents an important metric for song development and performance (Bouchard and Brainard, 2013; Kao and Brainard, 2006; Sakata et al., 2008; Sakata and Vehrencamp, 2012). To compute the FF, we calculated the autocorrelation of a segment of the sound waveform and defined the FF as the distance, in Hz, between the zero-offset peak and the highest peak in the autocorrelation function. To improve the resolution and accuracy of frequency estimates, we performed a parabolic interpolation of the peak of the autocorrelation function (de Cheveigné and Kawahara, 2002). Each rendition of a syllable was visually screened to ensure that we analyzed only examples devoid of sound artifacts that could affect FF calculations (e.g., sound of movement, female calls in background; 113.0 ± 6.3 renditions per syllable). For each syllable, we computed the mean and variability of FF across renditions, two aspects of song that change over time and across social contexts (e.g., Hampton et al., 2009; Kao and Brainard, 2006; Kelly and Sober, 2014; Sakata et al., 2008). The mean FF of syllables that we measured ranged from 0.4 to 4.5 kHz. Because of this wide range of values, we also calculated the percent change in mean FF across social contexts and ages for data visualization and analysis.

Changes to song tempo (sequence durations) were analyzed using methods comparable to previous studies (Cooper et al., 2012; James and Sakata, 2015, 2014). Specifically, we measured the full duration of each bird's motifs from the onset of the first syllable in the motif to the offset of the last syllable in the motif (range: 200 – 746 ms). Because recording conditions could vary between young and older recordings, we computed sequence durations after normalizing sound envelopes. For this, we extracted the waveform of the sequence and then rectified, smoothed (5 ms square window), and resampled (1 kHz) the waveform. Thereafter, we normalized the envelope between zero and one using the minimum and peak values. After this normalization, the amplitude traces across ages and contexts were comparable, and we applied a common threshold on these normalized traces to find syllable onsets and offsets for each recording). Because sequence durations increase as the song progresses (Chi and Margoliash, 2001; Cooper and Goller, 2006; Glaze and Troyer, 2006) and because song bout durations change as a function of social context and age (James et al., 2018; Kao and Brainard, 2006; see also Results), we restricted our analysis to the first occurrence of the sequence in each song bout. Given the wide range of motif durations, we also calculated the percent change in durations across social contexts and ages for data visualization and analysis.

To analyze experimental changes to syllable sequencing within a motif, we measured commonly used metrics for assessing zebra finch song (e.g., Horita et al., 2008; Scharff and Nottebohm, 1991; Tanaka et al., 2016; Zevin et al., 2004). Sequence linearity is calculated as the total number of transition types divided by the total number of syllable types for each bird. Sequence consistency is a related measure that is calculated as the total number of 'typical' transitions divided by the total number of transitions across all songs analyzed for each individual. Here, the most common transition from each syllable type is defined as the 'typical' transition. Finally, the weighted transition entropy for syllable sequencing was computed for each bird. The transition entropy for each syllable type in a bird is measured with the following formula:

transition entropy = Σ-pi*log2(pi)

where the sum is over all transitions produced from the syllable, and pi is the probability of the ith transition across all renditions of the syllable. Syllables with more variable transitions will have higher entropy values.

Context-dependent changes have been analyzed and reported for these birds in a previous publication (James et al., 2018). However, age-dependent changes and the degree to which context-dependent changes predict age-dependent changes have not been reported.

4.3 Statistical analyses

We analyzed the similarity of UD_{older} songs and FD_{young} songs to UD_{young} song (SAP similarity score) using a t-test or mixed effects model. For the mixed effects model, we used a full-factorial mixed effect model with "Method of Tutoring" (social vs. non-social) and "Similarity Comparison" [context (UD_{young}-FD_{young}) vs. age (UD_{young}-UD_{older})] as the independent factors, and Bird ID as the random factor (Figure 2C). We analyzed the degree to which changes in motif structure in non-socially tutored birds reflected a convergence toward the tutor song using a mixed effects model with age of recording (UD_{young} vs UD_{older}) as the independent factor, and Bird ID as the random factor.

For the analysis of song features, we first simultaneously analyzed context- and age-dependent changes to song features (n=17 birds). For this, we ran full-factorial, mixed-effects models with method of tutoring (social vs. non-social) and recording condition (UD_{young}, FD_{young}, or UD_{older}) as the independent variables, and BIRD ID as a random effect (analogous to a repeated-measures design). Because we measured multiple syllables per bird for our analyses of fundamental frequency (mean and CV of FF), we included SYLLABLE ID nested within BIRD ID in these models. The method of tutoring only affected the degree to which syllables in the motif changed over time (see Results). The method of tutoring did not significantly affect or significantly interact with recording condition to affect song features for the remaining song features (mixed-effects models; Supplementary Table 1). Consequently, for simplicity, we only report the effect of condition within the model in our Results section (see Supplementary Table 1 for the full models). We also analyzed the degree to which the duration between the UD_{young} and UD_{older} recordings explained variation in the magnitude of age-dependent changes to song and found no significant effect for any of the features measured (Spearman's ρ or mixed effects models; $p>0.10$ for all). Tukey's HSD tests were used for post-hoc comparisons.

To analyze the relationship between individual variation in context- and age-dependent changes, we computed and correlated the percent change from UD_{young} to FD_{young} song (context-dependent change) and from UD_{young} to UD_{older} song (age-dependent change). In cases in which multiple syllables were measured per bird (FF), we computed the average percent change across syllables. For measures of sequencing within the motif (linearity, consistence, and entropy), we computed and correlated the change (difference) across contexts and age. A preliminary analysis indicated that the relationship between context- and age-dependent changes did not differ between socially or non-socially tutored birds (interaction term in ANCOVA; Supplementary Table 2), so we simplified our analysis by excluding method of tutoring. We used Spearman's ρ, a non-parametric correlation coefficient, to assess the degree to which context-dependent changes could predict age-dependent changes because it is more robust to outlying values and more sensitive to nonlinear relationships than parametric correlations (e.g., Pearson's product-moment correlation).

Analyses were conducted using JMP 10 (SAS Institute, Cary, NC), and $\alpha = 0.05$ for all tests.

5. Acknowledgements

We thank J. Dai and Y. Chen for help with data collection. We also thank S.C. Woolley for feedback on the manuscript. Research was supported by a NSERC Discovery Grant (J.T.S.; 2016-05016), a FQRNT Projet de Recherche en équipe (J.T.S.; PR-189949) and a Heller Award (L.S.J.; Biology, McGill University).

6. References

Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80: 494–506, 2013.

Alward BA, Balthazart J, Ball GF. Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate*. Proc. Natl. Acad. Sci.* 110: 19573–19578, 2013.

Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl. Acad. Sci.* 106: 12518–12523, 2009.

Araki M, Bandi MM, Yazaki-Sugiyama Y. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–1287, 2016.

Arnold AP. The effects of castration on song development in zebra finches (*Poephila guttata*). *J. Exp. Zool.* 191: 261–277, 1975.

Aronov D, Fee MS. Natural changes in brain temperature underlie variations in song tempo during a mating behavior. *PLoS One* 7: 1–10, 2012.

Beecher MD, Brenowitz EA. Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* 20: 143–149, 2005.

Bolhuis JJ, Okanoya K, Scharff C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11: 747–759, 2010.

Bouchard KE, Brainard MS. Neural encoding and integration of learned probabilistic sequences in avian sensory-motor circuitry. *J Neurosci* 33: 17710–17723, 2013.

Brainard MS, Doupe AJ. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404: 762–766, 2000.

Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu. Rev. Neurosci.* 36: 489–517, 2013.

Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, Doupe AJ. Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J. Neurosci.* 21: 2501–2517, 2001.

Bruno JH, Tchernichovski O. Regularities in zebra finch song beyond the repeated motif. *Behav. Processes in press*, 2017.

Chen Y, Matheson LE, Sakata JT. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci.* 113: 6641–6646, 2016.

Chi Z, Margoliash D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32: 899–910, 2001.

Cooper BG, Goller F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J. Neurophysiol.* 95: 3798–3809, 2006.

Cooper BG, Méndez JM, Saar S, Whetstone AG, Meyers R, Goller F. Age-related changes in the Bengalese finch song motor program. *Neurobiol. Aging* 33: 564–568, 2012.

Costa RM, Cohen D, Nicolelis MAL. Differential corticostriatal plasticity during fast and slow motor skill learning in mice. *Curr. Biol.* 14: 1124–1134, 2004.

Dalla C, Shors TJ. Sex differences in learning processes of classical and operant conditioning. *Physiol. Behav.* 97: 229–238, 2009.

de Cheveigné A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. *J. Acoust. Soc. Am.* 111: 1917–1930, 2002.

Derégnaucourt S, Poirier C, Van der Kant A, Van der Linden A, Gahr M. Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* 107: 210–218, 2013.

Dhawale AK, Smith MA, Ölveczky BP. The role of variability in motor learning. *Annu. Rev. Neurosci.* 40: 479–498, 2017.

Doupe AJ, Kuhl PK. Birdsong and humman speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* 22: 567–631, 1999.

Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* 10: 732–739, 2000.

Drake C, Palmer C. Skill acquisition in music performance: relations between planning and temporal control. *Cognition* 74, 1–32, 2000.

Eales LA. Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim. Behav.* 33: 1293–1300, 1985.

Ericsson KA. Deliberate practice and acquisition of expert performance: A general overview. *Acad. Emerg. Med.* 15: 988–994, 2008.

Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100: 363–406, 1993.

Fujimoto H, Hasegawa T, Watanabe D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J. Neurosci.* 31: 10023–10033, 2011.

Funabiki Y, Funabiki K. Song retuning with tutor model by adult zebra finches. *Dev. Neurobiol.* 68: 645–655, 2008.

Glaze CM, Troyer TW. Temporal structure in zebra finch song: Implications for motor coding. *J. Neurosci*: 26, 991–1005, 2006.

Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C. Brain structure predicts the learning of foreign speech sounds. *Cereb. Cortex* 17: 575–582, 2007.

Golestani N, Paus T, Zatorre RJ. Anatomical correlates of learning novel speech sounds. *Neuron* 35: 997–1010, 2002.

Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J. Neurophysiol.* 101: 3235–3245, 2009.

Heinig A, Pant S, Dunning JL, Bass A, Coburn Z, Prather JF. Male mate preferences in mutual mate choice: finches modulate their songs across and within male-female interactions. *Anim. Behav.* 97: 1–12, 2014.

Herholz SC, Zatorre RJ. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76: 486–502, 2012.

Hisey E, Kearney MG, Mooney R. A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat. Neurosci.* 21: 589–597, 2018.

Horita H, Wada K, Jarvis ED. Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *Eur. J. Neurosci.* 28: 2519–2532, 2008.

James LS, Dai JB, Sakata JT. Ability to modulate birdsong across social contexts develops without imitative social learning. *Biol. Lett.* 14, 2018.

James LS, Sakata JT. Learning biases underlie "universals" in avian vocal sequencing. *Curr. Biol.* 27: 3676–3682.e4, 2017.

James LS, Sakata JT. Predicting plasticity: acute context-dependent changes to vocal performance predict long-term age-dependent changes. *J. Neurophysiol.* 114: 2328–2339, 2015.

James LS, Sakata JT. Vocal motor changes beyond the sensitive period for song plasticity. *J. Neurophysiol.* 112: 2040–2052, 2014.

Johnson F, Soderstrom K, Whitney O. Quantifying song bout production during zebra finch sensory-motor learning suggests a sensitive period for vocal practice. *Behav. Brain Res.* 131: 57–65, 2002.

Jones AE, ten Cate C, Slater PJB. Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* 110: 354–369, 1996.

Kao MH, Brainard MS. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J. Neurophysiol.* 96: 1441–1455, 2006.

Keith N, Ericsson KA. A deliberate practice account of typing proficiency in everyday typists. *J. Exp. Psychol. Appl.* 13: 135–145, 2007.

Kelly CW, Sober SJ. A simple computational principle predicts vocal adaptation dynamics across age and error size. *Front. Integr. Neurosci.* 8: 1–9, 2014.

Kobayashi K, Uno H, Okanoya K. Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. *Neuroreport* 12: 353–358, 2001.

Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proc. Natl. Acad. Sci. U. S. A.* 108: 1687–92, 2011.

Kojima S, Doupe AJ. Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J. Neurophysiol.* 98: 2099–109, 2007.

Kosten TA, Kim JJ, Lee HJ. Early life manipulations alter learning and memory in rats. *Neurosci. Biobehav. Rev.* 36: 1985–2006, 2012.

Kubikova L, Bosikova E, Cvikova M, Lukacova K, Scharff C, Jarvis ED. Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. *Sci. Rep.* 4, 2014.

Landi SM, Baguear F, Della-Maggiore V. One week of motor adaptation induces structural changes in primary motor cortex that predict long-term memory one year later. *J. Neurosci*. 31, 11808–11813, 2011.

Leblois A, Perkel DJ. Striatal dopamine modulates song spectral but not temporal features through D1 receptors. *Eur. J. Neurosci.* 35: 1771–1781, 2012.

Leblois A, Wendel BJ, Perkel DJ. Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *J. Neurosci.* 30: 5730–5743, 2010.

Lipkind D, Zai AT, Hanuschkin A, Marcus GF, Tchernichovski O, Hahnloser RHR. Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nat. Commun.* 8: 1247, 2017.

Livingston FS, White SA, Mooney R. Slow NMDA-EPSCs at synapses critical for song development are not required for song learning in zebra finches. *Nat. Neurosci.* 3: 482–488, 2000.

Lombardino AJ, Nottebohm F. Age at deafening affects the stability of learned song in adult male zebra finches. *J. Neurosci.* 20: 5054–5064, 2000.

Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456: 189–194, 2008.

Makino H, Hwang EJ, Hedrick NG, Komiyama T. Circuit mechanisms of sensorimotor learning. *Neuron* 92: 705–721, 2016.

Matheson LE, Sun H, Sakata JT. Forebrain circuits underlying the social modulation of vocal communication signals. *Dev. Neurobiol.* 76, 47–63, 2016.

Mooney R. Neural mechanisms for learned birdsong. *Learn. Mem.* 16: 655–669, 2009.

Morris D. The reproductive behaviour of the zebra finch (*Poephila guttata*), with special reference to pseudofemale behaviour and displacement activities. *Behaviour* 6: 271–322, 1954.

Morrison RG, Nottebohm F. Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *J. Neurobiol.* 24: 1045–1064, 1993.

Murphy K, James LS, Sakata JT, Prather JF. Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. *J. Neurophysiol.* 118: 800–816, 2017.

Okanoya K. Song syntax in Bengalese finches: proximate and ultimate analyses, in: *Advances in the Study of Behavior*. Academic Press, New York, NY, 2004, 297–346.

Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS. Growth and splitting of neural sequences in songbird vocal development. *Nature* 528: 352–357, 2015.

Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3: 0902–0909, 2005.

Ota N, Soma M. Age-dependent song changes in a closed-ended vocal learner: elevation of song performance after song crystallization. *J Avian Biol* 45: 566–573, 2014.

Podos J, Lahti DC, Moseley DL. Vocal performance and sensorimotor learning in songbirds. *Adv Study Behav* 40: 159–195, 2009.

Prather J, Okanoya K, Bolhuis JJ. Brains for birds and babies: neural parallels between birdsong and speech acquisition. *Neurosci Biobehav Rev*, 2017.

Pytte CL, Gerson M, Miller J, Kirn JR. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Rajan R, Doupe AJ. Behavioral and neural signatures of readiness to initiate a learned motor sequence. *Curr Biol* 23: 87–93, 2013.

Riters L V., Eens M, Pinxten R, Duffy DL, Balthazart J, Ball GF. Seasonal changes in courtship song and the medial preoptic area in male European starlings (*Sturnus vulgaris*). *Horm Behav* 38: 250–261, 2000.

Rivera-Gutierrez HF, Pinxten R, Eens M. Tuning and fading voices in songbirds: age-dependent changes in two acoustic traits across the life span. *Anim Behav* 83: 1279–1283, 2012.

Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R. Identification of a motor-to-auditory pathway important for vocal learning. *Nat. Neurosci.* 20: 978–986, 2017.

Sakata JT, Brainard MS. Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *J Neurophysiol* 102: 2485–2497, 2009.

Sakata JT, Brainard MS. Real-time contributions of auditory feedback to avian vocal motor control. *J. Neurosci.* 26: 9619–9628, 2006.

Sakata JT, Crews D. Embryonic temperature shapes behavioural change following social experience in male leopard geckos, *Eublepharis macularius. Anim Behav* 66: 839–846, 2003.

Sakata JT, Crews D. Developmental sculpting of social phenotype and plasticity. *Neurosci Biobehav Rev* 28: 95–112, 2004.

Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99: 1700–11, 2008.

Sakata JT, Vehrencamp SL. Integrating perspectives on vocal performance and consistency. *J Exp Biol* 215: 201–209, 2012.

Salvante KG, Racke DM, Campbell CR, Sockman KW. Plasticity in singing effort and its relationship with monoamine metabolism in the songbird telencephalon. *Dev Neurobiol* 14, 2009.

Scharff C, Nottebohm FN. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11: 2896–2913, 1991.

Schmidt RA. *Motor control and learning*. Champaign, IL: *Human Kinetics*, 1988.

Shors TJ, Miesegaes G. Testosterone in utero and at birth dictates how stressful experience will affect learning in adulthood. *Proc Natl Acad Sci* 99: 13955–13960, 2002.

Smith GT, Brenowitz EA, Beecher MD, Wingfield JC. Seasonal changes in testosterone, neural attributes of song control nuclei, and song structure in wild songbirds. *J Neurosci* 17: 6001–10, 1997.

Sossinka R, Böhner J. Song types in the zebra finch. *Z Tierpsychol* 53: 123–132, 1980.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Tanaka M, Singh Alvarado J, Murugan M, Mooney R. Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. *Proc Natl Acad Sci* 113: E1720–E1727, 2016.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

Tchernichovski O, Nottebohm F, Ho C, Pesaran B, Mitra P. A procedure for an automated measurement of song similarity. *Anim Behav* 59: 1167–1176, 2000.

Teramitsu I, White SA. FoxP2 regulation during undirected singing in adult songbirds. *J Neurosci* 26: 7390–7394, 2006.

Ungerleider LG, Doyon J, Karni A. Imaging brain plasticity during motor skill learning. *Neurobiol Learn Mem* 78: 553–564, 2002.

Voorhuis TAM, De Kloet ER, De Wied D. Effect of a vasotocin analog on singing behavior in the canary. *Horm Behav* 25: 549–559, 1991.

Warren TL, Charlesworth JD, Tumer EC, Brainard MS. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32: 15414–15425, 2012.

Woolley SC. Dopaminergic regulation of vocal-motor plasticity and performance. *Curr Opin Neurobiol* 54: 127–133, 2019.

Woolley SC, Doupe AJ. Social context-induced song variation affects female behavior and gene expression. *PLoS Biol.* 6: 0525–0537, 2008.

Woolley SC, Kao MH. Variability in action: Contributions of a songbird cortical-basal ganglia circuit to vocal motor learning and control. *Neuroscience* 296: 39–47, 2015.

Yu AC, Margoliash D. Temporal hierarchical control of singing in birds. *Science* 273: 1871–1875, 1996.

Zann RA. The zebra finch: A synthesis of field and laboratory studies. Oxford, UK: Oxford University Press, 1996.

Zevin JD, Seidenberg MS, Bottjer SW. Limits on reacquisition of song in adult zebra finches exposed to white noise. *J Neurosci* 24: 5849–5862, 2004.

Figures

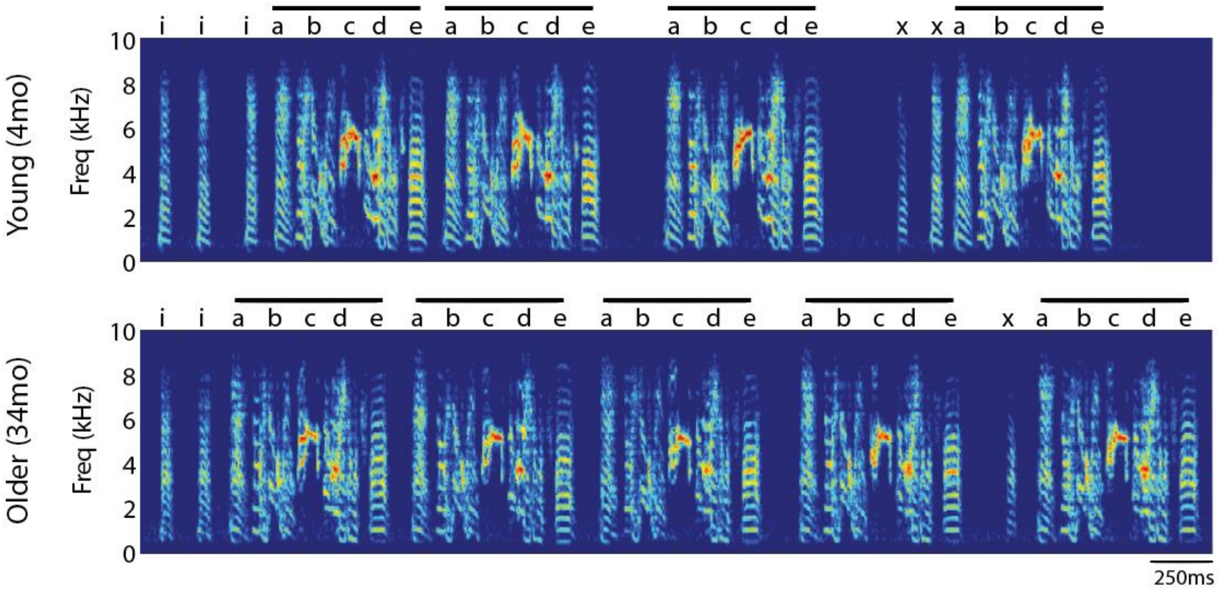


Figure 1: Examples of young (top) and older (bottom) adult recordings of a socially tutored zebra finch. Above each spectrogram (time on the x-axis, frequency on the y-axis, color indicates intensity) is the arbitrary labelling scheme for this bird. In our analyses of gross song structure (Section 2.1) we focused on the motif (e.g., 'abcde' for this bird). We analyzed a number of different aspects of song in our fine-scale analyses of zebra finch song (Section 2.2). For example, we analyzed the fundamental frequency (FF) of syllables with flat, harmonic structure (e.g., 'e'), the number of introductory notes that precede song ('i'), as well as the number of glue syllables repeated between motifs ('x'), and motif and song bout durations.

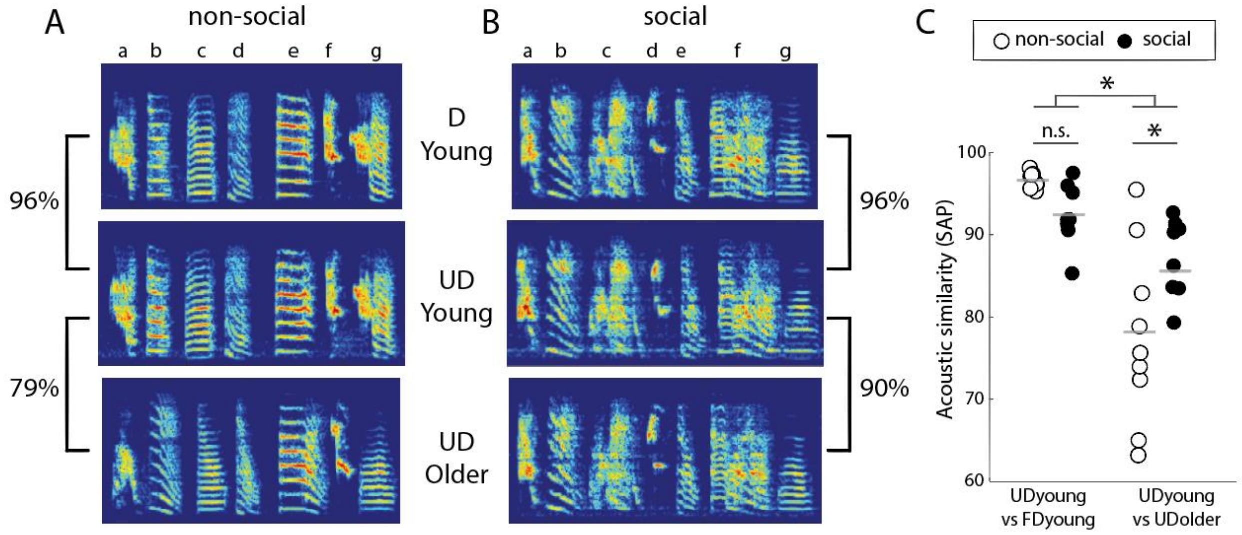


Figure 2: Examples of age- and context-dependent changes to motifs. A-B. Examples of an individual bird's motifs in each condition: Undirected (UD) song as a young adult (middle), female-directed (FD) song as a young adult (top) and UD song as an older adult (bottom). On the top portion of each spectrogram are the labels assigned to each syllable (the labels themselves are arbitrary). Percentages indicate the quantitative acoustic similarity across context (UD young to D young) and age (UD young to UD older; see Experimental Procedure). The bird in (A) was non-socially tutored while the bird in (B) was socially tutored. The spectral structure of syllables in the motif of the non-socially tutored bird (A) remained quite consistent across context but noticeably changed with age. In addition, the bird in this example dropped the syllable "g" from his motif and replaced it with a novel syllable. In contrast, all syllables in the motif of the socially tutored bird (B) remain quite similar across context and with age. C. The extent to which the acoustic structure of FD_{young} and UD_{older} song resembles UD_{young} song is reflected in the acoustic similarity score (Sound Analysis Pro; see Experimental Procedure). Acoustic similarity was higher across contexts (UD_{young} to FD_{young}) than across ages (UD_{young} to UD_{older}). Furthermore, acoustic similarity across age was lower for non-socially tutored birds than for socially tutored birds, with no difference between groups in acoustic similarity across context. Grey bars indicate means. "*" denotes contrasts with $p<0.05$.

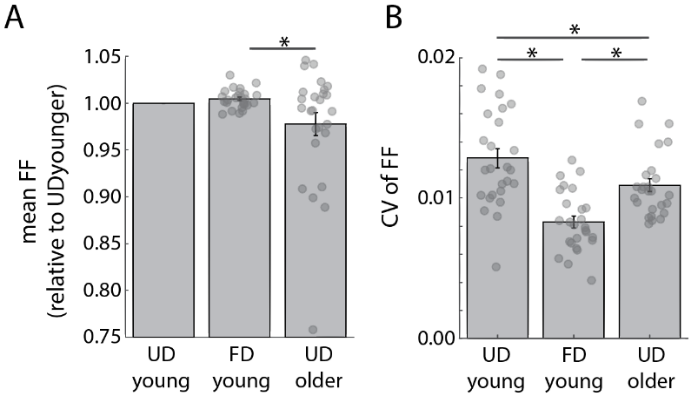


Figure 3: Age- and context-dependent changes to syllable structure. A. The mean FF of syllables with flat, harmonic structure was significantly lower for UD_{older} song than for FD_{young} song. While raw data were analyzed, plotted are values normalized to the FF of the syllable during UD_{young} song because of the wide range of FF values between syllables. B. The variability of the FF of harmonic syllables was highest for UD_{young} song, intermediate for UD_{older} song, and lowest for FD_{young} song ($p<0.05$ for all contrasts). "*" denotes contrasts with $p<0.05$.

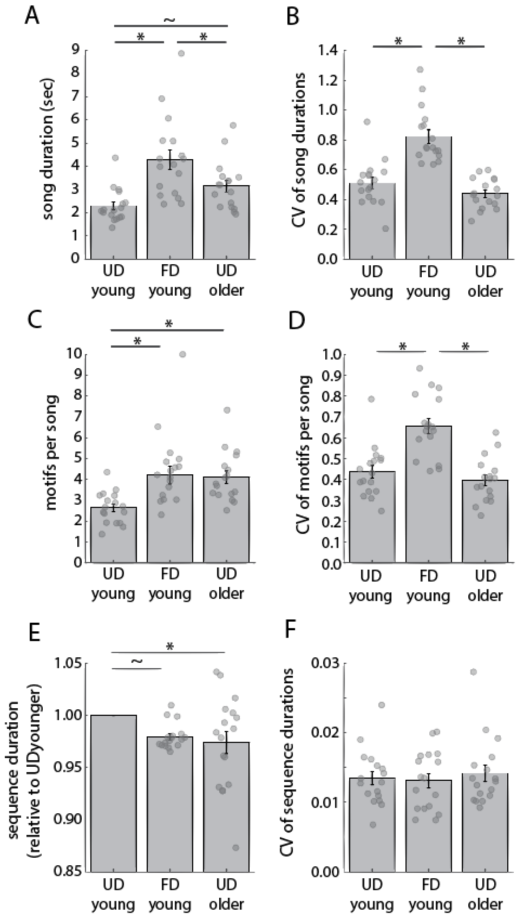


Figure 4: Age- and context-dependent changes to song bout and sequence durations. A. Song bout durations were significantly longer during $\mathrm{FD}_{\mathrm{young}}$ song than during both $\mathrm{UD}_{\mathrm{young}}$ and $\mathrm{UD}_{\mathrm{older}}$ songs, and there was a trend for $\mathrm{UD}_{\mathrm{older}}$ song bouts to be longer than $\mathrm{UD}_{\mathrm{young}}$ song bouts. B. The CV of song bout durations was significantly higher during $\mathrm{FD}_{\mathrm{young}}$ song than during both $\mathrm{UD}_{\mathrm{young}}$ and $\mathrm{UD}_{\mathrm{older}}$ songs. C. The number of motifs per song bout was significantly higher for $\mathrm{FD}_{\mathrm{young}}$ and $\mathrm{UD}_{\mathrm{older}}$ songs than for $\mathrm{UD}_{\mathrm{young}}$ songs. D. The CV of the number of motifs per song bout was significantly higher during $\mathrm{FD}_{\mathrm{young}}$ song than during both $\mathrm{UD}_{\mathrm{young}}$ and $\mathrm{UD}_{\mathrm{older}}$ songs (see also Stepanek and Doupe, 2010). E. Sequence durations were significantly shorter during $\mathrm{UD}_{\mathrm{older}}$ song than during $\mathrm{UD}_{\mathrm{young}}$ song, and there was a trend for sequence durations to be shorter for $\mathrm{FD}_{\mathrm{young}}$ song than $\mathrm{UD}_{\mathrm{young}}$ song. While raw data were analyzed, plotted are values normalized to sequence durations during $\mathrm{UD}_{\mathrm{young}}$ song because of the wide range of sequence durations across birds. F. There was no significant difference in the CV of sequence durations across recording conditions. "*" denotes contrasts with $p<0.05$, "∼" denotes contrasts with $p<0.10$.

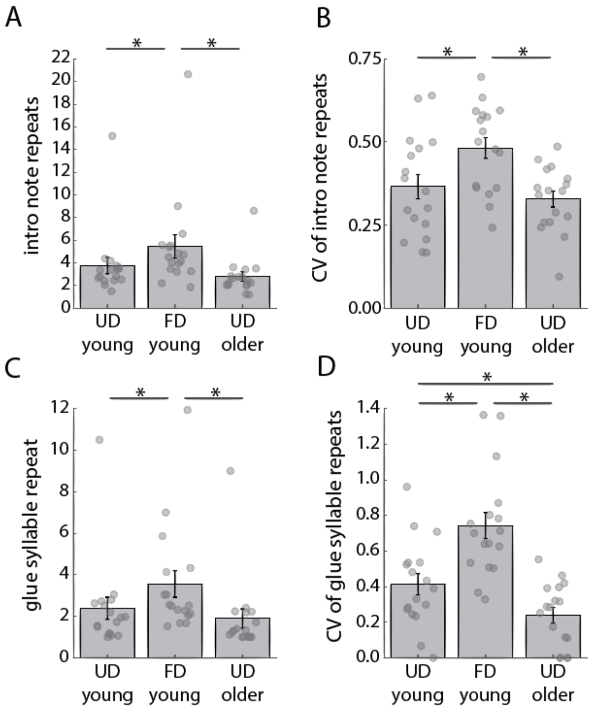


Figure 5: Age- and context-dependent changes to syllable repetition outside of the motif. A. Birds repeated introductory notes ("intro notes") more times before the onset of FD_{young} song than before the onsets of both UD_{young} and UD_{older} songs. B. The variability (CV) in the number of times introductory notes were repeated was significantly higher during FD_{young} song than during both UD_{young} and UD_{older} songs. C. Glue syllables were repeated more times between motifs during FD_{young} song than during both UD_{young} and UD_{older} songs. D. The variability (CV) in the number of times glue syllables were repeated was significantly higher during FD_{young} song than during both UD_{young} and UD_{older} songs, and significantly higher during UD_{young} song than UD_{older} song. "*" denotes contrasts with $p<0.05$.

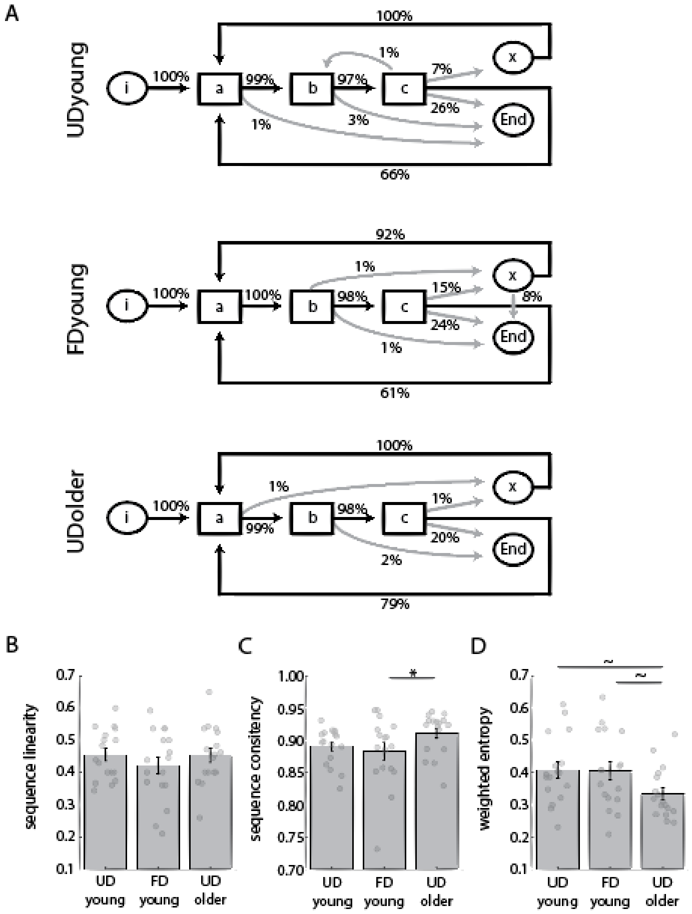


Figure 6: Age- and context-dependent changes to syllable sequencing variability. A. Example transition diagrams for one bird across all three conditions. Letters indicate syllables, and arrows represent transitions between syllables, with percentages indicating the local transition probability from one syllable to another syllable. “End” represents the termination of song. Letters in rectangles indicate “motif syllables” whereas “i” represents introductory notes and “x” represents glue syllables (see Figure 1). Black lines indicate ‘typical’ (i.e., the most common) transitions used for the calculation of sequence consistency (see Experimental Procedure). Repetition of introductory notes and glue syllables are not indicated (see Figure 5 for analyses of such repetition). B. There was no significant difference in sequence linearity across recording conditions. C. Sequence consistency was significantly higher during $\mathrm{UD}_{\mathrm{older}}$ songs than during $\mathrm{FD}_{\mathrm{young}}$ song. D. The weighted sequence entropy tended to be lower during $\mathrm{UD}_{\mathrm{older}}$ songs than both $\mathrm{UD}_{\mathrm{young}}$ and $\mathrm{FD}_{\mathrm{young}}$ song. “*” denotes contrasts with $p<0.05$, “~” denotes contrasts with $p<0.10$.

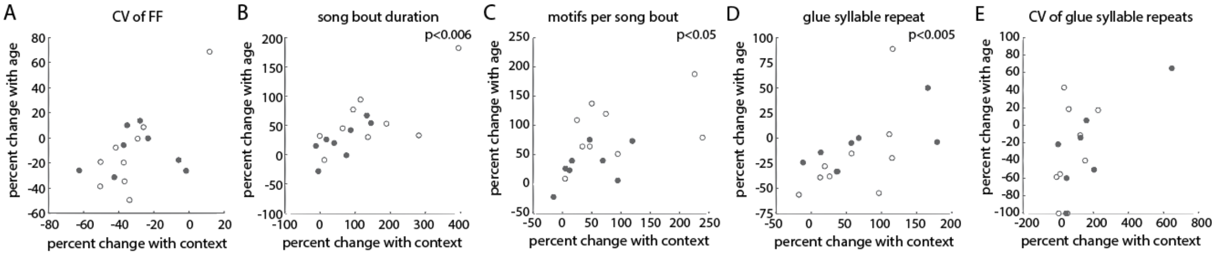


Figure 7: Correlations between context- and age-dependent changes to song. A-E. Significant correlations for the percent change of five song features (Spearman's ρ; $p<0.05$): CV of FF (average within each bird; A), song bout duration (B), motifs per song bout (C), glue syllable repetition (D) and the CV of glue syllable repetition (E). The correlations for A and E were not significant after the removal of the outlying value. Open circles = socially tutored birds; filled circles = non-socially tutored birds. P-values are listed only for relationships that remained significant after the removal of an outlying value (if there was one), and reported p-values are those obtained after the removal of the outlying value.

Appendix: Supplementary information

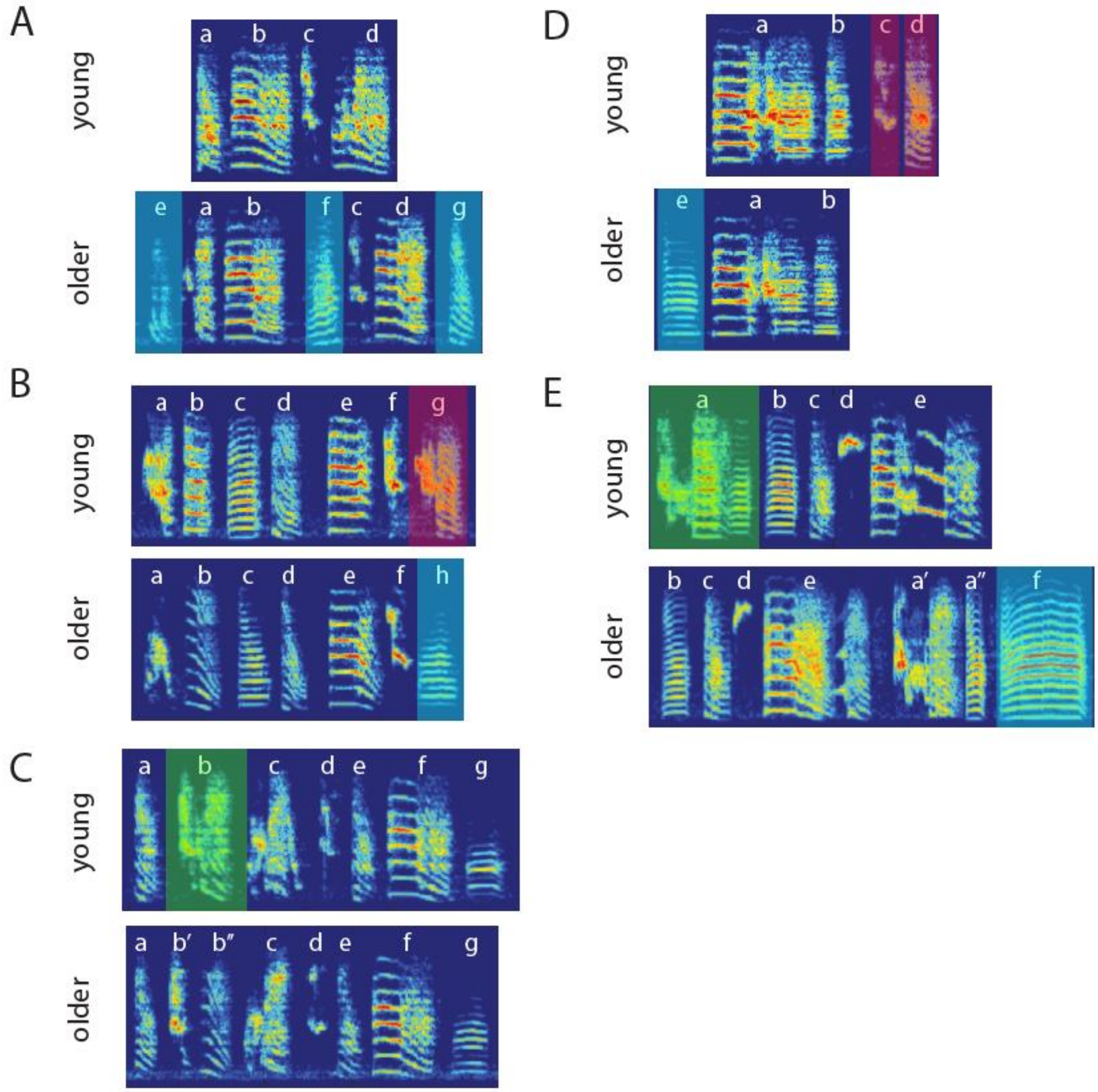
Supplementary Table 1
Full models for the mixed-effects models for all song features. These are full-factorial, mixed-effects models with method of tutoring (social vs. non-social) and recording condition (UD_{young}, FD_{young}, or UD_{older}) as the independent variables, and BIRD ID as a random effect (analogous to a repeated-measures design). Because we measured multiple syllables per bird for our analyses of fundamental frequency (mean and CV of FF), we included SYLLABLE ID nested within BIRD ID in these models. For simplicity, we only reported the effect of condition in the Results.

			Mean		**CV**	
Feature	**Factor**	**dfs**	**F**	**P**	**F**	**P**
Introductory Note	Tutoring	1,15	0.59	0.4535	0.28	0.6039
	Condition	2,30	13.48	**<.0001**	11.81	**0.0002**
	Tutoring*Condition	2,30	0.08	0.9257	0.57	0.5735
Song Duration	Tutoring	1,15	2.89	0.11	1.56	0.2312
	Condition	2,30	14.69	**<.0001**	43.87	**<.0001**
	Tutoring*Condition	2,30	2.16	0.1332	0.59	0.5597
Motifs per song	Tutoring	1,15	1.26	0.28	0.03	0.8556
	Condition	2,30	10.96	**0.0003**	21.30	**<.0001**
	Tutoring*Condition	2,30	1.97	0.1567	0.51	0.6035
Glue Syllables	Tutoring	1,15	0.41	0.5296	1.96	0.1823
	Condition	2,30	18.16	**<.0001**	29.23	**<.0001**
	Tutoring*Condition	2,30	0.08	0.9252	0.58	0.5634
Fundamental Frequency	Tutoring	1,24	0.48	0.4939	0.61	0.4423
	Condition	2,48	3.24	**0.0477**	28.89	**<.0001**
	Tutoring*Condition	2,48	0.15	0.8597	1.47	0.2409
Sequence Duration	Tutoring	1,15	11.48	**0.0041**	0.95	0.3456
	Condition	2,30	3.99	**0.0291**	0.37	0.6917
	Tutoring*Condition	2,30	0.01	0.9921	1.20	0.3138
Sequence Linearity	Tutoring	1,15	0.85	0.3718		
	Condition	2,30	0.74	0.484		
	Tutoring*Condition	2,30	0.33	0.7235		
Sequence Consistency	Tutoring	1,15	0.40	0.5356		
	Condition	2,30	3.69	**0.0368**		
	Tutoring*Condition	2,30	0.55	0.5803		
Sequence Entropy	Tutoring	1,15	0.54	0.474		
	Condition	2,30	3.89	**0.0314**		
	Tutoring*Condition	2,30	0.05	0.9469		

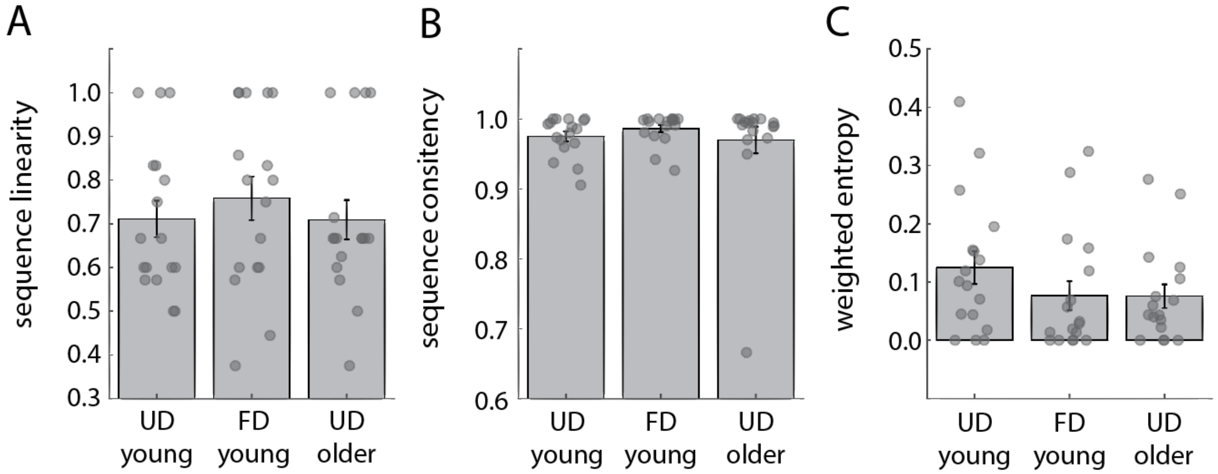
Supplementary Table 2

Full models for the analyses of covariance (ANCOVAs) for all song features. For this preliminary analysis, we assessed whether the relationship between context- and age-dependent changes differed between birds that were socially vs. non-socially tutored using a full-factorial ANCOVAs with method of tutoring as the independent variable, the change (percent or difference, see Statistical Analysis) across social contexts as the covariate, and the change with age as the dependent variable. For all but one song feature, there was no significant interaction between the method of tutoring and the relationship between context- and age-dependent changes. However, the relationship between context- and age-dependent changes in the CV of FF was significantly different between socially and non-socially tutored birds (p=0.0086). This interaction was driven by one outlying value, and after removing this value from the analysis, there interaction was no longer significant (p=0.3092). These analyses support the exclusion of method of tutoring in the analysis of the relationship between context- and age-dependent changes in the main text.

			Mean		**CV**	
Feature	**Factor**	**dfs**	**F**	**P**	**F**	**P**
Introductory Note	Tutoring	1,16	2.00	0.1808	0.57	0.4637
	AvgMeanContext	1,16	1.22	0.2898	0.04	0.8425
	Tutoring*AvgMeanContext	1,16	0.03	0.8549	0.21	0.6552
Song Duration	Tutoring	1,16	0.12	0.7358	0.14	0.7177
	AvgMeanContext	1,16	8.89	**0.0106**	3.50	0.0841
	Tutoring*AvgMeanContext	1,16	0.16	0.6958	0.63	0.4408
Motifs per song	Tutoring	1,16	3.92	0.0693	0.02	0.8937
	AvgMeanContext	1,16	3.00	0.1068	0.93	0.3519
	Tutoring*AvgMeanContext	1,16	0.03	0.867	1.58	0.2306
Glue Syllables	Tutoring	1,16	0.17	0.6833	0.46	0.5118
	AvgMeanContext	1,16	8.20	**0.0133**	4.71	0.0507
	Tutoring*AvgMeanContext	1,16	0.68	0.4232	0.19	0.672
Fundamental Frequency	Tutoring	1,16	0.02	0.8986	0.08	0.7828
	AvgMeanContext	1,16	0.05	0.8212	10.66	**0.0061**
	Tutoring*AvgMeanContext	1,16	3.52	0.0835	9.57	**0.0086**
Sequence Duration	Tutoring	1,16	0.00	0.9517	0.54	0.4766
	AvgMeanContext	1,16	1.59	0.2294	1.86	0.1957
	Tutoring*AvgMeanContext	1,16	0.05	0.825	1.93	0.188
Sequence Linearity	Tutoring	1,16	0.04	0.8379		
	AvgMean_Context	1,16	1.75	0.2082		
	Tutoring*AvgMeanContext	1,16	0.82	0.381		
Sequence Consistency	Tutoring	1,16	1.42	0.2548		
	AvgMean_Context	1,16	0.21	0.6581		
	Tutoring*AvgMeanContext	1,16	0.01	0.9178		
Sequence Entropy	Tutoring	1,16	0.02	0.887		
	AvgMeanContext	1,16	2.25	0.1577		
	Tutoring*AvgMeanContext	1,16	4.05	0.0653		



Supplementary Figure 1
Examples of the motifs of non-socially tutored birds that changed over time. Each pair of spectrograms (A-E) depicts a bird's motif when he was a young adult (top) and an older adult (bottom). Arbitrary letters above each spectrogram indicate the label of each syllable. Syllables highlighted in blue indicate "novel" syllables (i.e., syllables that were produced in the motif of the bird when it was an older adult but not when it was a young adult); syllables highlighted in red indicate "dropped" syllables (i.e., syllables that were produced in the motif of the bird when it was a young adult but not when it was an older adult); and, syllables highlighted in green indicate syllables that were split over time. Letters with an apostrophe indicate syllables that seem to be split from a syllable produced when the bird was younger.



Supplementary Figure 2
Sequence linearity (A), sequence consistency (B) and weighted entropy (C) scores for an analysis of syllable transitions within the motif (see Figure 6 for analysis of all transitions within the song bout). Here we analyzed transitions from syllables within the motif to other syllables within the motif or to the end of song (i.e., motif truncations). We did not consider transitions from the last syllable in the motif. We observed no significant main effect of condition (p>0.05 for all) nor any pairwise comparison (p>0.10 for all) for all three features.

Linking statement from chapter 3 to chapter 4

In the three previous chapters we have documented the predictability of age-dependent changes to song by analyzing context-dependent changes to song in two species. From here, we sought to better understand the predictability of song learning across development. For this, we focused on whether zebra finches exhibit internal biases that guide sequence learning using an experimentally controlled tutoring design. Here, we individually tutored zebra finches with random sequences, and asked whether internal filters biased birds to produce similar types of song sequences (despite the fact that the stimulus provided no information on which sequence to produce).

Chapter 4

Learning biases underlie “universals” in avian vocal sequencing

Logan S. James and Jon T. Sakata

Published: *Current Biology* 2017

Summary

Biological predispositions in vocal learning have been proposed to underlie commonalities in vocal sequences, including for speech and birdsong, but cultural propagation could also account for such commonalities (Bickel 2013, Chomsky 1980, Greenberg 1963, Yang et al. 2017). Songbirds such as the zebra finch learn the sequencing of their acoustic elements ("syllables") during development (Brainard and Doupe 2013, Doupe and Kuhl 1999, Lipkind et al. 2013, Prather et al. 2017). Zebra finches are not constrained to learn a specific sequence of syllables, but significant consistencies in the positioning and sequencing of syllables have been observed between individuals within populations and between populations (Lachlan et al. 2016, Lipkind et al. 2013, Zann 1996). To reveal biological predispositions in vocal sequence learning, we individually tutored juvenile zebra finches with randomized and unbiased sequences of syllables and analyzed the extent to which birds produced common sequences. In support of biological predispositions, birds tutored with randomized sequences produced songs with striking similarities. Birds preferentially started and ended their song sequence with particular syllables, consistently positioned shorter and higher frequency syllables in the middle of their song, and sequenced their syllables such that pitch alternated across adjacent syllables. These patterns are reminiscent of those observed in normally tutored birds, suggesting that birds "creolize" aberrant sequence inputs to produce normal sequence outputs. Similar patterns were also observed for syllables that were not used for tutoring (i.e., unlearned syllables), suggesting that motor biases could contribute to sequence learning biases. Furthermore, zebra finches spontaneously produced acoustic patterns that are commonly observed in speech and music, suggesting that sensorimotor processes that are shared across a wide range of vertebrates could underlie these patterns in humans.

Results and discussion

Songbirds such as the zebra finch offer a powerful opportunity to conduct controlled experiments of vocal sequence learning and uncover biological predispositions in vocal learning. This is because songbirds, like humans, learn the sequencing of acoustic elements ("syllables") during development (Brainard and Doupe 2013, Chen et al. 2013, Doupe and

Kuhl 1999, Lipkind et al. 2013, Prather et al. 2017). Furthermore, while species like the zebra finch can learn to produce a wide range of vocal patterns, there exists some evidence that particular acoustic patterns and sequences of syllables are more common than expected (Lachlan et al. 2016, Zann 1996). However, experimental tutoring is required to determine the contribution of learning predispositions (independent from cultural forces) to the production of common vocal patterns.

To reveal biases in vocal sequence learning, we tutored naive juvenile zebra finches (i.e., reared without song exposure during their critical period for song learning) with randomized sequences of syllables that lacked sequencing consistency and then assessed the degree to which tutored individuals developed songs with similar syllable sequences (see STAR Methods). We tutored juveniles with synthesized songs consisting of five species-typical syllables (labeled as "a"–"e") arranged in every possible five-syllable sequence (n = 120 possible sequences; Figures 1A–1C). Importantly, each bird heard the 120 possible sequences in equal proportions (Figure 1C), and the order of sequence presentations was randomized across individuals and days of tutoring. Consequently, the tutoring stimuli provided no bias or predictability in syllable sequencing (e.g., syllable position within a sequence or directionality of syllable transitions). If biological predispositions bias the acquisition of vocal sequences, we should observe significant consistencies in syllable positioning or sequencing across tutored birds, despite being tutored by stimuli that provided no such consistencies.

In support of this hypothesis, experimentally tutored birds produced songs with significant similarities in the positioning and sequencing of syllables. Despite being tutored with randomized sequences, birds demonstrated significant learning of the acoustic structure of syllables (see STAR Methods), and like normally tutored birds, most experimentally tutored birds (45 out of 51 birds) produced one dominant sequence of syllables ("motif"; Figures 2A, S1, S2, and S3). We first investigated similarities in how tutored birds positioned syllables within their motif (i.e., beginning, middle, or end). The positioning of three of the five tutored syllables significantly deviated from chance. For example, of the birds that produced the "d" syllable in their motif, 4% produced the "d" at the beginning, 46% produced the “d” in the middle, and 50% produced the “d” at the end. This skewed distribution of the “d” syllable across positions was significantly different than the distribution expected by chance ($p = 0.0328$; as determined using Monte Carlo simulations; see STAR Methods) and was driven by the facts that the “d” was produced more than expected at the end of the motif and less than expected at the beginning ($p < 0.02$ for each; observed likelihoods in each position were, respectively, 3.43σ and -2.30σ away from the chance value; Figure 2B). The distributions of the syllables “e” and “b” also significantly differed from chance (Figure 2B): the syllable “e” was significantly more likely to be produced in the middle of the motif than expected by chance, whereas the syllable “b” was significantly more likely to be produced at the beginning of the motif. Consequently, despite being tutored with stimuli in which each syllable was equally likely to occupy each position, birds preferentially produced syllables in particular positions within the motif.

Importantly, these patterns of syllable positioning resemble those observed among populations of normally reared zebra finches (Lachlan et al. 2016, Zann 1996). Across wild populations of zebra finches, birds frequently terminate their motif with distance calls (our syllable “d”) and produce a high-pitched syllable (our syllable “e”) in the middle (Zann 1996). As such, our data strongly suggest that the biological predispositions in vocal sequence learning could account for natural commonalities in vocal positioning.

Positional variation in the duration, pitch, and amplitude of sounds is prevalent across speech and music (Fougeron and Keating 1997, Oller 1973, Patel 2010, Vaissière 1983). For example, utterances in the final position of phrases tend to be longer and lower in pitch than utterances in medial positions, and music phrases tend to have arch-like pitch contours such that notes in the beginning and final positions tend to be lower in pitch. We assessed the degree to which our experimentally tutored birds organized the syllables in their songs with such positional variation. Consistent with these patterns, syllables at the end of the motif were significantly longer than syllables in the middle or beginning of the motif ($p < 0.02$ for each contrast; Figure 2C). Syllables at the beginning of the motif were also significantly longer than syllables in the middle of the motif ($p = 0.007$). Pitch (mean frequency) also varied across motif position, with beginning and end syllables being lower in pitch than middle syllables ($p < 0.04$ for each; Figure 2D). Amplitude did not significantly vary across motif positions ($p = 0.2478$) but generally demonstrated an ascending pattern from beginning to end (Figure 2E).

To assess whether similar positional variation is observed in the songs of zebra finches that were tutored with species-typical songs, we applied the same analyses to zebra finch songs from an online repository ($n = 61$; see STAR Methods). We found patterns of positional variation that resembled those found in our experimentally tutored birds. For example, syllable duration significantly varied across position, with end syllables being significantly longer than syllables in the beginning or middle of the motif ($p < 0.0001$ for each contrast; Figure 2F). Pitch also significantly varied across motif positions (Figure 2G), and, similar to the songs of experimentally tutored birds, beginning syllables were lower in pitch than middle syllables ($p = 0.0006$). While end syllables were not significantly different in pitch than middle syllables, end syllables were higher in pitch than beginning syllables ($p = 0.0043$). Mean amplitude significantly varied across motif position within the songs of finches tutored with species-typical songs, with end syllables being the loudest, followed by middle syllables and then beginning syllables ($p < 0.0003$ for each contrast; Figure 2H). A similar, albeit statistically non-significant, pattern was observed in the songs of experimentally tutored birds. Because patterns of positional variation in acoustic features were similar between birds tutored with randomized sequences and birds tutored with normal zebra finch song, this suggests that experimentally tutored birds spontaneously organized their syllables into patterns typically found in zebra finch song.

In addition to the positioning of syllables, we also analyzed the degree to which the sequencing of syllables within the songs of experimentally tutored birds followed predictable patterns. We first analyzed predictability (i.e., asymmetries) in the directionality of syllable transitions within the 10 pairs of tutored syllables (e.g., transitioning from “a” to “e” versus from “e” to “a”; Figures 3A–3C). Because birds were tutored with sequences that were symmetric in the directionality of syllable transitions, the null model (i.e., no bias in sequence learning) predicts equal probabilities of transitions in each direction. However, the observed pairwise transitions were significantly asymmetric (p = 0.0312; Figure 3D; see STAR Methods), indicating that birds produced similar sequences. One striking result that exemplifies this predictability was that four birds produced the exact same four-syllable sequence, “bead” (Figure 2A). The likelihood that the same four-syllable sequence is produced by four or more birds, by chance, is <0.2% (Monte Carlo simulation), underscoring that syllable sequencing is non-random and that biological predispositions contribute to vocal sequence learning.

Speech sounds are commonly sequenced in alternating patterns of stressed (generally higher in pitch, louder, and longer) and unstressed syllables (Goedemans and van der Hulst 2013, Langus et al. 2016, Liberman and Prince 1977). However, the extents to which songbirds produce sequences with alternating acoustic patterns and to which biological predispositions in vocal learning contribute to this patterning remain largely unknown. To this end, we first analyzed a large corpus of songs of zebra finches that were tutored with species-typical songs (see STAR Methods). While we did not observe significant patterning of duration (p = 0.2096) or amplitude (p = 0.0838), birds tutored with normal song alternated the mean frequency of syllables in their motif more than expected by chance (Monte Carlo simulation; p = 0.0223; Figures 4A and 4B). To reveal the contribution of learning predispositions to this patterning, we assessed the degree to which birds tutored with randomized sequences similarly produced songs with alternating patterns. We found that experimentally tutored birds significantly alternated the pitch (p = 0.0018), but not duration or (p = 0.9311) amplitude (p = 0.2570), of syllables in their motif (Monte Carlo simulation; Figure 4C). These data strongly suggest that learning biases could contribute to the typical alternation of pitch in zebra finch song.

Motor and/or sensory biases could underlie such biological predispositions in learning. For example, the organization of the motor system has been hypothesized to influence and constrain the acquisition and production of acoustic patterns within birdsong, speech, and music (MacNeilage et al. 2000, Patel 2010, Podos 1996, Tierney et al. 2011, Zollinger and Suthers 2004). To assess the degree to which motor biases contribute to biological predispositions in vocal sequence learning, we analyzed positional variation in the acoustic features of syllables that experimentally tutored birds produced and that were not heard in the tutor stimulus set ("novel syllables"; see STAR Methods; Figures S2 and S3). As observed in the previous analyses of experimentally or normally tutored birds (Figure 2), novel syllables that were produced at the end of the motif were longer than those produced in the middle or beginning of the motif, and novel syllables at the beginning or end of the motif were lower in pitch than those in the middle (Figure S4). Because positional variation in utterances was observed even for vocalizations that were not directly heard during tutoring, these data suggest that motor biases contribute to the acoustic patterns commonly observed in our birds. It is possible, for example, that such acoustic patterns are easier for the motor system to produce (MacNeilage et al. 2000).

Despite these findings, it remains possible that biases in auditory processing and memory could also contribute to the common vocal patterns. Just as previous studies demonstrate "innate" tuning for species-typical acoustic features (Araki et al. 2016, Doupe and Kuhl 1999, Gómez et al. 2014, Marler and Peters 1997), our data suggest that the auditory system of juvenile zebra finches could preferentially process and learn particular types of acoustic sequences. Recent studies on prosodic processing in zebra finches support this idea. When asked to categorize patterns of human speech sounds or zebra finch syllables, zebra finches preferentially use pitch cues (relative to duration or amplitude) to infer patterns (Spierings et al. 2017, Bion et al. 2011). Interestingly, birds tutored with randomized sequences spontaneously arranged their syllables such that pitch but not amplitude or duration alternated, highlighting a parallel in the prominence of pitch to production and perception in zebra finches. Similar parallels in production and processing are observed in humans, as speakers of languages in which pitch is used as a primary prosodic cue tend to use pitch cues for speech segmentation (Bion et al. 2011, Jusczyk et al. 1999). In addition, these data suggest the intriguing possibility that females could prefer males that produce songs with alternation of and positional variation in pitch.

Regardless of the mechanism, we provide compelling experimental evidence that the nervous system is endowed with predispositions that bias vocal sequence learning in a songbird species that lacks specific rules for vocal sequence production (Gardner et al. 2005, Lachlan et al. 2016, Plamondon et al. 2008, Zann 1996). Tutoring birds with random, unbiased sequence inputs led to the production of non-random, biased sequence outputs that were common across individuals, similar to patterns found in birds tutored with species-typical songs, and similar to those commonly observed in speech and music (Bickel 2013, Chomsky 1980, Goedemans and van der Hulst 2013, Patel 2010, Savage et al. 2015, Tierney et al. 2011). Consequently, these data indicate that biological predispositions influence not only the learning of acoustic structure (Fehér et al. 2009, Marler and Peters 1977, ter Haar et al. 2014) but also the acquisition of syllable sequencing. That the patterning of acoustic features is similar between the songs of experimentally and normally tutored birds indicates that juvenile zebra finches “normalize” atypical inputs of syllable sequences into species-typical sequences. This process resembles the “creolization” of human languages and acoustic features of birdsong (Fehér et al. 2009, Gardner et al. 2005, Senghas and Coppola 2001, Thomason and Kaufman 1988). Children who are exposed to novel, simplified forms of languages that lack properties commonly observed across established languages reshape and “creolize” this language into one that is more likely to contain common linguistic patterns (Senghas and Coppola 2001, Thomason and Kaufman 1988), and juvenile zebra finches that are tutored by conspecifics that produce acoustically atypical songs produce syllables with more normal acoustic structure as adults (Fehér et al. 2009). It will be important for future studies to assess the extent to which these learning biases can be overcome and to explicitly evaluate the motor and sensory underpinnings of these biases. Given the parallels to patterns in speech and music, our findings suggest that sensorimotor processes that are shared across a wide range of vertebrates underlie “universals” in phonological, prosodic, and musical patterns in humans (Liberman and Mattingly 1985, Mol et al. 2017, Tierney et al. 2011, Zollinger and Suthers 2004).

STAR methods

Experimental Model and Subject Details

Adult zebra finches for breeding were obtained from commercial breeders (Exotic Wings and Things, Ontario, Canada). Naive juvenile zebra finches were bred and raised in our lab in custom-built sound-attenuating chambers (TRA Acoustics, Ontario, Canada). Fathers were removed from the nest ≤5 days post-hatch (dph), and, thereafter, experimental juveniles were raised solely by their mother or a foster mother (females do not produce song) (Fehér et al. 2009, Zann 1996). The critical period for song learning for zebra finches opens ~20 dph; thus, song exposure before 5 dph does not lead to significant song learning (Dunn and Zann 1996, Fehér et al. 2009, Roper and Zann 2006). After reaching nutritional independence (~35–50 dph), juvenile males were individually housed in sound-attenuating chambers for tutoring. Juveniles were individually housed because social interactions among juveniles can influence song development (Derégnaucourt and Gahr 2013, Tchernichovski and Nottebohm 1998, Volman and Khanna 1995). Animals used in this experiment were not part of any other experiment prior to tutoring. All birds were housed on a 14:10 light:dark cycle and provided with food and water ad libitum. All procedures were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

Method Details

Song Tutoring

Birds were tutored using operantly triggered playbacks of synthesized song stimuli for >30 days (Clayton 1989, Tchernichovski et al. 1999). Song stimuli were synthesized to resemble natural zebra finch song, albeit with no sequence consistency across renditions. To create song stimuli for tutoring, five typical zebra finch song elements (termed “syllables,” contiguous epochs of sound separated by short silent gaps (>5 ms)) were selected from the songs of different males from multiple lab populations. These five syllables corresponded to syllable types most commonly observed in natural populations of zebra finches, four of which were termed the “primary elements” of zebra finch phrases (Zann 1996). Using the nomenclature of Zann (1996), these syllables can be described as a short sweep (“a”), a noise-noise syllable (“b”), a short stack (“c”), a distance call (“d”), and a high note (“e”; Figure 1). To create tutoring stimuli, syllables were bandpass-filtered (0.3–10 kHz), normalized by their maximum amplitude, and ramped to avoid sharp onsets and offsets in amplitude

We tutored juveniles with stimuli that contained randomized and unbiased syllable sequences to assess biological predispositions in sequence learning (n = 51). To this end, we constructed all possible five-syllable sequence ("motif") variants that contained each of the syllables only once (n = 120 motif variants; Figure 1). We created motif variants in this manner because individual syllables are rarely produced more than once in a zebra finch's motif (Clayton 1989, Fehér et al 2009, Sturdy et al. 1999). Because zebra finches typically produce bouts of song that consists of multiple renditions of a motif, we synthesized song stimuli such that each song bout contained four motifs. However, in contrast to normal zebra finch song bouts in which the same motif is repeated, synthesized song bouts consisted of four distinct motif variants. The sequence of motif variants per song bout was randomized across days for each bird and across birds. Moreover, to ensure that each motif variant was heard equally and randomly, we presented the stimuli in blocks of 120 motifs so that every motif variant was heard once before any variant was repeated.

The timing of syllables and motif variants in song stimuli followed a species-typical pattern: syllables within the motif were each separated by a silent gap of 30 ms, whereas the gap between syllables across adjacent motifs (i.e., from offset of last syllable in the motif to onset of first syllable of subsequent motif) was 100 ms.

Birds were operantly tutored using perch hops (Adret 1993, Ten Cate 1991), and each perch hop triggered the playback of one song bout. To improve learning outcomes, song playbacks were spaced out such that juveniles were allowed to hear only 10 operantly triggered song playbacks (i.e., 40 motif variants) within each of three time periods in the day (morning, noon, and afternoon; Tchernichovski et al. 1999), for a maximum of 120 motifs (one of each variant) per day. Birds were housed individually throughout tutoring and until 4 months of age.

Sound Analysis Pro 2011 (SAP; http://soundanalysispro.com) was used for song tutoring. Stimuli were played out of an Avantone Pro Mixcube speaker (Avantone, NY) connected to a Crown XLS 1000 amplifier (Crown Audio, IN). Song playback was operantly triggered by perch hops using custom-built perches connected to a National Instruments PCI-6503 I/O card (National Instruments, TX).

Song Analysis

We recorded the spontaneously produced songs of tutored birds using an omnidirectional microphone (Countryman Associates, Menlo Park, CA) and an amplitude-based recording system (SAP; 44.1 kHz; bandpass-filtered from 0.3-10 kHz). Songs were visualized and analyzed offline using custom-written software in MATLAB (Mathworks, Natick, MA). We analyzed the adult songs (n > 30 songs/bird) of tutored birds when they were 4 months old, an age at which song is typically stable and stereotyped (Brainard and Doupe 2002, Mooney 2009). We identified and labeled syllables and motifs following automated, amplitude-based segmentation of audio files. As indicated above, zebra finch songs consist of syllables that are arranged into stereotyped sequences called "motifs," and each adult zebra finch repeats a single motif within his song. An individual's motif is readily identifiable because it is repeated multiple times within a song bout (Brainard and Doupe 2002, Mooney 2009, Scharff and Nottebohm 1991, Zann 1996) and because the duration of gaps between motifs is usually longer and more variable than gap durations within the motif (Glaze and Troyer 2006, Vu et al. 1994). To identify motifs, authors independently examined multiple renditions of each bird's song. Both researchers identified and agreed upon a stereotyped motif for 45 of the 51 tutored birds, and the songs of these birds resembled the songs of normally reared zebra finches. The remaining 6 tutored birds either produced songs with sufficient sequence variability to make motif identification difficult or produced songs with multiple motif variants (i.e., songs in which authors could not agree upon a motif; see Figure S1 for examples). Such variability is not unexpected since operant tutoring can lead to aberrant learning outcomes in a subset of birds [e.g., Lipkind et al. (2013), Derégnaucourt et al. (2013)]. The songs of these six birds with more variable sequencing were not analyzed because of ambiguity in demarcating the beginning and end of the motif.

The 45 birds used in the analysis produced motifs with, on average, 4.4 ± 0.2 syllables (mean ± SEM; range: 2–8), with average gap durations (i.e., intervals between syllables within the motif) of 27.1 ± 1.2 ms. Thirty-seven of the 45 birds (82%) produced a direct transition from the end of one motif to the beginning of the subsequent motif, and the average inter-motif interval (duration from the offset of the last syllable in the motif to the onset of the first syllable in the subsequent motif) for these birds was 77.5 ± 7.5 ms.

After identifying motifs, we labeled each syllable within the motif using quantitative and qualitative approaches. We first identified the median rendition of each distinct syllable of a bird's motif. To identify the median syllable rendition, we used MATLAB scripts to quantify acoustic features of every rendition of each syllable in the bird's motif across >30 randomly selected songs (duration, mean frequency, spectral density, spectral temporal entropy, and loudness entropy; Sakata and Brainard 2006, Stepanek and Doupe 2010), converted these feature values into z-scores for each syllable in each bird's motif, and then identified the individual syllable rendition that was closest to the median centroid of the distributions. Once the median rendition for each syllable was identified, five individuals with expertise studying birdsong were presented with individual spectrograms of each median syllable rendition and asked to determine whether the syllable matched those in the tutoring stimulus set (syllables "a" through "e") or should be considered distinct from the five tutor syllables (i.e., "novel"). Experts were blind to the identity of the bird and the position of the syllable within the motif (i.e., raters categorized individual syllables that were spliced out from the motif). The label chosen by at least three of the five raters was selected as the final label for the syllable. There was a single syllable in which there was no consensus across raters, and this syllable was categorized as a "novel" syllable. On average, each rater's labels were consistent with the final labels (i.e., labels used in analyses) for 90% of the syllables.

The degree of syllable learning in tutored birds with stereotyped motifs (n = 45) was comparable to that observed in previous studies of experimental song tutoring (Chen et al. 2016, Zann 1996); on average, 72% of all syllables produced by these birds were categorized as syllables that were learned from the tutor stimulus set. To further assess the extent of learning, we employed quantitative analyses of acoustic similarity using SAP (Tchernichovski et al. 2000). We measured the spectral similarity of each pupil's motif to a motif from the tutor stimulus. The SAP "similarity score" does not take into account similarities in the sequencing of syllables between compared motifs; therefore, we randomly picked one of the 120 motifs in the tutor stimulus to use for all comparisons. On average, the motifs of our tutored birds had SAP similarity scores of 55.9 ± 2.6% (mean ± SEM), and these scores were significantly and positively correlated with our estimate of the number of syllables learned from the tutor stimulus set ($r = 0.64$, $p < 0.0001$). Furthermore, these similarity scores were comparable to those reported for birds that were operantly tutored in other studies; for example, Derégnaucourt et al. (2013) found that pupils tutored by operant conditioning of natural song stimuli had an average SAP similarity score of 61.1 ± 3.4% (values extracted from their figure; Derégnaucourt et al. 2013).

In addition, the songs of our experimentally tutored birds were more similar to the tutor stimulus set than were the songs of zebra finches not tutored with these stimuli. For this analysis, we used SAP to measure the similarity of motifs of zebra finches from five different laboratories to our tutor stimulus set (http://people.bu.edu/timothyg/song_website/index.html; see "Comparisons with zebra finches tutored with normal songs" below for further details). These similarity scores provide an estimate of the extent to which birds could, by chance, resemble the tutor stimulus set. Overall, the motifs of birds not tutored by our tutor stimulus set had similarity scores of 32.4 ± 1.9%, which was significantly lower than the scores of our tutored birds ($t_{104} = 7.39$, $p < 0.0001$). Altogether, these analyses indicate that the extent of learning of our experimentally tutored birds was significantly greater than chance and comparable to the degree of learning observed in related studies [e.g., Derégnaucourt et al. (2013)].

To assess the robustness of the findings reported in the main text, we also conducted the same set of analyses on motifs that were independently identified by two songbird researchers who were unaware of the experiment. These researchers were asked to identify motifs by examining spectrograms of the adult songs of tutored birds. Overall, there was 86% similarity in syllable composition of motifs across the two researchers and

88% similarity to the dataset analyzed in the main text. Moreover, results for these analyses were highly consistent with those reported in the main text, with 89% of analysis yielding $p < 0.05$ or $p < 0.10$ for all significant effects reported in the main text.

Comparisons with Zebra Finches Tutored with Normal Songs

To corroborate patterns that were commonly observed in our experimentally tutored birds, we analyzed the extent of positional variation in and alternation of song features within the songs of zebra finches tutored with normal songs. We examined zebra finch songs from a repository that consisted of songs from five different laboratories (http://people.bu.edu/timothyg/song_website/index.html); songs from our laboratory were not included in the repository. While specific information about how these birds were tutored was not provided, these recordings were obtained from labs that generally raise and tutor birds with typical zebra finch song (i.e., stereotyped sequences of syllables). We independently examined the repository and identified a subset of birds in which we could readily identify a motif. Songs were included in the analysis if (1) the song file contained at least two full motifs (since only one file existed per bird) and (2) we agreed on the exact sequence of syllables that comprised the motif (n = 61 birds).

Because these songs were obtained from birds raised in different laboratories, we examined the extent to which positional variation in or alternation of song features varied across labs. There was no significant interaction between lab of origin and motif position or alternation in the dataset ($p > 0.2$ for each), suggesting that the nature and extent of positional variation and alternation were comparable across lab populations.

Quantification and Statistical Analysis

Monte Carlo Simulations

We conducted Monte Carlo simulations to determine the extent to which adult birds tutored with randomized sequences produced similar acoustic patterns within their motifs (Manly 1997). These simulations were used to compute the likelihood of observing a particular pattern by chance and were important because individual birds produced different repertoires of syllables. These simulations were also important to account for the number of times each syllable was produced within the motif (i.e., syllables were occasionally produced more than once in a motif). The simulation consisted of 10,000 iterations, and for each iteration we randomized the sequence of syllables within each bird's motif. We then computed the likelihood of observing a particular phenomenon across motifs (e.g., percent of motifs that ended with a particular syllable) during each iteration, and then generated a distribution of these likelihoods across the 10,000 iterations. These Monte Carlo simulations were the basis for generating null distributions against which we assessed the significance of our observations.

Skew in Syllable Position

Monte Carlo simulations were first used to quantify the degree to which the observed skew in syllable positioning within the motif (beginning, middle, and end) deviated from chance. The first syllable of the motif was designated as positioned at the beginning, the final syllable of the motif was designated as positioned at the end, and the remaining syllables were designated as positioned in the middle. We first analyzed the degree to which the distribution of each syllable across motif positions deviated from random using Likelihood ratio tests. Using the data from Monte Carlo simulations, we computed the expected distribution (i.e., null distribution) of syllables across motif positions if syllable sequencing was random. For example, to generate the null distribution for the "d" syllable, we computed the proportion of motifs (across birds) in which the "d" syllable was situated in the beginning, middle, or end of the motif for each iteration of the Monte Carlo simulation. We then calculated the average proportion across all 10,000 iterations for each of the three positions to generate the expected probability of observing a "d" syllable in each of these positions. Because the sequence of syllables within each bird's motif was randomized for each iteration of the simulation, this allowed us to compute the expected probability of a "d" occupying a particular position if sequence learning was random (i.e., if there were no biological predispositions for sequence learning). We then multiplied these average proportions by the number of birds that produced the "d" syllable to generate the expected number of birds that would have produced the "d" syllable at the beginning, middle, or end of their motif if sequencing was random. Finally, we compared these expected numbers to the observed number of birds that produced the “d” syllable at the beginning, middle, or end of the motif using a Likelihood ratio test. This procedure was repeated for each of the five tutor syllables.

To supplement the Likelihood ratio tests, we used Monte Carlo simulations to assess whether the observed likelihoods that individual syllables were produced at the beginning, middle, or end of the motif were each significantly different than chance. Using the syllable “d” as an example again, we observed that, of birds that produced the “d” syllable in their motif, 4% produced the “d” at the beginning of the motif, 46% of birds produced the “d” in the middle of the motif, and 50% of birds produced the “d” at the end of the motif. To assess whether each of these observed percentages were significantly different than chance, we created three null distributions from the Monte Carlo simulation, with distributions representing the likelihood of observing a “d” at the beginning, in the middle, or at the end by chance. We then compared the observed likelihoods of observing a “d” at the beginning, middle, and end against their respective null distribution. We deemed the observed percentage as significant if it lay outside the 95th percentile of values from the simulation (two-tailed). For example, 50% (13 of the 26) of the birds that produced a “d” in their motif positioned the “d” at the end of the motif, but in only 0.14% (14 of the 10,000) simulated iterations did we observe that ≥ 50% of motifs ended with “d”; in other words, by chance alone, it was highly unlikely that 50% of birds ended their motif with a “d.” Consequently, our results indicate a significant bias for zebra finches to produce the “d” syllable at the end of the motif. Conversely, only 3.8% (1 of the 26) of these birds produced the “d” at the beginning of their motif, and in only 1.2% of the simulated iterations did we observe that ≤3.8% of motifs began with “d”; therefore, our results indicate a significant bias for birds to not produce a “d” at the beginning of the motif.

Transitions between Syllables

We also analyzed the extent of patterning in syllable transitions within the motif. For example, we analyzed the degree to which birds exhibited consistent directionality (i.e., predictability) in sequencing for each of the 10 pairs of tutored syllables. To this end, we compared the observed asymmetry in syllable sequencing to the distribution of sequence asymmetries across the iterations of the Monte Carlo simulation. To compute sequence asymmetry, we first calculated, for each pair of syllables, the proportion of transitions in one direction (e.g., "e" to "a") out of the total transitions in either direction (e.g., sum of "e" to "a" and "a" to "e"); this value should be 0.5 if transitions were equally likely in both directions. We computed the asymmetry in pairwise transitions as the absolute deviation of the observed proportion from 0.5 ("asymmetry value"). For example, across the 12 birds that produced motifs with a direct transition between "e" and "a" (in either direction), 8 of those were from "e" to "a"; therefore, the asymmetry value for the "e"–"a" pair was 0.167 ($|(8/12)-0.5|$). To compute the overall degree of asymmetry in syllable sequencing, we averaged the asymmetry values across the 10 syllable pairs ("overall asymmetry value"). To compute the significance of the overall asymmetry value that we observed, we compared the observed overall asymmetry value to the distribution of overall asymmetry values derived from the Monte Carlo simulation. To generate this distribution, we calculated the overall asymmetry value on each of the 10,000 iterations of the simulation, and then computed the likelihood of observing an overall asymmetry value that was equal to or greater than the overall asymmetry value observed in our data (Figure 3D).

Alternation in Acoustic Structure

We ran Monte Carlo simulations to assess the significance of acoustic alternations within each bird's motif (Figures 3E and 3F). We measured the mean frequency, duration, and mean amplitude, three features that are commonly found to alternate in speech, for each rendition of each syllable in every bird's motif (Sakata and Brainard 2006, Stepanek and Doupe 2010). Then we computed the median value for each feature across all renditions of the syllable and used these values in the analysis. We plotted the change in mean frequency, for example, from the previous syllable$_{(t-1)}$ to the current syllable$_{(t)}$ against the change from the current syllable$_{(t)}$ to the next syllable$_{(t+1)}$ across all birds to measure the extent of pitch alternation between adjacent syllables in the motif. If there is acoustic alternation from syllable-to-syllable, we would expect a negative relationship between the change from $\text{syllable}_{(t-1)}$ to $\text{syllable}_{(t)}$ and from $\text{syllable}_{(t)}$ to $\text{syllable}_{(t+1)}$. For each acoustic feature, we then computed the r^2 value of the relationship between the two changes ("observed r^2") and compared this observed r^2 value to the distribution of r^2 values obtained for each of the 10,000 simulated iterations. For example, the observed r^2 value for the alternation in mean frequency was 0.476, and in only 18 of the 10,000 iterations did we observe an r^2 value equal to or greater than the observed r^2 ($p = 0.0018$). Therefore, the observed extent of pitch alternation in the songs of experimentally tutored birds was significantly greater than that expected by chance alone.

Analysis of Variation in Acoustic Features across Positions

We analyzed how acoustic features of syllables varied depending on the position in the motif (beginning, middle, or end) in our birds tutored with randomized sequences of syllables as well as birds tutored with species-typical song. We measured the duration, mean frequency, and mean amplitude of each rendition of each syllable in every bird's motif, and then analyzed how the medians of these features varied depending on syllable position. We used mixed effects models with position (Beginning, Middle or End) as the independent variable and individual bird as a random variable (to account for the fact that multiple syllables were analyzed per bird) to analyze systematic variation in acoustic structure across positions. Such mixed effects models (with birdID as a random variable) are repeated-measures analyses that provide the flexibility to accommodate the facts that most birds have multiple syllables defined as "middle syllables" and that different birds have different numbers of middle syllables. Tukey's HSD tests were used for post hoc contrasts. Significant differences between motif positions are identical if we compare values for beginning and end syllables to the average value for all middle syllables for each bird (i.e., each bird has a single data point per position).

For birds tutored with the same set of syllables (i.e., our experimentally tutored birds), we examined the degree to which positional variation in acoustic structure was due to variation in the types of syllables placed at different positions in the motif as well as variation in the acoustic structure of individual syllables according to position. Indeed, acoustic features of individual speech sounds systematically vary according to position (Fougeron and Keating 1997, Oller 1973). To differentiate between these possibilities, we conducted an additional analysis in which "syllable identity" ("a," "b," "c," "d" or "e") was included as an additional factor in the model. Acoustic features did not significantly vary across position when "syllable identity" was included in model, suggesting that positional variation in acoustic structure is primarily driven by variation in the placement of particular syllables at different parts of the motif.

Statistical Analyses

All analyses were done using MATLAB (Mathworks, Natick, MA) or JMP v. 11 (Cary, NC). We set $\alpha = 0.05$ for all tests.

Acknowledgements

We thank K. Onishi, M. Kao, R. Krahe, and S.C. Woolley for constructive input and feedback throughout the experiment and H. Goad and L. White for formative discussions about speech and language patterns. We also thank L. Matheson, H. Sun, J. Dai, Y. Chen, and G. Lafon for help with data analysis and manuscript preparation; T. Gardner for collecting and making available the large database of zebra finch songs; and three anonymous reviewers for their helpful feedback on the manuscript. Research was supported by funding from the Natural Sciences and Engineering Research Council (J.T.S.; #05016), the Centre for Research on Brain, Language, and Music, and an award from the Heller Family Fellowship (L.S.J.).

References

Adret P. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim Behav* 46: 149–159, 1993.

Araki M, Bandi MM, Yazaki-Sugiyama Y. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–1287, 2016.

Bickel B. Distributional biases in language families. In *Language typology and historical contingency* 415–444, 2013.

Bion RAH, Benavides-Varela S, Nespor M. Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Lang Speech* 54: 123–140, 2011.

Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, Doupe AJ. Translating birdsong: Songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36: 489–517, 2013.

Chen Y, Matheson LE, Sakata JT. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc Natl Acad Sci* 113: 6641–6646 2016.

Chomsky N. Rules and representations. *Behav Brain Sci* 3: 1, 1980.

Deregnaucourt S, Gahr M. Horizontal transmission of the father's song in the zebra finch (*Taeniopygia guttata*). *Biol Lett* 9: 20130247, 2013.

Derégnaucourt S, Poirier C, Van der Kant A, Van der Linden A, Gahr M. Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J Physiol Paris* 107: 210–218, 2013.

Clayton NS. The effects of cross-fostering on selective song learning in Estrildid Finches. *Behaviour* 109: 163–174, 1989.

Doupe AJ, Kuhl PK. Birdsong and humman speech: Common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Dunn AM, Zann RA. Undirected song in wild zebra finch flocks: contexts and effects of mate removal. *Ethology* 102: 529–539, 2010.

Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O. De novo establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–568, 2009.

Fougeron C, Keating PA. Articulatory strengthening at edges of prosodic domains. *J Acoust*

Soc Am 101: 3728–3740, 1997.

Gardner TJ, Naef F, Nottebohm F. Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308: 1046–1049, 2005.

Glaze CM, Troyer TW. Temporal structure in zebra finch song: Implications for motor coding. *J Neurosci* 26: 991–1005, 2006.

Goedemans R, **van der Hulst H**. Rhythm Types (2013). In: *The world atlas of language structures online*, edited by Dryer MS, Haspelmath M. Max Planck Institute for Evolutionary Anthropology: http://wals.info/chapter/17.

Gomez DM, Berent I, Benavides-Varela S, Bion RAH, Cattarossi L, Nespor M, Mehler J. Language universals at birth. *Proc Natl Acad Sci* 111: 5837–5841, 2014.

Greenberg J. Some universals of grammar with particular reference to the order of meaningful elements. *Universals Lang* 2: 73–113, 1963.

Jusczyk PW, Houston DM, Newsome M. The beginnings of word segmentation in English-learning infants. *Cogn Psychol* 39: 159–207, 1999.

Lachlan RF, van Heijningen CAA, ter Haar SM, ten Cate C. Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front Psychol* 7: 980, 2016.

Langus A, Mehler J, Nespor M. Rhythm in language acquisition. *Neurosci Biobehav Rev*, 2016.

Liberman M, Prince A. On stress and linguistic rhythm. *Linguist Inq* 8: 249–336, 1977.

Liberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition* 21: 1–36, 1985.

Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013.

MacNeilage PF, Davis BL, Kinney A, Matyear CL. The Motor Core of Speech: A Comparison of Serial Organization Patterns in Infants and Languages. *Child Dev* 71: 153–163, 2000.

Manly B. *Randomization, bootstrap and Monte Carlo methods in biology.* London, UK. Chapman & Hall, 1997.

Marler P, Peters S. Selective vocal learning in a sparrow. *Science* 198: 519–521, 1977.

Mol C, Chen A, Kager RWJ, ter Haar SM. Prosody in birdsong: A review and perspective. *Neurosci Biobehav Rev*, 2017.

Mooney R. Neural mechanisms for learned birdsong. *Learn Mem* 16: 655–669, 2009.

Patel A. *Music, language, and the brain.* Oxford, UK. Oxford University Press, 2010.

Plamondon SL, Goller F, Rose GJ. Tutor model syntax influences the syntactical and phonological structure of crystallized songs of white-crowned sparrows. *Anim Behav* 76, 1815–1827, 2008.

Podos J. Motor constraints on vocal development in a songbird. *Anim Behav* 51: 1061–1070, 1996.

Oller DK. The effect of position in utterance on speech segment duration in English. *J Acoust Soc Am* 54: 1235–1247, 1973.

Prather J, Okanoya K, Bolhuis JJ. Brains for birds and babies: Neural parallels between birdsong and speech acquisition. *Neurosci Biobehav Rev*, 2017.

Roper A, Zann R. The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* 112: 458–470, 2006.

Sakata JT, Brainard MS. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006.

Savage PE, Brown S, Sakai E, Currie TE. Statistical universals reveal the structures and functions of human music. *Proc Natl Acad Sci* 112: 8987–8992, 2015.

Scharff C, Nottebohm FN. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11: 2896–2913, 1991.

Senghas A, Coppola M. Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychol Sci* 12: 323–328, 2001.

Spierings MJ, ten Cate C. Zebra finches are sensitive to prosodic features of human speech. *Proc R Soc B Biol Sci* 281: 20140480, 2014.

Spierings M, Hubert J, ten Cate C. Selective auditory grouping by zebra finches: testing the iambic–trochaic law. *Anim Cogn* 20: 665–675, 2017.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Sturdy CB, Phillmore LS, Weisman RG. Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 113: 194–203, 1999.

Tchernichovski O, Nottebohm F. Social inhibition of song imitation among sibling male zebra finches. *Proc Natl Acad Sci* 95: 8951–8956, 1998.

Tchernichovski O, Lints T, Mitra PP, Nottebohm F. Vocal imitation in zebra finches is inversely related to model abundance. *Proc Natl Acad Sci* 96: 12901–12904, 1999.

Tchernichovski O, Nottebohm F, Ho C, Pesaran B, Mitra P. A procedure for an automated measurement of song similarity. *Anim Behav* 59: 1167–1176, 2000.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

ten Cate C. Behaviour-contingent exposure to taped song and zebra finch song learning. *Anim Behav* 42: 857–859, 1991.

ter Haar SM, Kaemper W, Stam K, Levelt CC, ten Cate C. The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*). *Proc R Soc B Biol Sci* 281: 20141860, 2014.

Thomason S, and Kaufman T. *Language contact, creolization, and genetic linguistics.* Berkeley, University of California Press, 1988

Tierney AT, Russo FA, Patel AD. The motor origins of human and avian song structure. *Proc Natl Acad Sci* 108: 15510–15515, 2011.

Vaissière J. Language-independent prosodic features. In *Prosody: Models and measurements*, 1983, 53–66.

Volman SF, Khanna H. Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 109: 211–221, 1995.

Vu ET, Mazurek ME, and Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14: 6924–6934, 1994.

Yang C, Crain S, Berwick RC, Chomsky N, Bolhuis JJ. The growth of language: Universal Grammar, experience, and principles of computation. *Neurosci Biobehav Rev*, 2017.

Zann RA. *The zebra finch: A synthesis of field and laboratory studies* Oxford, UK. Oxford University Press, 1996.

Zollinger SA, Suthers RA. Motor mechanisms of a vocal mimic: implications for birdsong production. *Proc R Soc B Biol Sci* 271: 483–491, 2004.

Figures

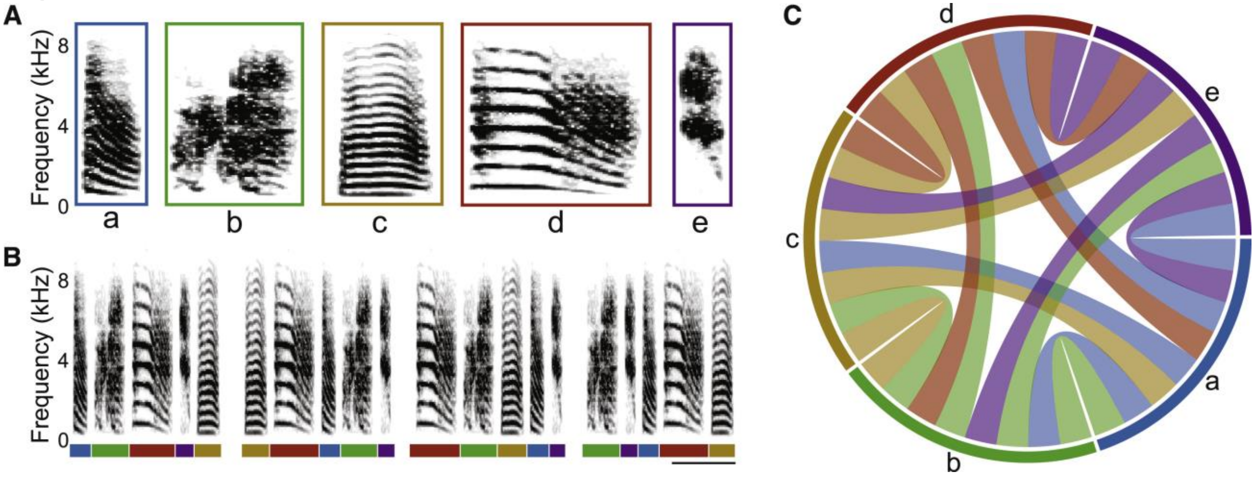


Figure 1. Experimental Tutoring with Randomized and Unbiased Sequences of Syllables

(A) Spectrograms of the five zebra finch “syllables” (i.e., contiguous epochs of sound) in the tutor stimulus set (arbitrarily labeled “a,” “b,” “c,” “d,” and “e”). These five syllables are acoustically distinct from each other and are types commonly observed in zebra finch songs.

(B) Example of tutoring song bout. Five zebra finch syllable exemplars were arranged into 120 possible five-syllable motif variants in which each syllable appears only once. Four distinct motif variants were presented in each song bout. Scale bar represents 250 ms.

(C) A circle plot summarizing transition probabilities between syllables within the motifs of the stimulus set. Outer-ring colors represent the five syllables, and line colors represent transitions from that syllable to others (e.g., blue lines represent transitions from “a” to other syllables). The width of the line is proportional to transition probability. All lines for the stimulus set have the same width because all transitions were heard with equal probabilities.

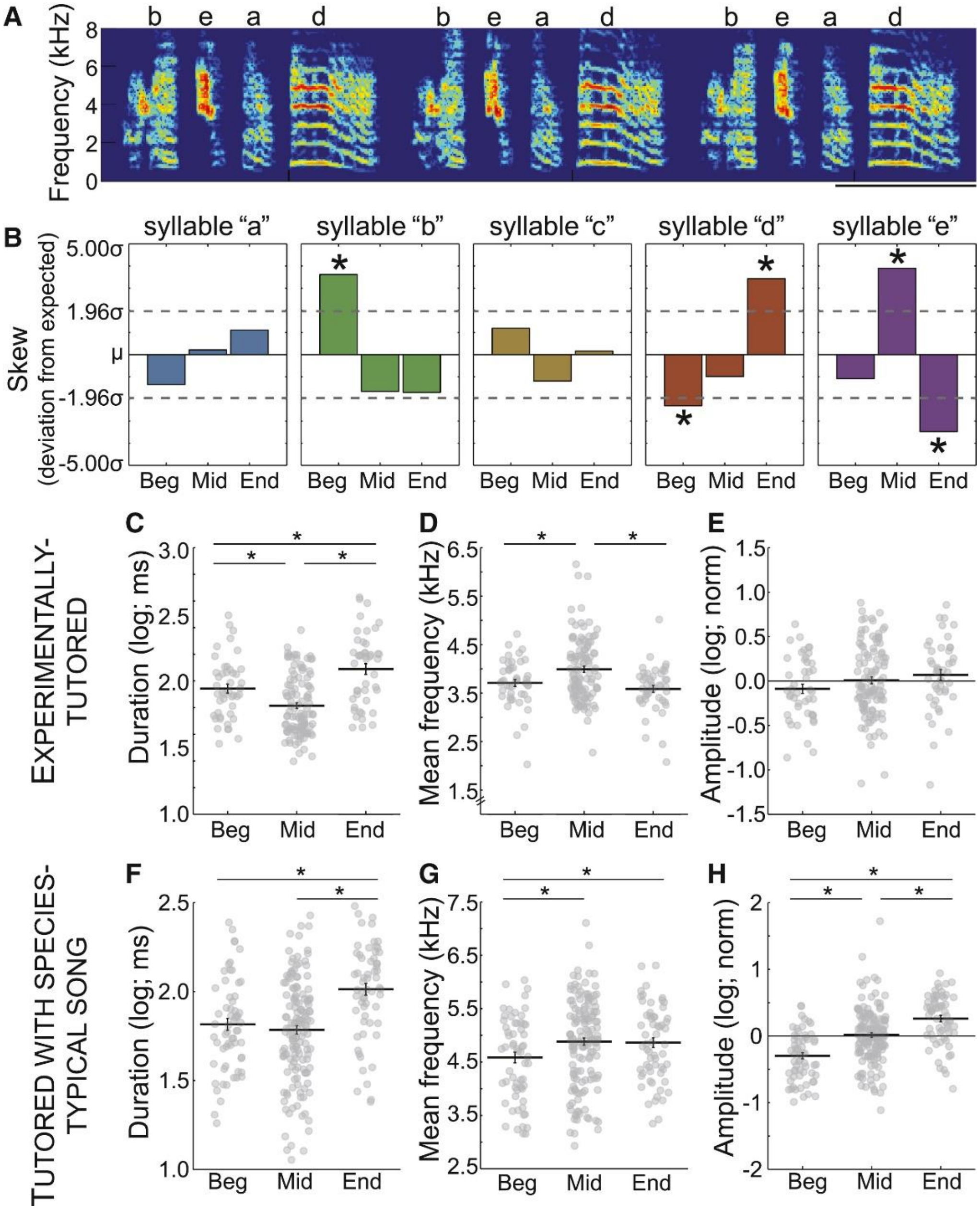


Figure 2. Skew in the Positioning of Syllables and Positional Variation in Acoustic Features

(A) Representative spectrogram of an experimentally tutored bird's song, which was composed of the motif "bead." Scale bar represents 250 ms.

(B) Bar graphs depicting the skew in the likelihood that individual syllables were positioned at the beginning ("Beg"), middle ("Mid"), or end ("End") of the motif (see STAR Methods). Skew was computed as the standardized difference between the observed number of times a syllable occupied a particular position and the number expected by chance (across 10,000 iterations of the Monte Carlo simulation). Skew is plotted for all five syllables across these three positions, and the significance of the skew was computed using Monte Carlo simulations ($p < 0.05$). Dashed lines indicate ±1.96 standard deviations ($p = 0.05$) from the expected (chance) probability (mean of the permutations; solid line). Tutored birds were significantly more likely to position the “b” syllable at the beginning of the motif ($p = 0.0009$), the “e” syllable in the middle ($p < 0.0001$), and the “d” syllable at the end ($p = 0.0014$). Birds were also significantly less likely to produce the “d” syllable at the beginning ($p = 0.0126$) and the “e” syllable at the end of the motif ($p = 0.0001$).

(C–E) The acoustic structure of syllables significantly differed across motif positions for experimentally tutored birds (beginning, middle, or end; n = 200 syllables). Syllables at the beginning or end of the motif were significantly longer in duration (C) and lower in pitch (D; mean frequency) than syllables produced in the middle ($p < 0.02$ for each). Syllables at the end of the motif were also significantly longer than syllables at the beginning of the motif ($p < 0.01$). Syllables did not significantly vary in amplitude across positions (E). Amplitude values for each bird are normalized (“norm”; mean subtracted) by the average amplitude of all syllables within his song.

(F–H) The acoustic structure of syllables also significantly varied across motif positions in the songs of birds tutored with species-typical song (n = 267 syllables; see STAR Methods). The duration of syllables significantly varied across positions (F), with syllables at the end of the motif being longer than syllables in the middle or at the beginning of the motif ($p < 0.0001$ for each). The pitch of syllables varied by position within the motif (G), with syllables at the beginning of the motif being lower in pitch than syllables in the middle or at the end of the motif ($p < 0.005$ for each). The mean amplitude of syllables significantly varied across positions in the motif (H), with syllables at the end of the motif being louder than syllables in the middle or at the beginning of the motif and syllables in the middle of the motif being louder than syllables at the beginning of the motif ($p < 0.003$ for each). Plotted are mean-subtracted amplitude values (see E).

(C–H) Dots represent median values for each syllable, horizontal lines indicate the means of the distributions, and vertical lines indicate standard errors of the distribution. For all panels, the asterisk (∗) denotes $p < 0.05$. See also Figures S1–S4 and Table S1.

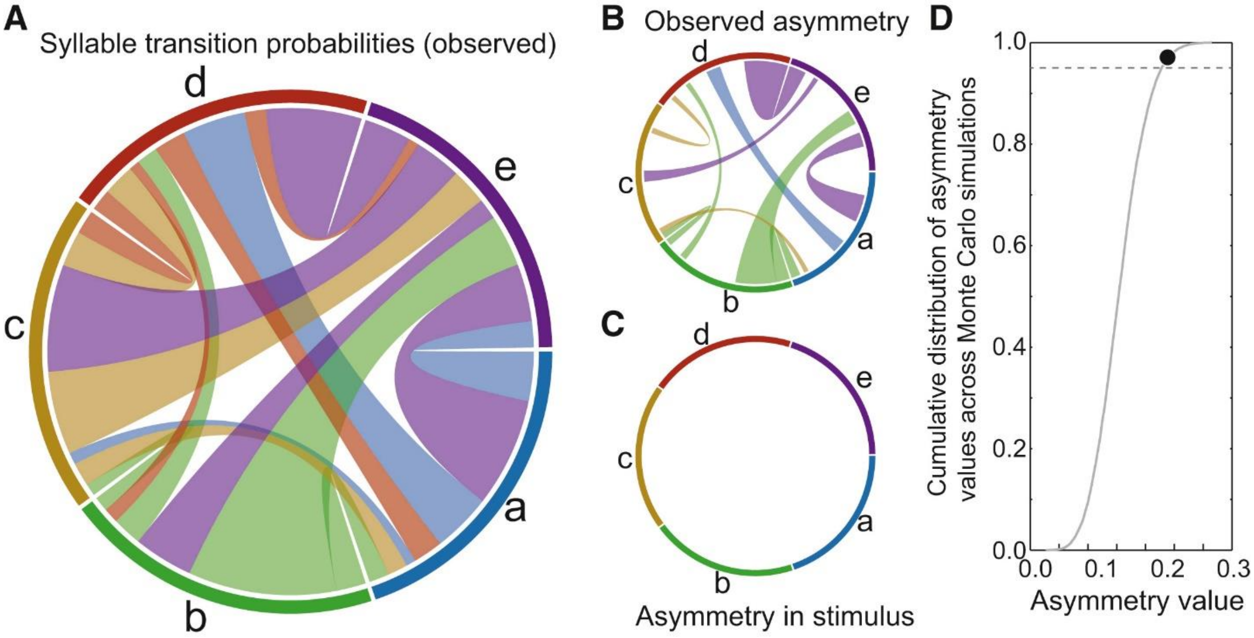


Figure 3. Significant Patterning of Syllable Transitions Following Tutoring with Randomized Sequences

(A) Circle plot depicts the syllable transition probabilities within motifs of tutored birds (see Figure 1C for tutoring stimulus).

(B and C) Asymmetries in the directionality of syllable transitions (e.g., difference in transition probabilities from “a” to “e” and from “e” to “a”) within the motifs of tutored birds (B) and tutoring stimulus (C). Line thickness scales with the magnitude of asymmetry in the directionality of syllable transitions (e.g., thick purple line between “e” and “a” in B represents a bias for transitions from “e” to “a” over “a” to “e”).

(D) Overall, there was a significant asymmetry in the directionality of syllable transitions ($p = 0.0312$; Monte Carlo simulations). Solid line depicts the cumulative distribution of overall asymmetry values found in the simulations. The dashed line represents the 95th percentile from the simulations, and the dot indicates the overall asymmetry value observed across experimentally tutored birds.

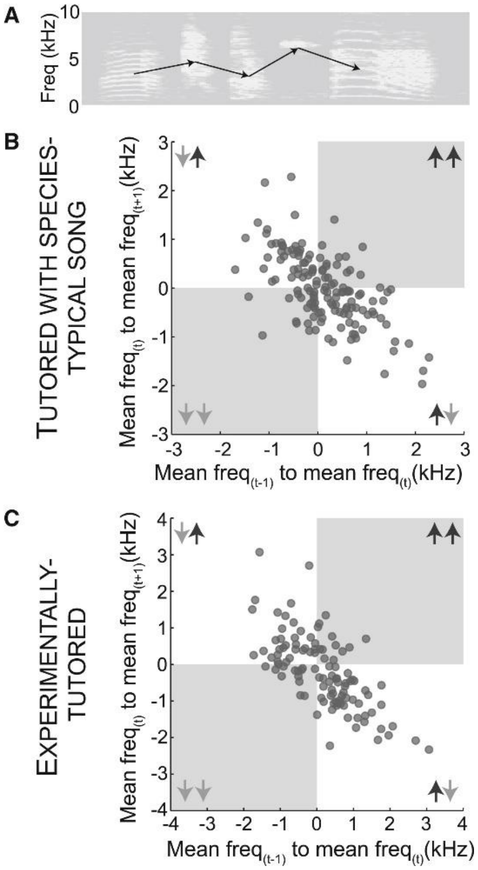


Figure 4. Mean Frequency Significantly Alternates across Syllables within Zebra Finch Motifs

(A) A bird's motif that demonstrates alternation across syllables. Arrows depict the change in mean frequency across adjacent syllables.

(B) The mean frequency of syllables significantly alternated within the motifs of birds tutored with species-typical song. Plotted is the change in mean frequency (in kHz) from the previous syllable$_{(t-1)}$ to the current syllable$_{(t)}$ (x axis) and the change from the current syllable$_{(t)}$ to the following syllable$_{(t+1)}$ (y axis). Arrows (in the corners) indicate the direction of change, with the first and second arrows of the pair indicating, respectively, change from syllable$_{(t-1)}$ to syllable$_{(t)}$ and change from syllable$_{(t)}$ to syllable$_{(t+1)}$. The significant negative relationship indicates that pitch alternated across syllables more than expected by chance.

(C) Birds tutored with randomized sequences of syllables also significantly alternated the mean frequency of syllables across the motif ($r^2 = 0.476$; $p = 0.0018$; Monte Carlo simulations). Data for experimentally tutored birds are plotted the same way in (B).

Appendix: Supplemental figures and table

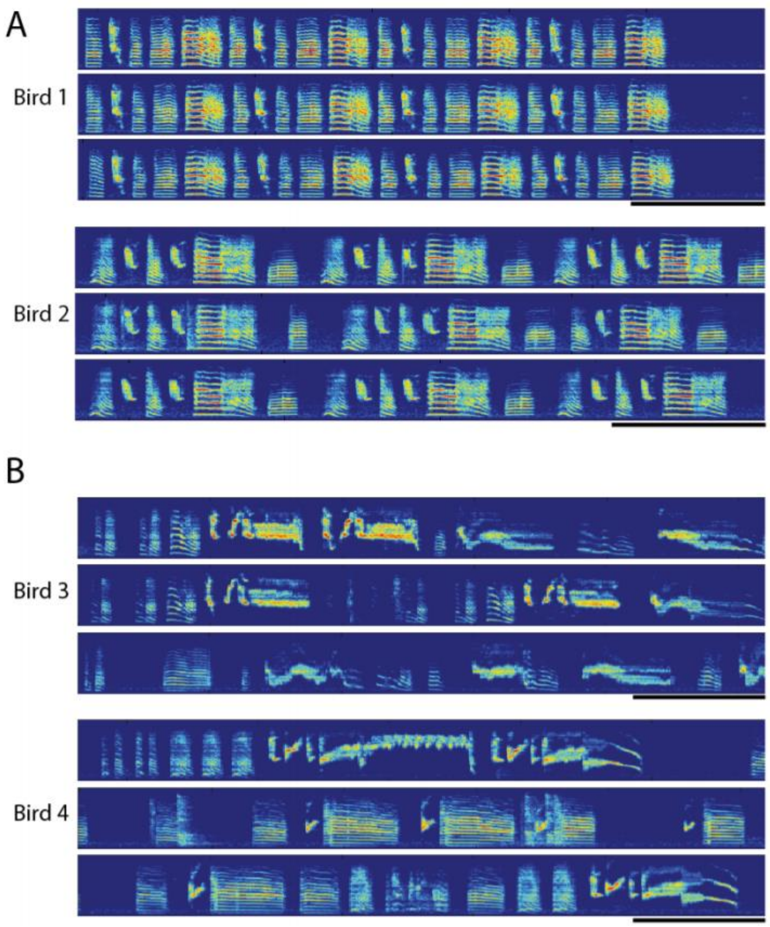


Figure S1: Examples of birds who were included and excluded from analysis (Related to Figure 2)

Spectrograms (time on the x-axis, frequency on the y-axis (0-10 kHz), color representing amplitude) of song bouts of example birds that were included (A) or excluded (B) from the analysis based on the ability to reliably detect stereotyped motifs.

(A) Representative spectrograms of three song renditions for two tutored birds ("Bird 1" and "Bird 2") that produced consistent motifs at 4 months of age.

(B) Examples of three song renditions from two tutored birds that produced variable sequences in their adult songs ("Bird 3" and "Bird 4"), making it difficult to reliably detect stereotyped motifs that are typical of zebra finch song. Scale bars = 500 ms.

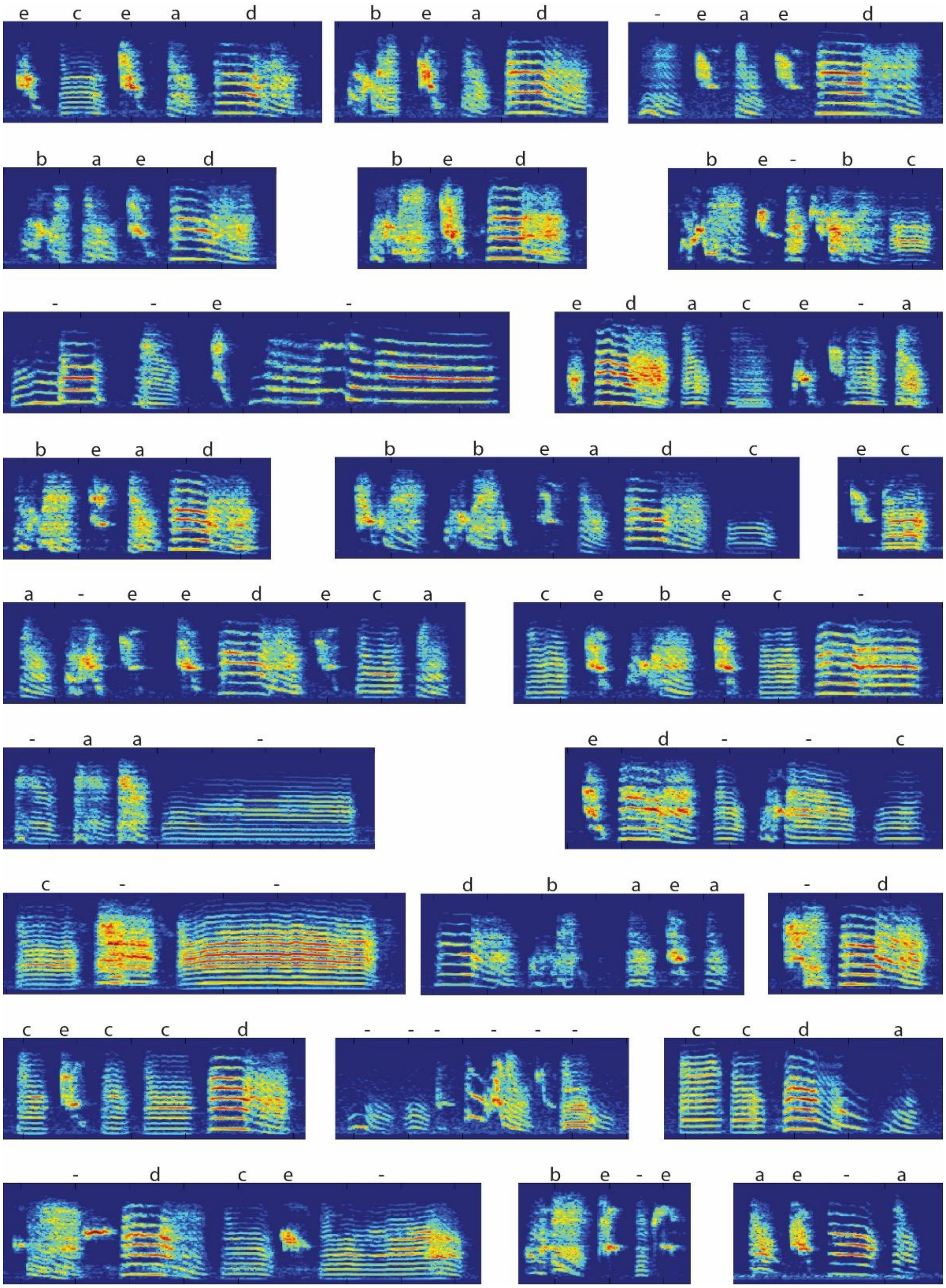


Figure S2: Example spectrograms of motifs of all birds included in the analysis part 1 (Related to Figures 2 and 3)

Letters and symbols above spectrograms denote the syllable labels used in our analysis ("-" denotes syllables that were not considered part of the tutoring set ("novel syllables")). N = 24 of 45 birds.

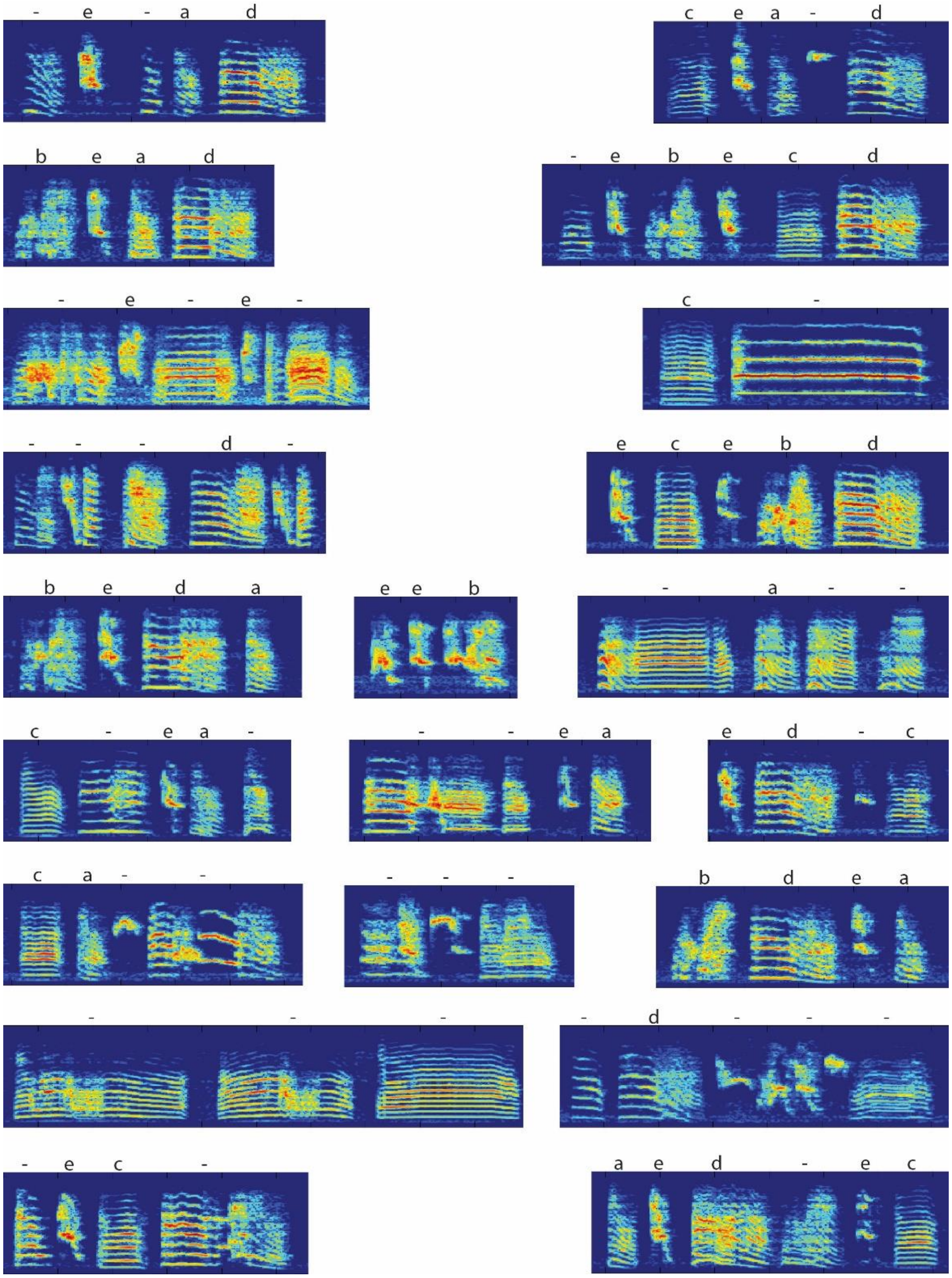


Figure S3: Example spectrograms of motifs of all birds included in the analysis part 2 (Related to Figures 2 and 3)

Letters and symbols above spectrograms denote the syllable labels used in our analysis ("-" denotes syllables that were not considered part of the tutoring set ("novel syllables")). N = 21 of 45 birds.

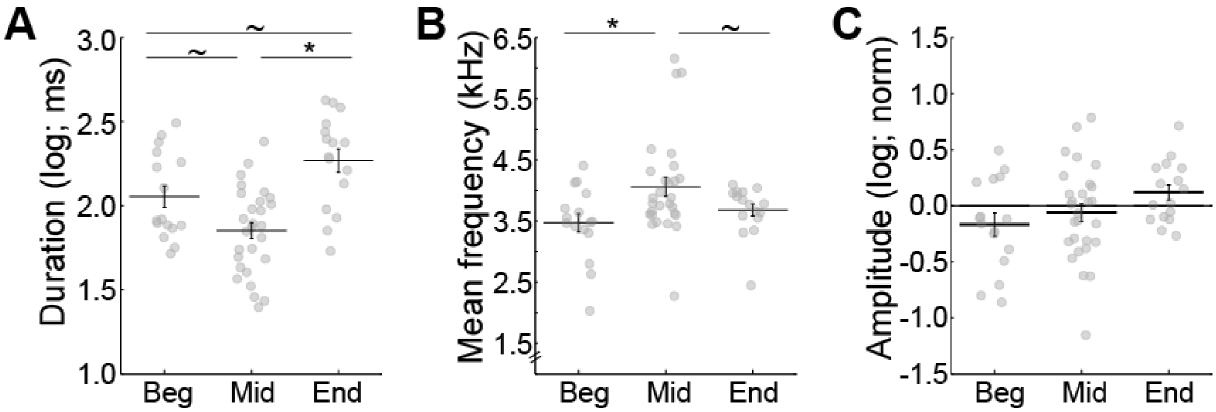


Figure S4: Positional variation in acoustic features of syllables that were produced by birds but not present in tutor stimulus (“novel syllables”; n = 60; Related to Figure 2).

(A) Novel syllables positioned at the end of the motif were significantly longer than those in the middle ($p < 0.0001$; Tukey’s HSD) and marginally longer than those at the beginning ($p = 0.0669$). Novel syllables at the beginning were marginally longer than those in the middle ($p = 0.0574$).

(B) Novel syllables positioned in the middle of the motif were higher in mean frequency than those at the beginning of the motif ($p=0.0084$) or those at the end of the motif ($p = 0.0542$).

(C). Amplitude of novel syllables did not vary across motif positions. Amplitude values for each bird are normalized (“norm”; mean-subtracted) by the average amplitude of all syllables within his song. For all panels, dots represent median values for each syllable, horizontal lines indicate the means of the distributions, “*” denotes $p < 0.05$, and “~” denotes $p < 0.07$.

Source	Comparison	Duration	Mean Frequency	Amplitude
Experimental	Beg vs Mid syllables	*Beg>Mid: 65.5%*	*Mid>Beg: 65.5%*	
	Beg vs End syllables	*End>Beg: 66.7%*		
	Mid vs End syllables	*End>Mid: 72.7%*	*Mid>End: 72.7%*	
ZF Library	Beg vs Mid syllables		*Mid>Beg: 64.1%*	*Mid>Beg: 66.9%*
	Beg vs End syllables	*End>Beg: 78.7%*	*End>Beg: 70.5%*	*End>Beg: 78.7%*
	Mid vs End syllables	*End>Mid: 67.6%*		*End>Mid: 68.3%*

Table S1: Summary of trends regarding significant differences in acoustic features between motif positions (Related to Figure 2). Values in the table represent the percent of syllables that followed patterns with respect to other syllables in their corresponding motifs. Percentages are reported on a per syllable basis (instead of, for example, on a per bird basis) because birds produce multiple middle syllables in their song; therefore, for this calculation, each beginning or end syllable is compare to each of the multiple middle syllables to compute the percent of comparisons that follow a trend. Percentages are only provided for contrasts that are significantly different (mixed effects model followed by Tukey's HSD, $p < 0.05$). "Experimental" birds refer to zebra finches that we tutored with randomized sequences (see Figure 1), whereas "ZF Library" birds refer to those tutored with species-typical songs that we obtained from an online repository (see STAR Methods).

Linking statement from chapter 4 to chapter 5

In Chapter 4, we discovered many common patterns across birds tutored with random sequences, indicating that zebra finches possess internal biases that guide developmental sequence learning. From here, we sought to better understand potential mechanisms that underlie such biases. We hypothesized that motor and/or auditory biases could contribute to these common patterns that we observed. To partially disentangle these effects, we compared the birds from Chapter 4 to birds that had impoverished auditory experiences throughout development, with our hypothesis being that the songs of these birds would be primarily the result of motor biases.

Chapter 5

Motor contributions to vocal learning biases in songbirds

Logan S. James, Ronald Davies Jr., Chihiro Mori, Kazuhiro Wada and Jon T. Sakata

In preparation for: *Proceedings of the Royal Society B*

Abstract

Biological predispositions in learning can bias and constrain the cultural evolution of social and communicative behaviors in a wide range of animals, including humans and songbirds, and such predispositions have been proposed to lead to the emergence of behavioral and cultural “universals”. For example, juvenile zebra finches that are individually tutored with randomized sequences of vocal elements (“tutored birds”) converge onto similar types of acoustic patterns that resemble universal patterns found in wild populations. Despite extensive demonstrations of learning biases, little is known about the biological mechanisms underlying predispositions in learning. Here we assessed the extent to which motor biases could contribute to the acquisition of universal acoustic patterns in zebra finch song. The songs of birds that are raised without auditory exposure to song throughout development (“untutored birds”) are thought to reflect motor biases in vocal production that could shape song acquisition. Consequently, we investigated the extent to which untutored birds produced songs that contained universal acoustic patterns exhibited by tutored birds. While the songs of untutored birds consisted of vocal elements (“syllables”) that had some distinct acoustic and sequencing features from those of tutored birds, the acoustic patterning within their songs was similar in a number of ways to that of tutored birds, including positional variation in the frequency, duration, and entropy of syllables within the song motif. This was true even for untutored birds that were unable to hear themselves throughout development. As such, these data provide compelling evidence for motor contributions to vocal learning and to the emergence of universals in culturally transmitted behaviors.

Introduction

Biological predispositions can sculpt the acquisition of many important behaviors. For example, the process of filial imprinting is critical for young animals to learn the identity of their caregivers, and young chicks are predisposed to imprint on visual stimuli that resemble adult conspecifics (Bolhuis and Trooster 1988; Di Giorgio et al. 2017; Horn and McCabe 1984; Sluckin and Salzen 1961). Relatedly, songbirds use learned songs during social interactions with conspecifics, and many species of songbirds demonstrate significant preferences to learn conspecific song over heterospecific song (Catchpole and Slater 2008; Doupe and Kuhl 1999; Wheatcroft and Qvarnström 2015). Many have theorized that internal biases guide the learning of human music, speech, and language and that these biases contribute to "universal" patterns across cultures (Chomsky 1980; Mehr et al. 2018; Yang et al. 2017). Indeed, controlled experiments provide support for a causal role of internal learning biases (separate from cultural propagation) in the acquisition and transmission of universal patterns in humans (Kirby 2017; Lumaca and Baggio 2017; Ravignani et al. 2016). However, despite the numerous documentation of learning biases, relatively little is known about the mechanisms that contribute to biases in learning and transmission.

Songbirds provide a powerful model system to not only reveal biological predispositions in learning but also to discover mechanisms that contribute to learning biases (Fehér et al. 2016; Gardner et al. 2005; Marler and Peters 1988; Plamondon et al. 2010). For example, when juvenile zebra finches are individually tutored with randomized and unbiased sequences of species-typical vocal elements ("syllables"), they produced convergent songs with number of common acoustic patterns (James and Sakata 2017). Birds tutored with randomized and unbiased sequences ("tutored birds") produced songs in which the syllable positioned at the end of the motif (i.e., the primary sequence that is repeated throughout the song) tends be the longest in duration and in which syllables positioned in the middle of the motif tend to be higher in mean frequency than syllables at the beginning or end of the motif. In addition, birds tutored in this manner produce songs in which the mean frequency of syllables alternate across adjacent syllables within the motif. Interestingly, this positional variation and alternation in acoustic structure mirror universal patterns observed across populations of zebra finches (Lachlan et al. 2016; Zann 1996) as well as universal patterns observed in human speech and music (Mol et al. 2017; Tierney et al. 2011). As such, understanding the mechanisms that contribute to vocal learning biases in songbirds like the zebra finch could provide insight into mechanisms that contribute to universals in speech and music.

A substantial component of song learning involves learning of how to produce the sounds memorized for acoustic communication (Brainard and Doupe 2002, 2013; Fee and Goldberg 2011; Mooney 2009; Murphy et al. 2017; Tschida and Mooney 2012), and the organization of central and peripheral aspects of the vocal motor system has been proposed to regulate various aspects of song learning. For example, species variation in morphological features of the vocal periphery have been found to covary with species variation in the spectral properties of song (Düring and Elemans 2016; Riede and Goller 2014; Suthers and Goller 1997). Additionally, neuromuscular and morphological aspects of the vocal motor system have been proposed to constrain an individual's ability to produce the temporal patterns of species-typical song (Gardner et al. 2005; Marler 1997; Mori et al. 2018; Mori and Wada 2015; Podos 1996; Podos et al. 2009; Price 1979).

To reveal motor contributions to universal patterns of zebra finch song (James and Sakata 2017; Lachlan et al. 2016; Zann 1996), we analyzed the songs of individuals with impoverished auditory experiences during the critical period for song learning ("untutored birds"; Marler 1997). This included birds that were simply isolated from any experience with an adult male song during the critical period for song learning ("untutored-hearing birds") as well as individuals that were isolated from adult male song and also prevented from hearing their own song during the critical period for song learning ("untutored-deafened birds"). Because these birds were not exposed to species-typical exemplars of song during development, the songs of these birds should reflect "innate" biases in song production that could sculpt the trajectory of vocal learning (Marler 1970, 1997). We specifically analyzed positional variation and alternation of various acoustic features in the songs of untutored birds, and compared these to universal patterns produced by tutored birds.

Materials and Methods

Animals and recording

To understand motor contributions to vocal learning biases and to the transmission of "universal" sequence patterns, we analyzed the song motif of zebra finches (*Taeniopygia guttata*) that remained untutored throughout development ("untutored birds"; n=22) and compared their song motif to those of zebra finches that were tutored with randomized and unbiased sequences of syllables ("tutored birds"; n=45; James and Sakata 2017). All animal experiments were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care (McGill University) or the Committee on Animal Experiments (Hokkaido University) in accordance with Japanese regulations for animal welfare (e.g., Law for the Humane Treatment and Management of Animals in Japan; after partial amendment number 68, 2005).

Tutored birds were tutored with synthesized songs that consisted of five canonical zebra finch syllables that were sequenced in a randomized and unbiased manner (see Supplementary Information for further description of stimuli; James and Sakata (2017) for full tutoring methods). Such a tutoring regime was used to reveal biological predispositions in vocal sequence learning. We analyzed recordings of tutored birds with readily identifiable "motifs" (see below for definition) that were 4 months old.

We compared the songs produced by tutored birds to those of birds that were not exposed to conspecific adult song throughout the developmental period for learning. Untutored birds were initially raised by both parents but raised by only their mother in a sound-attenuating chamber from 9.2 ± 0.5 (mean ± SEM) days of age (range: 5-12 days). Because only male zebra finches learn to produce complex songs, this protocol ensured that fathers were removed well before the opening of the critical period for song learning (Brainard and Doupe 2013; Roper and Zann 2006). When these juveniles could feed themselves, they were housed individually throughout development until their songs were recorded in adulthood ("untutored-hearing birds"; n=14 birds). Most (11 out of 14) of the untutored-hearing birds were recorded when they were 4-5 months old. Two of the three remaining birds were recorded at 11-12 months of age; the exact age of the third bird was unknown but he was recorded sometime between 5-12 months of age. We analyzed recordings of untutored-hearing birds from multiple lab populations that produced readily identifiable motifs (10 from the University of California, San Francisco, 2 from Hokkaido University, and 2 from McGill University; see Supplementary Information). The songs of some of these birds were included in distinct types of analyses in previous studies (Chen et al. 2016; Kojima and Doupe 2007; Mori and Wada 2015).

Untutored birds engage in vocal practice during development, and the act of hearing one's self during vocal practice could shape the nervous system and vocal patterning. Therefore, to further evaluate motor contributions to acoustic patterning, we also analyzed acoustic patterns within the songs of untutored birds that were deprived of auditory feedback from early in development, before the opening of the critical period for song learning ("untutored-deafened birds"). Untutored-deafened birds (n=8) were deafened before fledging (17–23 days post-hatch) by bilateral cochlear extirpation as described previously [Supplementary Information; see also Mori and Wada (2015) for details on experimental procedure]. After bilateral cochlear removal, the birds were returned to their nest until 32–41 days post-hatch, and were then housed in group cages with other deafened birds. An important requirement of our analysis is that birds produce a stereotyped motif, and while birds that are deafened early in development eventually produce stereotyped motifs, it takes longer for them to exhibit stereotyped vocal patterns (Mori and Wada 2015). We analyzed acoustic patterning within the motifs of untutored-deafened birds that were 11.7 ± 1.5 (mean ± SEM) months old. The songs of all untutored-deafened birds were described in previous publication dealing with the neurogenomic aspects of song development (Mori and Wada 2015) but the analyses were distinct from the current investigation.

All (n=67 total) were individually housed in sound-attenuation boxes during song recording (i.e., all songs are undirected songs) and were recorded using the Sound Analysis Pro program (Tchernichovski et al. 2000).

Analysis of acoustic features

Zebra finch songs consist of a single "motif" that is repeated within the song bout. Motifs in zebra finch song are readily identifiable because of the stereotypy of syllable sequencing within the motif, because gap durations between syllables within the motif are generally shorter and less variable than those between syllables outside the motif, and because adult zebra finches usually terminate their song bout at the end of a motif (Figure 1; see Supplementary Information). We analyzed variation in the acoustic features of syllables depending on their position in the motif (beginning, middle, or end) as well as alternation of these features across adjacent syllables. We measured the duration, mean frequency, mean amplitude, spectral entropy, spectrotemporal entropy, and amplitude entropy of each rendition of each syllable in every bird's motif (James and Sakata 2017; Piristine et al. 2016; Sakata and Brainard 2006; Stepanek and Doupe 2010; Wohlgemuth et al. 2010; see Supplementary Information for feature definitions), and then analyzed positional variation and alternation in the medians of these features (James and Sakata 2017). For the analysis of positional variation, we considered the first syllable of the motif as the "beginning syllable", the last syllable of the motif as the "end syllable", and syllables between the beginning and end syllables as "middle syllables".

Machine learning techniques

We used random forests (RFs), a machine learning algorithm, to classify syllables as beginning, middle or end syllables within each bird. Random forests were conducted in R with the randomForest package (Liaw and Wiener 2002), with the default settings except Ntree = 250. Prior to training on the algorithms, we normalized the data for each acoustic feature by subtracting the mean and dividing by the standard deviation. We trained each algorithm on 67% of the data (randomly selected) for an individual bird, then assessed the accuracy of the trained algorithm on a test set of data (the remaining 33%). Performance of the algorithm on the test set are reported in the manuscript.

To determine the contribution each feature to the categorization of position in the RF algorithm (for each bird), we extracted the “importance score” of each feature in the algorithm. The importance score represents the efficacy of a particular feature at classifying data (the greater the efficacy, the higher the importance score). This score is measured in arbitrary units; therefore, for our statistical analyses, we normalized the importance scores within each bird by dividing each importance score by the maximum importance score for that bird (i.e., “relative importance score”, in which the most important feature for that bird is scored as a “1”).

Statistical analyses

We used linear mixed-effects models (LMMs) using the R package lme4 (Bates et al. 2015) to analyze all data dealing with positional variation and relative importance scores. The data for duration and mean amplitude were highly right-skewed, so we log-transformed those data to improve normality. All three measures of entropy measures were highly left-skewed, and to normalize these data, we used a reflected log, which is defined as:

$$\text{transformedEntropy}_i = -\log(K - \text{Entropy}_i)$$

where “K” is defined as the maximum entropy in the dataset plus 1 and “i” refers to the i^{th} data point (Tabachnick et al. 2007).

For the analysis of positional variation in birds that were tutored with random sequences (“tutored birds”; as reported in James and Sakata 2017), we ran LMMs with Position (beginning, middle or end) as the independent variable and individual bird ID as a random variable (to account for the fact that multiple syllables were analyzed per bird). We also analyzed positional variation across different groups of birds (e.g., untutored-hearing vs. untutored-deafened), and these models included Group, Position, and the interaction between Group and Position as independent variables (i.e., full factorial model) and bird ID as a random variable.

Similar LMMs were used to analyze variation in relative importance scores. To analyze variation in relative importance scores across acoustic features in tutored birds, we ran an

LMM with Feature as the independent variable and bird ID as a random variable, and in analyses comparing variation in relative importance scores across different groups of birds, we ran LMMs with Group, Feature, and the interaction between Group and Feature as independent variables and bird ID as a random variable.

For all tests, a type II Wald chi-square test using the R package 'car' was used to determine the significance of each term (Fox et al. 2013), and Tukey's HSD tests with the Holm correction were used for post-hoc contrasts with the R package 'multcomp' (Hothorn et al. 2008).

We ran Monte Carlo simulations to assess the extent to which acoustic features alternated within each bird's motif (James and Sakata 2017). We first computed the observed differences in the median value of an acoustic feature across adjacent syllables. We plotted the change in an acoustic feature from the previous $\text{syllable}_{(t-1)}$ to the current $\text{syllable}_{(t)}$ against the change from the current $\text{syllable}_{(t)}$ to the next $\text{syllable}_{(t+1)}$ across all birds to measure the extent to which the feature alternated between adjacent syllables in the motif. If there is acoustic alternation from syllable-to-syllable, we would expect a negative relationship between the change from $\text{syllable}_{(t-1)}$ to $\text{syllable}_{(t)}$ and from $\text{syllable}_{(t)}$ to $\text{syllable}_{(t+1)}$. For each acoustic feature, we then computed the r^2 value of the relationship between the two changes ("observed r^2"). Finally, we compared this observed r^2 value to the distribution of r^2 values obtained from 10,000 Monte Carlo simulations in each the sequence of syllables within each bird' motif was randomized. Because the distribution from the Monte Carlo simulations represents the "null distribution" or the likelihood of randomly obtaining a particular r^2 value, we were able to compute the likelihood of obtaining an r^2 that was equal to or greater than the observed r^2, which represents the significance of the relationship. These analyses were conducted separately in tutored birds and in untutored birds.

Results

The principal aim of this study was to assess how motor biases could contribute to biological predispositions in the acquisition of acoustic patterning in the zebra finch song motif (James and Sakata 2017; Lachlan et al. 2016; Zann 1996). Zebra finches are the most extensively studied songbird, and the motif (a stereotyped sequence of syllables) is the fundamental unit of their song (Zann 1996). Because tutoring birds with randomized sequences of syllables reveals biological predispositions in the learning of acoustic patterns and "universals" in song production (James and Sakata 2017), we first extended existing analyses of acoustic patterning in zebra finches tutored with randomized sequences ("tutored birds"; n=45 birds). Thereafter, to assess motor contributions to the acoustic patterns observed in tutored birds, we conducted the same set of analyses in zebra finches with impoverished auditory experiences throughout development. Birds with impoverished auditory experiences were not exposed to conspecific song throughout development ("untutored birds") and could hear themselves ("untutored-hearing birds"; n=14 birds) or could not hear themselves ("untutored-deafened birds"; n=8 birds; Figure 1). The acoustic structure of syllables in the songs of untutored birds differed in a number of ways from song syllables in the songs of tutored birds (see below; Supplementary Information), and such differences reflect one of the consequence of vocal learning. Moreover, the songs of untutored birds are thought to reflect "innate" motor biases that can shape song learning (Marler 1970, 1997); therefore, we hypothesized that, if motor biases contributed to biological predispositions in the acquisition of acoustic patterning, untutored birds would produce songs with acoustic patterning that resembled those of birds tutored with randomized sequences of syllables.

Positional variation in features

The motif is the primary sequence of syllables that is repeated in the song of zebra finches, and birds tutored with randomized sequences of syllables produce motifs with systematic variation in the acoustic features of syllables across positions in the motif (James and Sakata, 2017). For example, the last syllable of the motif ("end syllable") of birds tutored with randomized sequences is longer in duration than the first syllable of the motif ("beginning syllable") and syllables produced between beginning and end syllables (i.e., "middle syllables"). In addition, beginning and end syllables are lower in pitch than middle syllables (Figure 2A-C). These patterns are also commonly observed in normally tutored birds and across wild populations of zebra finches (James and Sakata 2017; Zann 1996).

We extended these existing analyses of tutored birds by analyzing the degree to which measures of entropy varied across syllables occupying different positions in the motif (beginning, middle, or end; n=200 syllables). These measures of entropy (spectral, spectrotemporal, and amplitude entropy; see Supplementary Information) have previously been shown to change across development, to reliably categorize syllables, and to change with experimental manipulations of the nervous system (Piristine et al. 2016; Stepanek and Doupe 2010; Tchernichovski et al. 2001). There was significant variation across positions in the entropy of the amplitude envelope (i.e., amplitude entropy; linear mixed-effects model (LMM); $\chi^2_2=34.2$, $p<0.0001$; Figure 2D), with end syllables being higher in amplitude entropy than either middle (Tukey's HSD: $p<0.0001$) or beginning ($p=0.0001$) syllables, and in spectrotemporal entropy ($\chi^2_2=21.3$, $p<0.0001$; Figure 2E), with middle syllables being lower in spectrotemporal entropy than either end ($p<0.0001$) or beginning syllables ($p=0.0147$). There was no significant positional variation in the spectral entropy of syllables ($\chi^2_2=4.8$, $p=0.0906$; Figure 2F).

We next assessed the contribution of motor biases to such acoustic patterning by analyzing the extent of positional variation in untutored-hearing (n=61 syllables) and untutored-deafened (n=34 syllables) birds. Consistent with previous studies, some differences in acoustic features were observed between untutored-deafened and untutored-hearing birds (see Supplementary Information; Marler 1997), but importantly, no significant interactions between Group and Position were observed across all six features ($p>0.15$ for all). This indicates that both groups of untutored birds demonstrated similar types of positionally variation in acoustic features. Given the lack of significant interaction, this term was removed from the models for subsequent analyses.

There was significant positional variation in syllable durations within the songs of untutored birds (χ^2_2=7.6, p=0.0228; Figure 2G), and overall, end syllables were significantly longer than beginning syllables (p=0.0317) and middle syllables (p=0.0429).

The mean frequency of syllables of untutored birds significantly varied across positions in the motif (χ^2_2=11.6, p=0.0031; Figure 2H), with middle syllables being significantly higher in mean frequency than beginning syllables (p=0.0056) and end syllables (p=0.0419).

Amplitude entropy significantly varied across position in the motif (χ^2_2=11.4, p=0.0034; Figure 2J), with the amplitude entropy of end syllables being significantly higher than that of middle syllables (p=0.0027) and beginning syllables (p=0.0326). Similarly, there was a significant effect of position on spectrotemporal entropy (χ^2_2=14.8, p=0.0006; Figure 2K), with the spectrotemporal entropy of end syllables being significantly higher than that of beginning syllables (p=0.0031) and middle syllables (p=0.0009).

Syllable amplitude (χ^2_2=5.0, p=0.0837; Figure 2I) and spectral entropy (χ^2_2=2.6, p=0.2677; Figure 2L) did not significantly vary across positions within the motifs of untutored birds.

Many of the patterns observed in the motifs of untutored birds qualitatively resembled patterns observed in birds that were tutored with randomized sequences of syllables. For example, untutored and tutored birds produced motifs with end syllables that were longer in duration and lower in pitch than middle syllables. To statistically assess the extent to which these patterns differed, we analyzed the data from tutored and untutored birds simultaneously using the a full-factorial mixed effects model (both groups of untutored birds were merged because positional variation was not significantly different between hearing and deafened birds). While values for individual acoustic features differed between experimental groups (see Supplementary Information), importantly, there was no significant interaction between Group and Position within motif for any feature. Consequently, these analyses indicate that patterns of positional variation in acoustic features were comparable between untutored and tutored birds.

While positional variation in a number of acoustic features was similar between untutored and tutored birds, it should be emphasized that untutored and tutored birds produced songs that consisted of distinct acoustic syllable types (see Supplementary Information; Figure 1). For example, whereas tutored birds were biased to end their motifs with syllables that resembled ‘distance calls’ (James and Sakata 2017; Zann 1996), the end syllables of untutored birds would generally not be categorized as distance calls. In addition, the songs of untutored birds consist of syllables that are distinct from those used to tutor the tutored birds (Supplementary Information). Therefore, despite common acoustic patterning between untutored and tutored birds, tutoring leads to the production of songs with different syllables than the songs of untutored birds.

Machine learning techniques reveal the relative importance of acoustic features to predict syllable position within motifs

Given that a number of acoustic features significantly varied across motif positions for all three groups of birds, we sought to reveal the extent to which each acoustic feature predicted syllable position in the motif. For this, we employed a random forests (RF) algorithm to simultaneously analyze all acoustic features and reveal the feature(s) that provided the most predictive information about syllable position for that bird’s song (see Methods).

Overall, the RF algorithm performed well at predicting whether a syllable in an individual bird’s song was positioned in the beginning, middle or end of the motif (with accuracies >75% across each individual bird). Across all birds (n=67), the RF algorithms correctly predicted syllable position on 98.1±0.5% (mean±SEM) of the time (Figure 3A). While prediction accuracies were generally high across all groups of birds, prediction accuracies differed across tutored, untutored-hearing, and untutored-deafened birds ($\chi^2_2=17.0$, $p<0.0001$), with RF algorithms performing significantly better for tutored birds (98.9±0.3%) and untutored-hearing birds (98.6±0.6%) than for untutored-deafened birds (92.5±3.6%; $p<0.003$ for both contrasts).

Random forest algorithms allow one to extract information about the "importance" of each of the six features for predicting syllable position in each bird's motif ("relative importance score"; see Methods). Consequently, we extracted relative importance scores for each feature and analyzed differences in relative importance across features. We first ran these analyses for tutored birds and observed significant variation in the relative importance scores across features (χ^2_5=172.5, p<0.0001; Figure 3B). Duration was more important than all other features (e.g., spectrotemporal entropy) in predicting syllable position, and amplitude entropy was the least important feature.

We next assessed the pattern of relative importance scores for untutored birds. We ran a full-factorial mixed-effect model that included data for both groups of untutored birds (Feature, Group, and Feature x Group as independent variables) and observed a significant effect of Feature (χ^2_5=54.8, p<0.0001) with no significant effect of Group or Group x Feature. The latter indicates that the overall pattern of relative importance scores was not significantly different between untutored-hearing and untutored-deafened birds. Subsequently, we re-ran the model without the interaction term. Similar to tutored birds, duration was the most important feature, with post-hoc contrasts indicating that it was significantly more important than every other feature (χ^2_5=54.3, p<0.0001; Figure 3C).

Visual inspection of the distribution of relative importance scores across tutored and untutored birds suggest similarities as well as some differences in the relative importance of particular features in predicting syllable position. To quantitatively assess the degree to which the distribution of relative importance scores differed between tutored and untutored birds, we ran the same full-factorial model as described above but with tutored birds and untutored birds (because there was no interaction between untutored-hearing and untutored-deafened birds, they were combined into one group for this analysis). We observed a significant effect of Feature (χ^2_5=202.2, p<0.0001) and a significant interaction between Feature and Group (χ^2_5=16.1, p=0.0065). This significant interaction seems to be driven by the slightly different pattern of relative importance scores across tutored and untutored birds and by the relative importance of spectrotemporal entropy and amplitude entropy being higher for untutored birds than for tutored birds.

Alternation in acoustic features

Similar to speech and musical phrases, the motifs of birds tutored with randomized sequences of syllables consists of syllables that alternated in mean frequency (James and Sakata 2017; Figure 4A). However, consecutive syllables within the motifs of birds tutored with randomized sequences did not alternate in syllable duration or amplitude (James and Sakata 2017). As an extension of the previous analysis, we analyzed the degree to which the three measures of syllable entropy alternated across adjacent syllables in the motifs of tutored birds (see Methods) and observed significant alternation of amplitude entropy ($p=0.0020$; Figure 4B) and spectrotemporal entropy ($p=0.0005$; Figure 4C) but not of spectral entropy.

To analyze the extent to which motor biases could contribute to feature alternation, we examined the alternation of six acoustic features within the songs of untutored birds. (For this analysis, we combined the datasets for untutored birds because of the relatively small sample size for each group of untutored birds.) In contrast to the songs of tutored birds, the songs of untutored birds were not patterned in such a way that mean frequency ($p=0.5654$), amplitude entropy ($p=0.4809$), or spectrotemporal entropy ($p=0.4961$) alternated across consecutive syllables (Figure 4D-F). The other three features of song also did not significantly alternate within the songs of untutored birds ($p>0.20$ across all; data not shown).

Discussion

Biological predispositions in learning sculpt the acquisition of many socially important behaviours (Fehér et al. 2009; Kirby 2017; Kirby et al. 2008; Lumaca and Baggio 2017; Ravignani et al. 2016; Smith et al. 2017), but the mechanisms that underlie learning biases remain poorly understood. Tutoring juvenile zebra finches with randomized sequences of syllables revealed that biological predispositions bias vocal sequence learning (James and Sakata 2017), and here we investigated motor contributions to such vocal learning biases. Because the songs of untutored birds reflect “innate” predispositions in vocal production that could shape vocal learning (Marler 1997), we compared the acoustic patterning of song motifs of untutored birds with the acoustic patterning of song motifs of zebra finches tutored with randomized and unbiased sequences ("tutored birds"). Because auditory feedback is critical for song development, we investigated not only the songs of untutored birds that were able to hear themselves ("untutored-hearing birds") but also the songs of untutored birds that were deafened early in life ("untutored-deafened birds").

While untutored birds produced syllables that were spectrally distinct from those of tutored birds (Supplementary Information), both groups of untutored birds produced motifs with significant positional variation, and the magnitude and direction of this variation was similar between untutored birds that could or could not hear themselves. The lack of difference in acoustic patterning between the two groups of untutored birds contrasts with the numerous differences in the acoustic features of syllables that comprise the motifs of untutored birds that could or could not hear themselves (e.g., lower mean frequency; see also Marler 1997). Moreover, the patterns in positional variation for untutored birds resembled those for tutored birds. The same four acoustic features - syllable duration, mean frequency, and amplitude and spectrotemporal entropies – varied across motif position in untutored and tutored birds, and the pattern of positional variation was highly similar across birds. In particular, untutored and tutored birds produced syllables that are longer in duration and higher in amplitude entropy and spectrotemporal entropy at the end of their motif, as well as syllables with higher frequency in the middle of their motif. The fact that the positional variation in untutored birds resembles the patterns observed in tutored birds suggests that motor biases could contribute to vocal learning.

Motor biases have been proposed to influence song learning and could arise from neurobiological predispositions within forebrain, midbrain, and hindbrain circuits for vocal control (Brainard and Doupe 2002; Marler 1997; Marler and Sherman 1983, 1985; Mooney 2009; Mori and Wada 2015; Schmidt and Goller 2016) as well as biases in the vocal periphery (Larsen and Goller 1999, 2002; Podos 1996; Podos et al. 2009; Riede and Goller 2014; Suthers et al. 1999; Suthers and Goller 1997). Given that duration best predicted the position of syllables within the motif for both tutored and untutored birds, this suggests that neural, muscular, or morphological features that regulate respiratory patterns might be particularly influential in generating motor biases that shape vocal learning (Andalman et al. 2011; Cooper and Goller 2006; Schmidt and Goller 2016). For example, interactions between the dorsomedial nucleus of the intercollicular complex, hindbrain respiratory circuits, and HVC (used as a proper name) could be critical for creating motor biases that influence song learning.

Along with these proposed vocal motor biases, sensory (e.g., auditory) biases could also contribute to vocal learning biases. Auditory biases have been observed in songbirds using physiological, neural, and behavioural assays (Araki et al. 2016; Braaten and Reynolds 1999; Dooling and Searcy 1980; Nelson and Marler 1993; Prather et al. 2012; Whaling et al. 1997; Wheatcroft and Qvarnström 2015) and found to modulate the strength of learning (ter Haar et al. 2014; Marler 1997; Marler and Peters 1988; Mets and Brainard 2018). Just as auditory preferences for speech patterns and musical features have been proposed to contribute to universal patterns in speech and music (Bregman 1994; Trainor 2015; Werker and Hensch 2015), it is possible that the auditory system of zebra finches is biased to respond to songs with "universal" acoustic patterns and that songs that conform to universal patterns are more readily learned. As such, sensory biases could act in coordination with motor biases, reinforcing the same patterns observed in both tutored and untutored birds. It is also possible that auditory biases act independent of motor biases and could contribute to some of the differences between the songs of tutored and untutored birds (e.g., alternation of features).

While it is possible that non-auditory sources of feedback (e.g., somatosensory feedback) or early sensorimotor experiences (i.e., the production of unlearned calls) could contribute to the development of species-typical song patterns (Liu et al. 2009; Marler 2004; Suthers et al. 2002), these data suggest a mechanism by which the nervous system processes sensory inputs (especially aberrant sensory inputs) in order for the organism to produce species-typical song patterns. For example, when birds are tutored with sequences that are not species-typical or with stimuli that do not provide reliable information about syllable sequencing, our data suggests that the nervous system places sounds with particular acoustic features into positions in the motif that conform with motor biases. Such a model is consistent with previous observations that birds tutored with randomized and unbiased sequences can incorporate “novel syllables” (i.e., syllables that birds were not tutored with) into their motif and that these novel syllables were positioned with similar positional variation as learned syllables (James and Sakata 2017). More broadly speaking, it is possible that motor biases influence when and how specific sequences of syllables emerge over development in normally-reared birds (Lipkind et al. 2013; Tchernichovski et al. 2001).

Finally, these results support motor contributions to “universals” in the expression of a range of culturally-transmitted behaviours. Biological biases shape the cultural transmission of behaviors, including music, language and birdsong, and this process involves the regularization of sensory inputs by learners (Fehér et al. 2009; Ferdinand et al. 2019; Kirby 2017; Kirby et al. 2008; Lumaca and Baggio 2017; Ravignani et al. 2016; Tierney et al. 2011). Our results suggest that the organization of the motor system is involved in the regularization processes and that sensory inputs are transformed in a non-random manner to fit biases and constraints of the motor system (Boyd and Richerson 1988; Lumaca and Baggio 2017). This notion is consistent with the “gestural origins theory of speech” (Liberman et al. 1967) as well as related theories (e.g., frame-content theory) that propose that the building blocks and combinatorial structure of speech are derived from innate articulatory gestures (e.g., jaw movements for chewing) and pre-existing rhythmic behaviours of the jaw for eating, sucking, and breathing (MacNeilage 1998).

Acknowledgements

We thank M. Kao and S. Kojima for recordings for isolate birds from UCSF, and S.C. Woolley for input on data analysis and presentation. This work was supported by funding from the National Science and Engineering Research Council (#05016 to J.T.S.); MEXT/JSPS KAKENHI (JP17H06380 and JP19H04888 to K.W.), the Centre for Research on Brain, Language and Music (L.S.J.), and a Heller award (L.S.J.).

References

Andalman AS, **Foerster JN**, **Fee MS**. Control of vocal and respiratory patterns in birdsong: Dissection of forebrain and brainstem mechanisms using temperature. *PLoS One* 6, 2011.

Araki M, **Bandi MM**, **Yazaki-Sugiyama Y**. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–1287, 2016.

Bates D, **Maechler M**, **Bolker B**, **Walker S**. Package "lme4." *Convergence* 12, 2015.

Bolhuis JJ, **Trooster WJ**. Reversibility revisited: Stimulus-dependent stability of filial preference in the chick. *Anim Behav* 36: 668–674, 1988.

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

Braaten RF, **Reynolds K**. Auditory preference for conspecific song in isolation-reared zebra finches. *Anim Behav* 58: 105–111, 1999.

Brainard MS, **Doupe AJ**. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, **Doupe AJ**. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36: 489–517, 2013.

Bregman A. *Auditory scene analysis: The perceptual organization of sound*. MIT Press, 1994.

Catchpole CK, **Slater PJB**. *Bird song: biological themes and variations*. Cambridge, UK: Cambridge University Press, 2008.

Chen Y, **Matheson LE**, **Sakata JT**. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc Natl Acad Sci* 113: 6641–6646, 2016.

Chomsky N. Rules and representations. *Behav Brain Sci* 3: 1, 1980.

Cooper BG, **Goller F**. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J Neurophysiol* 95: 3798–3809, 2006.

Dooling R, **Searcy M**. Early perceptual selectivity in the swamp sparrow. *Dev Psychobiol* 13: 499–506, 1980.

Doupe AJ, **Kuhl PK**. Birdsong and humman speech: Common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Düring DN, **Elemans CPH**. Embodied motor control of avian vocal production. In:

Vertebrate sound production and acoustic communication. Springer, Cham, 2016, 119–157.

Fee MS, **Goldberg JH**. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* 198: 152–170, 2011.

Fehér O, **Ljubičić I**, **Suzuki K**, **Okanoya K**, **Tchernichovski O**. Statistical learning in songbirds: from self-tutoring to song culture. *Phil Trans R Soc B* , 2016.

Fehér O, **Wang H**, **Saar S**, **Mitra PP**, **Tchernichovski O**. De novo establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–568, 2009.

Ferdinand V, **Kirby S**, **Smith K**. The cognitive roots of regularization in language. *Cognition* 184: 53–68, 2019.

Fox J, **Friendly M**, **Weisberg S**. Hypothesis tests for multivariate linear models using the car Package. *R J* 5: 39, 2013.

Gardner TJ, **Naef F**, **Nottebohm F**. Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308: 1046–1049, 2005.

Di Giorgio E, **Loveland JL**, **Mayer U**, **Rosa-Salva O**, **Versace E**, **Vallortigara G**. Filial responses as predisposed and learned preferences: Early attachment in chicks and babies. *Behav Brain Res* 325: 90–104, 2017.

ter Haar SM, **Kaemper W**, **Stam K**, **Levelt CC**, **ten Cate C**. The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*). *Proc R Soc B Biol Sci* 281: 20141860, 2014.

Horn G, **McCabe BJ**. Predispositions and preferences. Effects on imprinting of lesions to the chick brain. *Anim Behav* 32: 288–292, 1984.

Hothorn T, **Bretz F**, **Westfall P**. Simultaneous inference in general parametric models. *Biometrical J* 50: 346–363, 2008.

James LS, **Sakata JT**. Learning biases underlie "universals" in avian vocal sequencing. *Curr Biol* 27: 3676-3682.e4, 2017.

Kirby S. Culture and biology in the origins of linguistic structure. *Psychon Bull Rev* 24: 118–137, 2017.

Kirby S, **Cornish H**, **Smith K**. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proc Natl Acad Sci* 105: 10681–10686, 2008.

Kojima S, **Doupe AJ**. Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J Neurophysiol* 98: 2099–109, 2007.

Lachlan RF, **van Heijningen CAA**, **ter Haar SM**, **ten Cate C**. Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front Psychol* 7: 980, 2016.

Larsen ON, **Goller F**. Role of syringeal vibrations in bird vocalizations. *Proc R Soc London Ser B Biol Sci* 266: 1609–1615, 1999.

Larsen ON, **Goller F**. Direct observation of syringeal muscle function in songbirds and a parrot. *J Exp Biol* 205: 25–35, 2002.

Liaw A, **Wiener M**. Classification and Regression by randomForest. *R news* 2: 18–22, 2002.

Liberman AM, **Cooper FS**, **Shankweiler DP**, **Studdert-Kennedy M**. Perception of the speech code. *Psychol Rev* 74: 431–461, 1967.

Lipkind D, **Marcus GF**, **Bemis DK**, **Sasahara K**, **Jacoby N**, **Takahasi M**, **Suzuki K**, **Feher O**, **Ravbar P**, **Okanoya K**, **Tchernichovski O**. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013.

Liu WC, **Wada K**, **Nottebohm F**. Variable food begging calls are harbingers of vocal learning. *PLoS One* 4, 2009.

Lumaca M, **Baggio G**. Cultural transmission and evolution of melodic structures in multi-generational signaling games. *Artif Life* 23: 406–423, 2017.

MacNeilage PF. The frame/content theory of evolution of speech production. *Behav. Brain Sci.* 21: 499–546, 1998.

Marler P. A comparative approach to vocal learning: Song development in white-crowned sparrows. *J Comp Physiol Psychol* 71: 1–25, 1970.

Marler P. Three models of song learning: evidence from behavior. *J Neurobiol* 33: 501–516, 1997.

Marler P. Bird calls: Their potential for behavioral neurobiology. In: *Annals of the New York Academy of Sciences*, 2004, 31–44.

Marler P, **Peters S**. The Role of Song Phonology and Syntax in Vocal Learning Preferences in the Song Sparrow, Melospiza melodia. *Ethology* 77: 125–149, 1988.

Marler P, **Sherman V**. Song structure without auditory feedback: emendations of the auditory template hypothesis. *J Neurosci* 3: 517–531, 1983.

Marler P, **Sherman V**. Innate differences in singing behaviour of sparrows reared in isolation from adult conspecific song. *Anim Behav* 33: 57–71, 1985.

Mehr SA, **Singh M**, **York H**, **Glowacki L**, **Krasnow MM**. Form and function in human song. *Curr Biol* 28: 356-368.e5, 2018.

Mets DG, **Brainard MS**. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proc Natl Acad Sci* 115: 421–426, 2018.

Mol C, **Chen A**, **Kager RWJ**, **ter Haar SM.** Prosody in birdsong: A review and perspective. *Neurosci Biobehav Rev* 81: 167–180, 2017.

Mooney R. Neural mechanisms for learned birdsong. *Learn Mem* 16: 655–669, 2009.

Mori C, **Liu W**, **Wada K.** Recurrent development of song idiosyncrasy without auditory inputs in the canary, an open-ended vocal learner. *Sci Rep* 8: 8732, 2018.

Mori C, **Wada K.** Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *J Neurosci* 35: 878–889, 2015.

Murphy K, **James LS**, **Sakata JT**, **Prather JF**. Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. *J Neurophysiol* 118: 800–816, 2017.

Nelson DA, **Marler P**. Innate recognition of song in white-crowned sparrows: a role in selective vocal learning? *Anim Behav* 46: 806–808, 1993.

Piristine HC, **Choetso T**, **Gobes SMH**. A sensorimotor area in the songbird brain is required for production of vocalizations in the song learning period of development. *Dev Neurobiol* 76: 1213–1225, 2016.

Plamondon SL, **Rose GJ**, **Goller F**. Roles of syntax information in directing song development in white-crowned sparrows (*Zonotrichia leucophrys*). *J Comp Psychol* 124: 117–32, 2010.

Podos J. Motor constraints on vocal development in a songbird. *Anim Behav* 51: 1061–1070, 1996.

Podos J, **Lahti DC**, **Moseley DL.** Vocal performance and sensorimotor learning in songbirds. In: *Advances in the Study of Behavior*, 2009, 159–195.

Prather JF, **Peters S**, **Mooney R**, **Nowicki S**. Sensory constraints on birdsong syntax: neural responses to swamp sparrow songs with accelerated trill rates. *Anim Behav* 83: 1411–1420, 2012.

Price PH. Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93: 260–277, 1979.

Ravignani A, **Delgado T**, **Kirby S**. Musical evolution in the lab exhibits rhythmic universals. *Nat Hum Behav* 1: 0007, 2016.

Riede T, **Goller F**. Morphological basis for the evolution of acoustic diversity in oscine songbirds. *Proc R Soc B Biol Sci* 281: 20132306–20132306, 2014.

Roper A, **Zann R**. The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* 112: 458–470, 2006.

Sakata JT, **Brainard MS**. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006.

Schmidt MF, **Goller F**. Breathtaking songs: coordinating the neural circuits for breathing and singing. *Physiology* 31: 442–451, 2016.

Sluckin W, **Salzen EA**. Imprinting and perceptual learning. *Q J Exp Psychol* 13: 65–77, 1961.

Smith K, **Perfors A**, **Fehér O**, **Samara A**, **Swoboda K**, **Wonnacott E**. Language learning, language use and the evolution of linguistic variation. *Philos Trans R Soc B Biol Sci* 372: 20160051, 2017.

Stepanek L, **Doupe AJ**. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Suthers R, **Goller F**, **Pytte C**. The neuromuscular control of birdsong. *Philos Trans R Soc London Ser B Biol Sci* 354: 927–939, 1999.

Suthers RA, **Goller F**. Motor correlates of vocal diversity in songbirds. In: *Current Ornithology*. Springer US, 1997, 235–288.

Suthers RA, **Goller F**, **Wild JM**. Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proc Natl Acad Sci* 99: 5680–5685, 2002.

Tabachnick B, **Fidell L**, **Ullman J**. *Using multivariate statistics*. Boston, MA: Pearson, 2007

Tchernichovski O, **Mitra PP**, **Lints T**, **Nottebohm F**. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

Tchernichovski O, **Nottebohm F**, **Ho C**, **Pesaran B**, **Mitra P**. A procedure for an automated measurement of song similarity. *Anim Behav* 59: 1167–1176, 2000.

Tierney AT, **Russo FA**, **Patel AD**. The motor origins of human and avian song structure. *Proc Natl Acad Sci* 108: 15510–15515, 2011.

Trainor LJ. The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philos Trans R Soc B Biol Sci* 370: 20140089–20140089, 2015.

Tschida K, **Mooney R**. The role of auditory feedback in vocal learning and maintenance. *Curr Opin Neurobiol* 22: 320–327, 2012.

Werker JF, **Hensch TK**. Critical periods in speech perception: new directions. *Annu Rev Psychol* 66: 173–196, 2015.

Whaling CS, **Solis MM**, **Doupe AJ**, **Soha JA**, **Marler P**. Acoustic and neural bases for innate recognition of song. *Proc Natl Acad Sci* 94: 12694–12698, 1997.

Wheatcroft D, **Qvarnström A**. A blueprint for vocal learning: auditory predispositions from brains to genomes. *Biol Lett* 11: 20150155, 2015.

Wohlgemuth MJ, **Sober SJ**, **Brainard MS**. Linked control of syllable sequence and phonology in birdsong. *J Neurosci* 30: 12936–49, 2010.

Yang C, **Crain S**, **Berwick RC**, **Chomsky N**, **Bolhuis JJ**. The growth of language: Universal Grammar, experience, and principles of computation. *Neurosci Biobehav Rev*, 2017.

Zann RA. *The zebra finch: A synthesis of field and laboratory studies*. Oxford, UK: Oxford University Press, 1996.

Figures

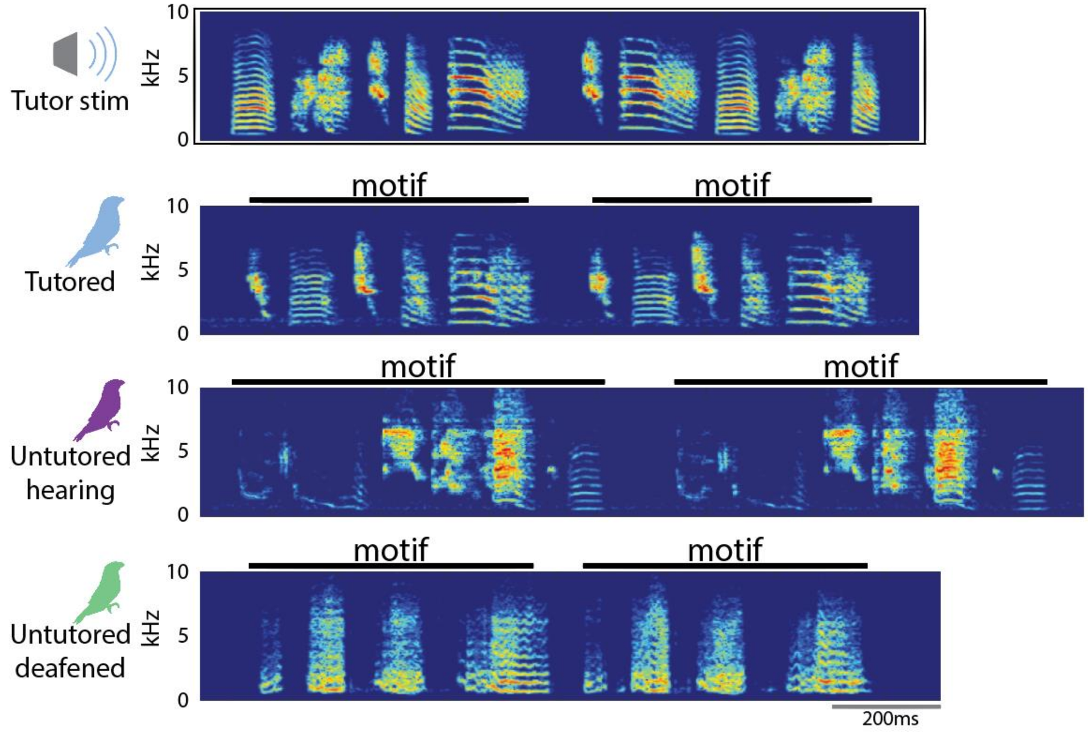


Figure 1: Example spectrograms of motifs from tutoring stimuli, tutored birds and untutored birds. Black bars above the spectrogram indicate the motif (i.e., repeated sequence of syllables). Bottom three rows depict, respectively, an example of song produced by a bird tutored with random sequences of syllables (James and Sakata 2017), an example of a song produced by a bird that remained untutored throughout development and was able to hear (untutored-hearing), and, an example of a song produced by a bird that remained untutored and was deafened before the critical period for song learning ("untutored-deafened").

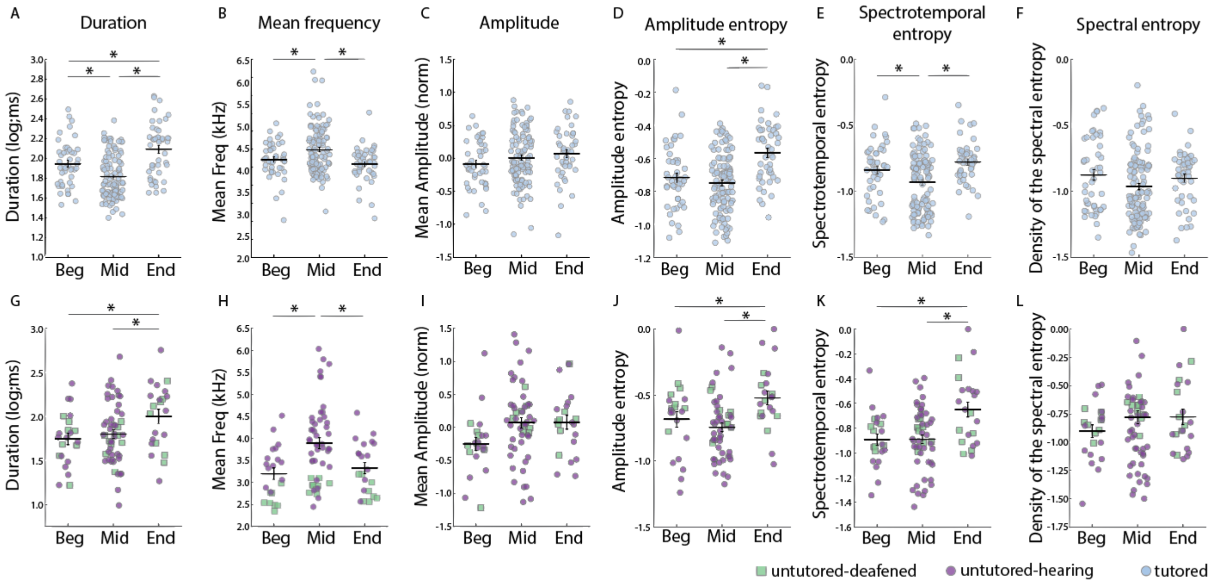


Figure 2. Positional variation of syllable features. Plotted are the features of each syllable produced by tutored birds (A-F) or untutored birds (G-L). Each point depicts the median value for a syllable produced by a bird. The horizontal line and error bar indicate the mean ± SEM. Asterisks indicate significant differences between positions (beginning, middle or end; $p<0.05$). (G-L) purple circles = untutored-hearing birds; green squares = untutored deafened birds. Amplitude was measured in arbitrary units (see Supplementary Information), and here we depicted the amplitude of syllables normalized to the maximum amplitude within the birds song (i.e., normalized amplitude) to better depict positional variation.

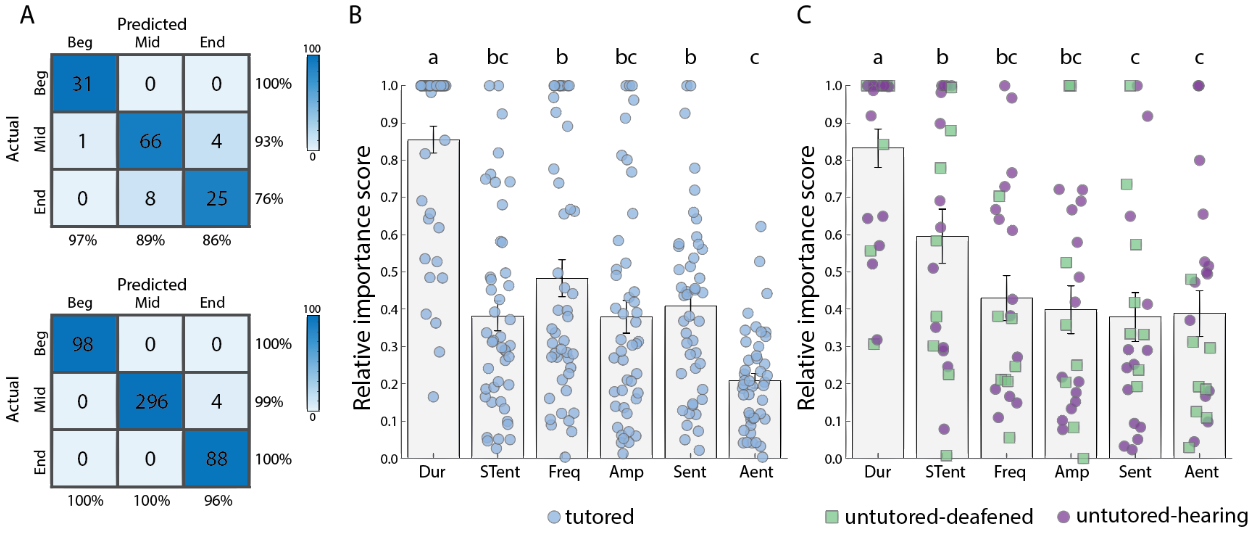


Figure 3: Machine learning approaches (random forests) reveal the relative importance of acoustic features in predicting syllable position across tutored and untutored birds. A. Two examples (from two different birds) of confusion matrices for the test sets of random forest (RF) algorithms. Each cell contains the number of syllables with a particular predicted and actual position. Cells along the diagonal indicate accurate predictions, and shading within each cell represents the percent of predictions that were accurate. To the right of the matrix is a summary of the percent of actual positions that were accurately predicted, and on the bottom of the matrix is a summary of the percent of predicted positions that were consistent with actual positions. B-C. The relative importance scores for each of the six features within a bird (Dur=duration, STent=spectrotemporal entropy; Freq=mean frequency; Amp=amplitude; Sent=spectral entropy; Aent= amplitude entropy). Tutored birds are depicted in (B) and untutored birds in (C). For (C), purple circles = untutored-hearing birds; green squares = untutored deafened birds. Features that share a letter above the bar indicate features that were not significantly different in relative importance. For example, duration was significantly different from all other features in both panels, whereas spectrotemporal entropy (STent) was significantly different from amplitude entropy (Aent) in untutored birds but not tutored birds.

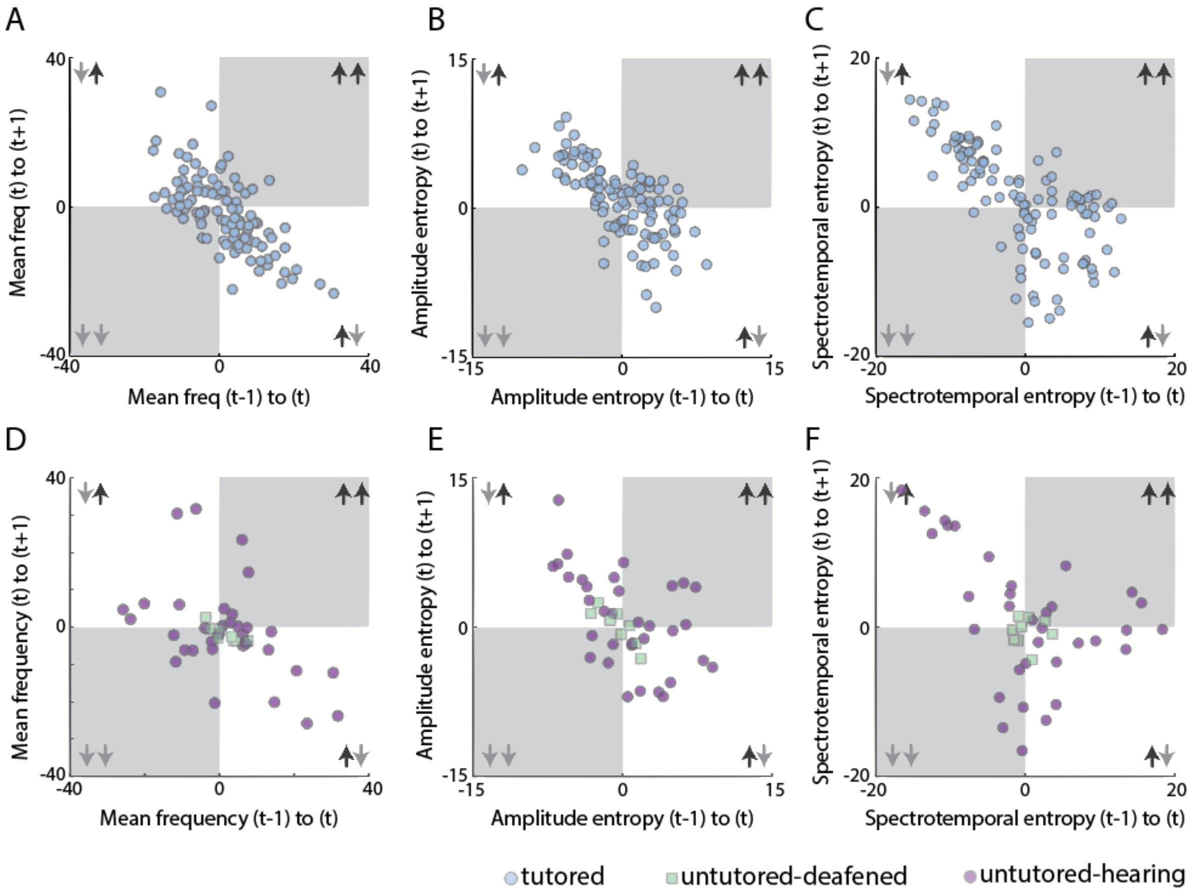


Figure 4: Alternation of acoustic features across adjacent syllables in the motif. The songs of tutored birds exhibited significant alternation of mean frequency (A; James and Sakata, 2017), amplitude entropy (B) and spectrotemporal entropy (C). However, none of these same features alternated within the songs of untutored birds (D-F).

Appendix: Supplementary Information

Tutoring

“Tutored birds” were tutored with randomized and unbiased sequences of syllables (James and Sakata, 2017). To this end, we constructed all possible five-syllable sequence (“motif”) variants that contained each of five species-typical syllables only once (n = 120 motif variants). Each song bout playback contained four motifs in a row. However, in contrast to normal zebra finch song bouts in which the same motif is repeated, synthesized song bouts consisted of four distinct motif variants. The sequence of motif variants per song bout was randomized across days for each bird and across birds. Moreover, to ensure that each motif variant was heard equally and randomly, we presented the stimuli in blocks of 120 motifs so that every motif variant was heard once before any variant was repeated.

The timing of syllables and motif variants in song stimuli followed a species-typical pattern: syllables within the motif were each separated by a silent gap of 30 ms, whereas the gap between syllables across adjacent motifs (i.e., from offset of last syllable in the motif to onset of first syllable of subsequent motif) was 100 ms.

Birds were operantly tutored using perch hops (Adret 1993; ten Cate 1991), and each perch hop triggered the playback of one song bout. Song playbacks were spaced out such that juveniles could hear only 10 operantly triggered song playbacks (i.e., 40 motif variants) within each of three time periods in the day (morning, noon, and afternoon). Birds were housed individually throughout tutoring and until 4 months of age.

Sound Analysis Pro 2011 (SAP; http://soundanalysispro.com) was used for song tutoring. Stimuli were played out of an Avantone Pro Mixcube speaker (Avantone, NY) connected to a Crown XLS 1000 amplifier (Crown Audio, IN). Song playback was operantly triggered by perch hops using custom-built perches connected to a National Instruments PCI-6503 I/O card (National Instruments, TX).

Removal of auditory feedback

Untutored-deafened birds (n=8) were deafened before fledging (17–23 days post-hatch) by bilateral cochlear extirpation as described previously [see Mori and Wada (2015) for details on experimental procedure]. Briefly, each bird was anesthetized with pentobarbital (6.48 mg/ml; 60 ul/10 g body weight) by intraperitoneal injection. After fixing the head in a custom-made stereotaxic apparatus with ear bars, a small window was made through the neck muscle and the skull near the end of the elastic extension of the hyoid bone. A small hole was then made in the cochlear dome, and the cochlea was pulled out with a fine hooked wire. The removed cochleae were confirmed by visual inspection under a dissecting microscope. After bilateral cochlear removal, the birds were returned to their nest until 32–41 days post-hatch, and were then housed in group cages with other deafened birds.

Identifying song motifs

We analyzed the adult songs (n>30 songs/bird) of tutored and untutored birds. We identified and labeled syllables and motifs following automated, amplitude-based segmentation of audio files. Zebra finch songs consist of syllables that are arranged into stereotyped sequences called "motifs", and each adult zebra finch repeats a single motif within his song. An individual's motif is readily identifiable because it is repeated multiple times within a song bout (ten Cate 1991; Scharff and Nottebohm 1991; Tchernichovski et al. 2000; Zann 1996) and because the duration of gaps between motifs is usually longer and more variable than gap durations within the motif (Glaze and Troyer 2006; Vu et al. 1994). To identify motifs and birds for analyses, two authors independently examined multiple renditions of each bird's song. Birds were included in the analysis only if both researchers identified and agreed upon a stereotyped motif.

Acoustic feature definitions

Syllable duration is defined as the interval from syllable onset to offset. Mean frequency is the power-weighted average frequency calculated over the spectrum of the entire syllable and is defined as:

$$\Sigma_f f^* p(f) / \Sigma_f p(f)$$

where f is frequency, p(f) is power as a function of frequency in the discrete Fourier transform of the entire syllable, and the sum is over the frequency range 0.3-8 kHz. Amplitude is computed as the area under the squared waveform of the syllable, smoothed with a 2-ms gaussian window. Spectral entropy is the entropy of the Fourier transform of the entire syllable and is defined as:

$$\Sigma_f -p(f)*\log_2 p(f)$$

where p(f) is the power at each frequency divided by the power of the entire syllable, and the sum is over the range 0.3-8 kHz. Amplitude entropy is the entropy of the normalized distribution of power as a function of time over the course of the syllable and is defined as:

$$\Sigma_t -p(t)*\log_2 p(t)$$

where p(t) is power at each time point divided by the power of the entire syllable and the sum is over the duration of the syllable. Spectrotemporal entropy is the entropy of the distribution of power in the spectrographic representation of the syllable and is defined as:

$$\Sigma_f\Sigma_t -p(f,t)*\log_2 p(f,t)$$

where p(f,t) is the power at each point in the time-frequency representation of the syllable divided by the total power of the syllable. For this measure, the spectral representation was derived using non-overlapping 8-ms Hanning windows to segment the time waveform of the syllable. All measures of entropy were normalized such that sounds with qualities similar to pure tones have values approaching 0 and sounds similar to white noise have values approaching 100.

Differences in song between birds tutored with randomized sequences and birds with impoverished auditory experiences

We compared the acoustic patterns of songs of birds that were tutored with randomized sequences of syllables (“tutored birds”) and of birds with impoverished auditory experiences throughout development (“untutored birds”). We analyzed the songs of untutored birds that could hear themselves throughout development (“untutored-hearing” birds) and untutored birds that were deprived of hearing throughout development (“untutored-deafened” birds). The songs of such untutored birds were analyzed because their songs are hypothesized to reflect biases in vocal production (Marler 1997). Overall, some acoustic features of syllables were different between untutored-hearing and untutored-deafened birds, whereas we were unable to detect significant differences between the motif syllables of untutored-hearing and tutored birds (Supplementary Table 1). For example, the motifs of untutored-deafened birds consisted of syllables that were lower in mean frequency than those of untutored-hearing and tutored birds and shorter in duration than those of tutored birds. These differences between deafened and hearing birds are consistent with previous song analyses (e.g., Price 1979; Scharff and Nottebohm 1991). The songs of untutored-deafened birds were also lower in amplitude than other birds, but this could be because of variation in recording condition.

Supplementary Table 1
Main effect of group [tutored (T), untutored-deafened (UD), untutored-hearing (UH)] and post-hoc contrasts (Tukey's HSD) for our six acoustic features.

Feature	DF	χ^2	p	Significant contrasts
Duration	2	5.8	0.0556	T>UD (p=0.0444)
Mean frequency	2	60.1	<0.0001	T>UD (p<0.0001); UH>UD (p<0.0001)
Mean amplitude	2	34.3	<0.0001	UH>UD (p=0.0002); T>UD (p<0.0001)
Amplitude entropy	2	1.9	0.3839	
Spectrotemporal entropy	2	3.4	0.1826	
Spectral entropy	2	5.3	0.0724	

Despite that we observed few overall differences between the acoustic features of the motif syllables between untutored and tutored birds, the songs produced by these two groups were notably distinct (Figure 1). For instance, the songs of tutored birds were far more similar to the tutor stimulus than the songs of either untutored-hearing or untutored-deafened birds. To quantify this, we computed the acoustic similarity (similarity score in Sound Analysis Pro; Tchernichovski et al. 2000) of ≥30 songs per bird to the tutoring stimulus motif. Tutored birds produced songs were significantly more similar to the tutoring stimulus [55.9 ± 2.6% (mean±SEM)] than the songs of untutored-hearing (25.8 ± 3.8%) and untutored-deafened birds (22.1 ± 3.1%; $F_{2,64}$=28.2, p<0.0001; Tukey's HSD: tutored significantly higher than both untutored groups (p<0.0001 for both)). Furthermore, isolate zebra finch song is often characterized by atypically long and structurally distinct syllable types (Fehér et al. 2009; Price 1979; Volman and Khanna 1995), and we observed similar syllable types in our untutored-hearing birds, albeit not within their motifs.

References

Adret P. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim Behav* 46: 149–159, 1993.

ten Cate C. Behaviour-contingent exposure to taped song and zebra finch song learning. *Anim Behav* 42: 857–859, 1991.

Fehér O, **Wang H**, **Saar S**, **Mitra PP**, **Tchernichovski O**. De novo establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–568, 2009.

Glaze CM, **Troyer TW**. Temporal structure in zebra finch song: Implications for motor coding. *J Neurosci* 26: 991–1005, 2006.

Marler P. Three models of song learning: evidence from behavior. *J Neurobiol* 33: 501–516, 1997.

Mori C, **Wada K**. Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *J Neurosci* 35: 878–889, 2015.

Price PH. Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93: 260–277, 1979.

Scharff C, **Nottebohm FN**. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11: 2896–2913, 1991.

Tchernichovski O, **Nottebohm F**, **Ho C**, **Pesaran B**, **Mitra P**. A procedure for an automated measurement of song similarity. *Anim Behav* 59: 1167–1176, 2000.

Volman SF, **Khanna H**. Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 109: 211–221, 1995.

Vu E, **Mazurek M**, **Kuo Y**. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14: 6924–6934, 1994.

Zann RA. *The zebra finch: A synthesis of field and laboratory studies*. Oxford, UK: Oxford University Press, 1996.

General conclusions

Summary

Through this thesis, we have provided novel understanding of how vocal communication is learned across development and subsequently modified over short- and long-term scales in adults. In chapter one, we revealed age-dependent plasticity of adult birdsong in the Bengalese finch. We found longitudinal increases in the tempo of song and decreases in the variability of the timing and sequencing of song. The nature of these changes is typical of the motor changes that generally occur through practice and reinforcement mechanisms (e.g., Adams 1987; Ericsson et al. 1993; Magill 2004), suggesting that these adult Bengalese finches are continuing to refine their songs across time. Furthermore, these results documented vocal motor plasticity beyond the general sensitive period for song learning in development.

In chapter two we hypothesized that short term context-dependent modulations of adult song could predict the changes we observed across age in chapter one, given the fact that previously reported context-dependent changes resembled the changes across age (Dunning et al. 2014; Hampton et al. 2009; Heinig et al. 2014; Sakata et al. 2008). In support of this, we found that a number of song features changed similarly between social context in young adults and across age within those same birds. In particular, the overall direction and magnitude of change for song tempo and sequencing were similar across age and social context. Furthermore, individual variation in the changes across age for song tempo, sequencing and spectral structure could be predicted by individual variation in the changes across social context.

In chapter three, we employed a comparative approach to assess whether the age-dependent changes to adult song, and the predictability of those changes observed in Bengalese finches, were similar in the closely related zebra finch (Murphy et al. 2017). We found that zebra finches similarly increased the tempo of their song over time and decreased the variability of syllable sequencing. Zebra finches also increased the length of their songs over time, and this was also the most predictable change from the modulation of song across social context. Furthermore, we found that developmental experiences did not affect many of the age-dependent changes to song.

In chapter four we investigated the predictability of developmental song learning by using a large tutoring experiment to ask whether zebra finches are biased to produce certain types of sequences. For this, we individually tutored birds with random sequences of five syllables and observed which sequence each bird learned and produced. We found many common patterns across these birds, including biases in the positioning of particular syllable types, and the transitions between those types. Furthermore, we found positional biases in various acoustic features such as longer durations at the end of song motifs and higher mean frequencies in the middle of song motifs. Interestingly, these acoustic patterns resemble common patterns observed in human speech and music (Fougeron and Keating 1997; Oller 1973; Patel 2010; Vaissière 1983).

Finally, in chapter five we began to uncover the mechanisms that underlie the learning biases we observed in chapter four. In particular, we sought to highlight the motor contributions to these biases by observing the song patterns produced by birds with impoverished auditory experiences ("untutored birds"). We found that birds that were untutored, regardless of whether they could hear or not, generally produced songs that followed some of the patterns we observed in the birds that were tutored with random sequences. For instance, these birds also produced syllables with higher mean frequency in the middle of the song motif, and syllables with longer duration at the end of the song motif. Through a machine learning analysis, we also found that duration appears to be a key acoustic feature in the organization of song for both tutored and untutored birds. These data suggest that the motor system contributes to the observed biases in sequencing in zebra finches.

Birdsong and human speech

These results further demonstrate the utility of songbirds as a model for human speech learning and production. In the first three chapters, we observed many long-term longitudinal changes to various song features in adult Bengalese and zebra finches beyond the typical sensitive period for song learning. Such plasticity in motor behaviours is often characterized by decreases in the variability of vocal production, which is what we observed for various features in both species. Similarly, variability in human speech production also decreases across time (Gerosa et al. 2007; Sankoff 2018; Smith et al. 1996), suggesting that similar age-dependent processes may be at work in both humans and birds. Because of the vast number of songbirds, this model system presents the opportunity for large, cross-species comparisons of motor plasticity in adulthood to identify common patterns (i.e., “universals” of vocal motor plasticity). Our species comparison between the Bengalese finch and zebra finch identified that both species increased their speed of production and decreased sequence variability across time. Furthermore, both species demonstrated some degree of link between short- and long-term changes to syllable sequencing. Given that song variability correlates with age in a variety of species (Botero et al. 2009; Ota and Soma 2014; Pytte et al. 2007; Rivera-Gutierrez et al. 2012), whether the increase in timing and predictability of changes to syllable sequencing are specific to these two species or represent more generalizable patterns will be of great interest for further research.

In the fourth chapter, we observed biases in vocal learning that led zebra finches to learn to produce certain types of sound sequences over others. Of particular interest is the fact that some of these biases (e.g., longer productions at the ends of phrases and alternations of pitch) are also observed in the speech and music patterns of humans (Bickel 2013; Chomsky 1980; Goedemans and van der Hulst 2013; Patel 2010; Savage et al. 2015; Tierney et al. 2011). This suggests the possibility that such vocal patterns could arise from fundamental neural or anatomical building blocks shared by humans and songbirds. Furthermore, the fact that we observed motor contributions to these learning biases in chapter five suggests that similar motor biases may underlie common patterns in human speech across languages. Again, the power of the songbird model system for large, cross-species analyses will be of great interest for future research into fundamental features of vocal learning and production.

Neurobiology of song learning and production

A powerful feature of the songbird model system is the ability for detailed research into the neurobiology of vocal learning, production and plasticity. Throughout this thesis we have identified many aspects of birdsong learning and plasticity that can provide important foundations for future neurobiological research. Birdsong is controlled primarily by two forebrain circuits: the vocal motor pathway (VMP) and anterior forebrain pathway (AFP; Brainard and Doupe 2002). Neurons in the VMP, which includes the HVC (used as proper name) and the robust nucleus of the arcopallium (RA), encode the motor commands for song and are functionally analogous to neurons in the mammalian premotor, supplementary, and primary motor cortical areas (Doupe and Kuhl 1999; Fee and Scharff 2010). The VMP is particularly important for the control of temporal features of song, including syllable timing and sequencing (Hahnloser et al. 2002; Kozhevnikov and Fee 2007; Zhang et al. 2017). As such, our data on acute and long-term changes to syllable timing and sequencing can provide insights into changes in these circuits and the mechanisms underlying behavioural plasticity and biases in learning.

From the first two chapters, we found that aspects of long-term changes to syllable sequencing, structure and timing in the Bengalese finch could be predicted by short-term social modulation of song, and we proposed that context- and age-dependent changes to syllable sequencing at branch points and to gap durations could be mediated by neurophysiological changes in HVC. Future research focusing on HVC activity across context and age will be useful for confirming this hypothesis. The comparative approach we employed for chapter three can also provide evidence for neural mechanisms that underlie behavioural differences between species. For example, gap durations decreased across age in Bengalese finches, and the age-dependent changes were significantly predicted by context-dependent changes, whereas we observed no evidence of age-dependent change or of predictability of gap durations in zebra finches. HVC neurons demonstrate significant context-dependent changes to the expression of the immediate early gene egr-1 in Bengalese finches, but not zebra finches, which suggests that differences in HVC activity could be underlie the difference in behaviour related to gap durations (Matheson et al. 2016; Murphy et al. 2017). Continued research on species comparisons in neural activity will be of great use for understanding the fundamental processes of vocal motor modulation and plasticity.

In chapters four and five, we found biases that can predict learning outcomes in zebra finches and began to uncover possible mechanisms that underlie such biases. Specifically, we found that birds with impoverished auditory experiences still demonstrated many of the common sequence patterns that were observed in birds that were experimentally tutored with random sequences. Therefore, we proposed that the motor system contributes to these common patterns (e.g., longer song elements at the end of song). However, it is also possible that the auditory system contributes to these patterns, and to test for this we propose an experiment in juvenile songbirds that have not been exposed to song. We would present these birds with many example songs, and record neuronal activity in HVC or the caudomedial nidopallium (NCM), a higher order auditory area that has been implicated in song learning and preference (Chen et al. 2017; Mello et al. 2004; Yanagihara and Yazaki-Sugiyama 2016). We expect that neurons will respond differentially to songs that either follow the common patterns we observed in chapters four and five compared to songs that follow opposite patterns, which would suggest that the auditory system contributes to the learning biases as well.

Conclusion

Overall, this thesis provides integrative and fundamental data on links between short- and long-term changes to communication. We have documented relationships between short-term social modulation of behaviour and long-term plasticity of behaviour for multiple aspects of vocal communication in two distinct species of songbird. Furthermore, we have experimentally uncovered learning biases within zebra finches that guide sequence learning, and have provided evidence that motor biases contribute this process. Finally, this thesis can serve as the basis for further research on mechanisms that drive behavioral plasticity, development, and control.

References

Adams JA. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychol Bull* 101: 41–74, 1987.

Bickel B. Distributional biases in language families. In: *Language typology and historical contingency*. 2013, 415–444.

Botero CA, **Rossman RJ**, **Caro LM**, **Stenzler LM**, **Lovette IJ**, **de Kort SR**, **Vehrencamp SL**. Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Anim Behav* 77: 701–706, 2009.

Brainard MS, **Doupe AJ**. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Chen Y, **Clark O**, **Woolley SC**. Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proc R Soc B Biol Sci* 284: 20170054, 2017.

Chomsky N. Rules and representations. *Behav Brain Sci* 3: 1, 1980.

Doupe AJ, **Kuhl PK**. Birdsong and humman speech: Common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Dunning JL, **Pant S**, **Bass A**, **Coburn Z**, **Prather JF**. Mate choice in adult female Bengalese finches: Females express consistent preferences for individual males and prefer female-directed song performances. *PLoS One* 9: e89438, 2014.

Ericsson KA, **Krampe RT**, **Tesch-Römer C**. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100: 363–406, 1993.

Fee MS, **Scharff C**. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J* 51: 362–377, 2010.

Fougeron C, **Keating PA**. Articulatory strengthening at edges of prosodic domains. *J Acoust Soc Am* 101: 3728–3740, 1997.

Gerosa M, **Giuliani D**, **Brugnara F**. Acoustic variability and automatic recognition of children's speech. *Speech Commun* 49: 847–860, 2007.

Goedemans R, **van der Hulst H**. Rhythm Types (2013). In: *The world atlas of language structures online*, edited by Dryer MS, Haspelmath M. Max Planck Institute for Evolutionary Anthropology: http://wals.info/chapter/17.

Hahnloser RHR, **Kozhevnikov AA**, **Fee MS**. An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature* 419: 65–70, 2002.

Hampton CM, **Sakata JT**, **Brainard MS**. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J Neurophysiol* 101: 3235–3245, 2009.

Heinig A, **Pant S**, **Dunning JL**, **Bass A**, **Coburn Z**, **Prather JF**. Male mate preferences in mutual mate choice: finches modulate their songs across and within male-female interactions. *Anim Behav* 97: 1–12, 2014.

Kozhevnikov AA, **Fee MS**. Singing-related activity of identified HVC neurons in the zebra finch. *J Neurophysiol* 97: 4271–4283, 2007.

Magill RA. *Motor learning and control: concepts and applications* (7th ed.). New York: McGraw-Hill, 2004.

Matheson LE, **Sun H**, **Sakata JT**. Forebrain circuits underlying the social modulation of vocal communication signals. *Dev Neurobiol* 76: 47–63, 2016.

Mello CV, **Velho TAF**, **Pinaud R**. Song-induced gene expression: A window on song auditory processing and perception. *Ann N Y Acad Sci* 1016: 263–281, 2004.

Murphy K, **James LS**, **Sakata JT**, **Prather JF**. Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. *J Neurophysiol* 118: 800–816, 2017.

Oller DK. The effect of position in utterance on speech segment duration in English. *J Acoust Soc Am* 54: 1235–1247, 1973.

Ota N, **Soma M**. Age-dependent song changes in a closed-ended vocal learner: elevation of song performance after song crystallization. *J Avian Biol* 45: 566–573, 2014.

Patel A. *Music, language, and the brain*. Oxford, UK: Oxford University Press, 2010.

Pytte CL, **Gerson M**, **Miller J**, **Kirn JR**. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Rivera-Gutierrez HF, **Pinxten R**, **Eens M**. Tuning and fading voices in songbirds: age-dependent changes in two acoustic traits across the life span. *Anim Behav* 83: 1279–1283, 2012.

Sakata JT, **Hampton CM**, **Brainard MS**. Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99: 1700–11, 2008.

Sankoff G. Language change across the lifespan. *Annu Rev Linguist* 4: 297–316, 2018.

Savage PE, **Brown S**, **Sakai E**, **Currie TE**. Statistical universals reveal the structures and functions of human music. *Proc Natl Acad Sci* 112: 8987–8992, 2015.

Smith BL, **Kenney MK**, **Hussain S**. A longitudinal investigation of duration and temporal variability in children's speech production. *J Acoust Soc Am* 99: 2344–2349, 1996.

Tierney AT, **Russo FA**, **Patel AD**. The motor origins of human and avian song structure. *Proc Natl Acad Sci* 108: 15510–15515, 2011.

Vaissière J. Language-independent prosodic features. In: *Prosody: Models and Measurements*, 1983, 53–66.

Yanagihara S, **Yazaki-Sugiyama Y**. Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. *Nat Commun* 7: 11946, 2016.

Zhang YS, **Wittenbach JD**, **Jin DZ**, **Kozhevnikov AA**. Temperature manipulation in songbird brain implicates the premotor nucleus HVC in birdsong syntax. *J Neurosci* 37: 2600–2611, 2017.

Full reference list

Aamodt SM, Kozlowski MR, Nordeen EJ, Nordeen KW. Distribution and developmental change in [3H]MK-801 binding within zebra finch song nuclei. *J Neurobiol* 23: 997–1005, 1992.

Abbott LF, Nelson SB. Synaptic plasticity: taming the beast. *Nature Neurosci* 3 Suppl: 1178–1183, 2000.

Achiro JM, Bottjer SW. Neural representation of a target auditory memory in a cortico-basal ganglia pathway. *J Neurosci* 33: 14475–14488, 2013.

Adams JA. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin* 101: 41–71, 1987.

Adret P. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim Behav* 46: 149–159, 1993.

Alcock J. *Animal Behavior*. 9th ed. Sunderland, MA: Sinauer Associates, Inc, 2009.

Ali F, **Otchy TM**, **Pehlevan C**, **Fantana AL**, **Burak Y**, **Ölveczky BP**. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80: 494–506, 2013.

Alward BA, Balthazart J, Ball GF. Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate*. Proc. Natl. Acad. Sci.* 110: 19573–19578, 2013.

Andalman AS, **Fee MS**. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci* 106: 12518–12523, 2009.

Andalman AS, Foerster JN, Fee MS. Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature. *PLoS One* 6: e25461, 2011.

Araki M, **Bandi MM**, **Yazaki-Sugiyama Y**. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–1287, 2016.

Arnold AP. The effects of castration on song development in zebra finches (*Poephila guttata*). *J. Exp. Zool.* 191: 261–277, 1975.

Aronov D, Fee MS. Natural changes in brain temperature underlie variations in song tempo during a mating behavior. *PLoS One* 7: 1–10, 2012.

Aronov D, Veit L, Goldberg JH, Fee MS. Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. *J Neurosci* 45: 16353–316368, 2011.

Ashmore RC, Wild JM, Schmidt MF. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 25: 8543–8554, 2005.

Balmer TS, Carels VM, Frisch JL, Nick TA. Modulation of perineuronal nets and parvalbumin with developmental song learning. *J Neurosci* 29:12878–12885, 2009.

Baptista LF, **Petrinovich L**. Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Anim Behav* 32: 172–181, 1984.

Basista MJ, Elliott KC, Wu W, Hyson RL, Bertram R, Johnson F. Independent premotor encoding of the sequence and structure of birdsong in avian cortex. *J Neurosci* 34: 16821–16834, 2014.

Bates D, **Maechler M**, **Bolker B**, **Walker S**. Package "lme4." *Convergence* 12, 2015.

Beecher MD, Brenowitz EA. Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* 20: 143–149, 2005.

Bickel B. Distributional biases in language families. In: *Language typology and historical contingency*. 2013, 415–444.

Bion RAH, **Benavides-Varela S**, **Nespor M**. Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Lang Speech* 54: 123–140, 2011.

Bolhuis JJ, **Johnson MH**, **Horn G**. Effects of early experience on the development of filial preferences in the domestic chick. *Dev Psychobiol* 18: 299–308, 1985.

Bolhuis JJ, Okanoya K, Scharff C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11: 747–759, 2010.

Bolhuis JJ, **Trooster WJ**. Reversibility revisited: Stimulus-dependent stability of filial preference in the chick. *Anim Behav* 36: 668–674, 1988.

Botero CA, **Rossman RJ**, **Caro LM**, **Stenzler LM**, **Lovette IJ**, **de Kort SR**, **Vehrencamp SL**. Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Anim Behav* 77: 701–706, 2009.

Bottjer SW, Altenau B. Parallel pathways for vocal learning in basal ganglia of songbirds. *Nat Neurosci* 13: 153–155, 2010.

Bouchard KE, Brainard MS. Neural encoding and integration of learned probabilistic sequences in avian sensory-motor circuitry. *J Neurosci* 33: 17710–17723, 2013.

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

Braaten RF, **Reynolds K**. Auditory preference for conspecific song in isolation-reared zebra finches. *Anim Behav* 58: 105–111, 1999.

Brainard MS, Doupe AJ. Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci* 1: 31–40, 2000.

Brainard MS, Doupe AJ. Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J Neurosci* 21: 2501–2517, 2001.

Brainard MS, **Doupe AJ**. What songbirds teach us about learning. *Nature* 417: 351–358, 2002.

Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci* 36:489–517, 2013.

Brainard MS, **Knudsen EI**. Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J Neurosci* 18: 3929–3942, 1998.

Bregman A. *Auditory scene analysis: The perceptual organization of sound*. 1994.

Brenowitz EA, **Beecher MD**. Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends Neurosci* 28: 127–132, 2005.

Briellmann RS, Saling MM, Connell AB, Waites AB, Abbott DF, Jackson GD. A high-field functional MRI study of quadri-lingual subjects. *Brain Lang* 89: 531–542, 2004.

Bruno JH, Tchernichovski O. Regularities in zebra finch song beyond the repeated motif. *Behav. Processes in press*, 2017.

Byers J, Hebets E, Podos J. Female mate choice based upon male motor performance. *Animal Behaviour* 79: 771–778, 2010.

ten Cate C. Behaviour-contingent exposure to taped song and zebra finch song learning. *Anim Behav* 42: 857–859, 1991.

ten Cate C, **Spierings M**. Rules, rhythm and grouping: auditory pattern perception by birds. *Anim Behav* 1–9, 2018.

Catchpole CK, Slater PJB. *Bird song biological themes and variations, 2nd edn*. Cambridge: Cambridge University Press, 2008.

Charlesworth JD, Tumer EC, Warren TL, Brainard MS. Learning the microstructure of successful behavior. *Nat Neurosci* 14: 373–380, 2011.

Charlesworth JD, Warren TL, Brainard MS. Covert skill learning in a cortical-basal ganglia circuit. *Nature* 486: 251–255, 2012.

Chen JL, Rae C, Watkins KE. Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage* 59: 1200–1208, 2012.

Chen Y, **Clark O**, **Woolley SC**. Courtship song preferences in female zebra finches are shaped by developmental auditory experience. *Proc R Soc B Biol Sci* 284: 20170054, 2017.

Chen Y, **Matheson LE**, **Sakata JT**. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc Natl Acad Sci* 113: 6641–6646, 2016.

de Cheveigné A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. *J Acoust Soc Am* 111: 1917–1930, 2002.

Chi Z, Margoliash D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32: 899–910, 2001.

Chomsky N. Rules and representations. *Behav Brain Sci* 3: 1, 1980.

Clayton NS. Song tutor choice in zebra finches and Bengalese finches: the relative importance of visual and vocal cues. *Behaviour* 104: 281–299, 1988.

Clayton NS. The effects of cross-fostering on selective song learning in Estrildid finches. *Behaviour* 109: 163–174, 1989.

Colonnese MT, Stallman EL, Berridge KC. Ontogeny of action syntax in altricial and precocial rodents: Grooming sequences of rat and guinea pig pups. *Behaviour* 133: 1165–1195, 1996.

Cooper BG, Goller F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J Neurophysiol* 95: 3798–3809, 2006.

Cooper BG, **Méndez JM**, **Saar S**, **Whetstone AG**, **Meyers R**, **Goller F**. Age-related changes in the Bengalese finch song motor program. *Neurobiol Aging* 33: 564–568, 2012.

Costa RM. A selectionist account of de novo action learning. *Curr Opin Neurobiol* 21: 579–586, 2011.

Creutzfeldt OD, Heggelund P. Neural plasticity in visual cortex of adult cats after exposure to visual patterns. *Science* 188: 1025–1027, 1975.

Dalla C, Shors TJ. Sex differences in learning processes of classical and operant conditioning. *Physiol. Behav.* 97: 229–238, 2009.

Dan Y, Poo MM. Spike timing-dependent plasticity of neural circuits. *Neuron* 44: 23–30, 2004.

Davids K, Bennett S, Newell K. *Movement system variability*. Champaign, IL: Human Kinetics. 2006.

Dayan E, Cohen LG. Neuroplasticity subserving motor skill learning. *Neuron* 72: 443–454, 2011.

Deregnaucourt S, Gahr M. Horizontal transmission of the father's song in the zebra finch (*Taeniopygia guttata*). *Biol Lett* 9: 20130247, 2013.

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O. How sleep affects the developmental learning of bird song. *Nature* 433: 710–716, 2005.

Derégnaucourt S, Poirier C, Van der Kant A, Van der Linden A, Gahr M. Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* 107: 210–218, 2013.

Dhawale AK, Smith MA, Ölveczky BP. The role of variability in motor learning. *Annu. Rev. Neurosci.* 40: 479–498, 2017.

Dooling R, **Searcy M**. Early perceptual selectivity in the swamp sparrow. *Dev Psychobiol* 13: 499–506, 1980.

Doupe AJ, **Kuhl PK**. Birdsong and humman speech: Common themes and mechanisms. *Annu Rev Neurosci* 22: 567–631, 1999.

Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. *Trends Neurosci* 28: 353–363, 2005.

Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr Opin Neurobiol* 10: 732–739, 2000.

Drake C, **Palmer C**. Skill acquisition in music performance: relations between planning and temporal control. *Cognition* 74: 1–32, 2000.

Dunn AM, Zann RA. Undirected song in wild zebra finch flocks: contexts and effects of mate removal. *Ethology* 102: 529–539, 2010.

Dunning JL, Pant S, Bass A, Coburn Z, Prather JF. Mate choice in adult female Bengalese finches: females express consistent preferences for individual males and prefer female-directed song performances. *PLoS One* 9: e89438, 2014.

Düring DN, **Elemans CPH**. Embodied motor control of avian vocal production. In: *Vertebrate sound production and acoustic communication*. Springer, Cham, 2016, 119–157.

Eales LA. Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim. Behav.* 33: 1293–1300, 1985.

Eckert P, **Rickford J**. *Style and sociolinguistic variation*. Cambridge University Press, 2001.

Ericsson KA. Deliberate practice and acquisition of expert performance: A general overview. *Acad Emerg Med* 15: 988–994, 2008.

Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Pyschol Rev* 100: 363–406, 1993.

Ericsson KA, Lehmann AC. Expert and exceptional performance: evidence of maximal adaption to task constraints. *Annu Rev Psychol*, 47: 363–406, 1996.

Erzurumlu RS, Gaspar P. Development and critical period plasticity of the barrel cortex. *Eur J Neurosci* 35: 1540–1553, 2012.

Fee MS, **Goldberg JH**. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* 198: 152–170, 2011.

Fee MS, Scharff C. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J* 51: 362–377, 2010.

Fehér O, **Ljubičić I**, **Suzuki K**, **Okanoya K**, **Tchernichovski O**. Statistical learning in songbirds: from self-tutoring to song culture. *Phil Trans R Soc B*, 2016.

Fehér O, **Wang H**, **Saar S**, **Mitra PP**, **Tchernichovski O**. De novo establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–568, 2009.

Fentress JC. Emergence of pattern in the development of mammalian movement sequences. *J Neurobiol* 23: 1529–1556, 1992.

Ferdinand V, **Kirby S**, **Smith K**. The cognitive roots of regularization in language. *Cognition* 184: 53–68, 2019.

Fernald A, **Taeschner T**, **Dunn J**, **Papousek M**, **de Boysson-Bardies B**, **Fukui I**. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J Child Lang* 16: 477, 1989.

Fougeron C, Keating PA. Articulatory strengthening at edges of prosodic domains. *J Acoust Soc Am* 101: 3728–3740, 1997.

Fox J, **Friendly M**, **Weisberg S**. Hypothesis tests for multivariate linear models using the car Package. *R J* 5: 39, 2013.

Fujimoto H, Hasegawa T, Watanabe D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* 31: 10023–10033, 2011.

Funabiki Y, Funabiki K. Song retuning with tutor model by adult zebra finches. *Dev. Neurobiol.* 68: 645–655, 2008.

Galef BG, **Laland KN**. Social learning in animals: empirical studies and theoretical models. *Bioscience* 55: 489, 2005.

Gardner TJ, **Naef F**, **Nottebohm F**. Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308: 1046–1049, 2005.

Gentner TQ, Hulse SH. Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Anim Behav* 56: 579–594, 1998.

Gerosa M, **Giuliani D**, **Brugnara F**. Acoustic variability and automatic recognition of children's speech. *Speech Commun* 49: 847–860, 2007.

Gil D, Slater PJB. Song organization and singing patterns of the willow warbler, *Phylloscopus trochilus. Behaviour* 137: 759–782, 2000.

Di Giorgio E, **Loveland JL**, **Mayer U**, **Rosa-Salva O**, **Versace E**, **Vallortigara G**. Filial responses as predisposed and learned preferences: Early attachment in chicks and babies. *Behav Brain Res* 325: 90–104, 2017.

Glaze CM, Troyer TW. Temporal structure in zebra finch song: implications for motor coding. *J Neurosci* 26: 991–1005, 2006.

Glaze CM, Troyer TW. Development of temporal structure in zebra finch song. *J Neurophysiol* 109: 1025–1035, 2013.

Gobes SM, Bolhuis JJ. Birdsong memory: a neural dissociation between song recognition and production. *Curr Biol* 17: 789–793, 2007.

Goedemans R, **van der Hulst H**. Rhythm Types (2013). In: *The world atlas of language structures online*, edited by Dryer MS, Haspelmath M. Max Planck Institute for Evolutionary Anthropology: http://wals.info/chapter/17.

Goldberg JH, Fee MS. Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. *J Neurophysiol* 105: 2729–2739, 2011.

Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C. Brain structure predicts the learning of foreign speech sounds. *Cereb Cortex* 17: 575–582, 2007.

Golestani N, Paus T, Zatorre RJ. Anatomical correlates of learning novel speech sounds. *Neuron* 35: 997–1010, 2002.

Goller F, Cooper BG. Peripheral motor dynamics of song production in the zebra finch. *Ann N Y Acad Sci* 1016: 130–152, 2004.

Gomez DM, **Berent I**, **Benavides-Varela S**, **Bion RAH**, **Cattarossi L**, **Nespor M**, **Mehler J**. Language universals at birth. *Proc Natl Acad Sci* 111: 5837–5841, 2014.

Grafton ST, Hazeltine E, Ivry RB. Motor sequence learning with the nondominant left hand. *Exp Brain Res* 146: 369–378, 2002.

Greenberg J. Some universals of grammar with particular reference to the order of meaningful elements. *Universals Lang* 2: 73–113, 1963.

ter Haar SM, Kaemper W, Stam K, Levelt CC, ten Cate C. The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*). *Proc R Soc B Biol Sci* 281: 20141860, 2014.

Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C. FoxP2 expression in avian vocal learners and non-learners. *J Neurosci* 24: 3164–3175, 2004.

Hahnloser RH, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419: 65–70, 2002.

Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci* 32: 11671–11687, 2012.

Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J Neurophysiol* 101: 3235–3245, 2009.

Hanuschkin A, Ganguli S, Hahnloser RH. A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models. *Front Neural Circuits* 7: 106, 2013.

Hebb DO. The organization of behavior: a neuropsychological theory. New York: Wiley, 1949.

Heinig A, Pant S, Dunning J, Bass A, Coburn Z, Prather JF. Male mate preferences in mutual mate choice: finches modulate their songs across and within male-female interactions. *Anim Behav* 97: 1–12, 2014.

Hensch TK. Critical period plasticity in local cortical circuits. *Nat Rev Neurosci* 6: 877–888, 2005.

Herholz SC, Zatorre RJ. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76: 486–502, 2012.

Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci* 8: 393–402, 2007.

Hisey E, **Kearney MG**, **Mooney R**. A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat Neurosci* 21: 589–597, 2018.

Horita H, Wada K, Jarvis ED. Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *Eur. J. Neurosci.* 28: 2519–2532, 2008.

Horn G, **McCabe BJ**. Predispositions and preferences. Effects on imprinting of lesions to the chick brain. *Anim Behav* 32: 288–292, 1984.

Hothorn T, **Bretz F**, **Westfall P**. Simultaneous inference in general parametric models. *Biometrical J* 50: 346–363, 2008.

Houde JF, **Jordan MI**. Sensorimotor adaptation in speech production. *Science* 279: 1213–1216, 1998.

Hund-Georgiadis M, von Cramon DY. Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res* 125: 417–425, 1999.

Immelmann K. Song development in the zebra finch and other estrilid finches. In: *Bird Vocalizations*, edited by Hinde RA. New York: Cambridge Univ. Press, 1969, 61–74.

James LS, Dai JB, Sakata JT. Ability to modulate birdsong across social contexts develops without imitative social learning. *Biol. Lett.* 14, 2018.

James LS, Sakata JT. Learning biases underlie “universals” in avian vocal sequencing. *Curr. Biol.* 27: 3676–3682.e4, 2017.

James LS, Sakata JT. Vocal motor changes beyond the sensitive period for song plasticity. *J. Neurophysiol.* 112: 2040–2052, 2014.

James LS, Sakata JT. Predicting plasticity: acute context-dependent changes to vocal performance predict long-term age-dependent changes. *J. Neurophysiol.* 114: 2328–2339, 2015.

Jäncke L, Shah NJ, Peters M. Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. Brain *Res Cogn Brain Res* 10: 177–183, 2000.

Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F. For whom the bird sings: context-dependent gene expression. *Neuron* 21: 775–788, 1998.

Jin DZ. Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Phys Rev E Stat Nonlin Soft Matter Phys* 80: 051902, 2009.

Jin H, Clayton DF. Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. *Neuron* 19: 1049–1059, 1997.

Johnson F, Soderstrom K, Whitney O. Quantifying song bout production during zebra finch sensory-motor learning suggests a sensitive period for vocal practice. *Behav Brain Res* 131: 57–65, 2002.

Jones AE, ten Cate C, Slater PJB. Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* 110: 354–369, 1996.

Jun JK, Jin DZ. Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. *PLoS ONE* 2: e723, 2007.

Jusczyk PW, **Houston DM**, **Newsome M**. The beginnings of word segmentation in English-learning infants. *Cogn Psychol* 39: 159–207, 1999.

Kakishita Y, Sasahara K, Nishino T, Takahasi M, Okanoya K. Ethological data mining: an automata-based approach to extract behavioral units and rules. *Data Mining and Knowledge Discovery* 18: 446–471, 2009.

van der Kant A, Derégnaucourt S, Gahr M, Van der Linden A, Poirier C. Representation of early sensory experience in the adult auditory midbrain: implications for vocal learning. *PLoS One* 8: e61764, 2013.

Kao MH, **Brainard MS**. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96: 1441–1455, 2006.

Kao MH, Doupe AJ, Brainard MS. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433: 638–643, 2005.

Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377: 155–158, 1995.

Karni A, Meyer G, Rey-Hipolito C, Jezzard P, Adams MM, Turner R, Ungerleider LG. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci USA* 95: 861–868, 1998.

Katahira K, Suzuki K, Okanoya K, Okada M. Complex sequencing rules of birdsong can be explained by simple hidden Markov processes. *PLoS One* 6: e24516, 2011.

Kato M, Okanoya K. Molecular characterization of the song control nucleus HVC in Bengalese finch brain. *Brain Res* 1360: 56–76, 2010.

Keith N, **Ericsson KA**. A deliberate practice account of typing proficiency in everyday typists. *J Exp Psychol Appl* 13: 135–145, 2007.

Kelly AM, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 15: 1089–1102, 2005.

Kelly CW, Sober SJ. A simple computational principle predicts vocal adaptation dynamics across age and error size. *Front Integr Neurosci* 8: 75, 2014.

Kirby S. Culture and biology in the origins of linguistic structure. *Psychon Bull Rev* 24: 118–137, 2017.

Kirby S, **Cornish H**, **Smith K**. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proc Natl Acad Sci* 105: 10681–10686, 2008.

Knudsen EI. Sensitive periods in the development of the brain and behavior. *J Cogn Neurosci* 16: 1412–1425, 2004.

Kobayashi K, Uno H, Okanoya K. Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. *Neuroreport* 12: 353–358, 2001.

Kolb B, **Gibb R**, **Robinson TE**. Brain plasticity and behavior. *Curr Dir Psychol Sci* 12: 1–5, 2003.

Kojima S, Doupe AJ. Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J Neurophysiol* 98: 2099–2109, 2007.

Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci USA* 108: 1687–1692, 2011.

Konczak J, van der Velden H, Jaeger L. Learning to play the violin: motor control by freezing, not freeing degrees of freedom. *J Motor Behav* 41: 243–252, 2009.

Konishi M. Birdsong: from behavior to neuron. *Annu Rev Neurosci* 8: 125–170, 1985.

Konishi M. The role of auditory feedback in birdsong. *Ann N Y Acad Sci* 1016: 463–475, 2004.

de Kort SR, **Eldermire ERB**, **Valderrama S**, **Botero CA**, **Vehrencamp SL.** Trill consistency is an age-related assessment signal in banded wrens. *Proc R Soc B Biol Sci* 276: 2315–2321, 2009.

Kosche G, Vallentin D, Long MA. Interplay of inhibition and excitation shapes a premotor neural sequence. *J Neurosci* 35: 1217–1227, 2015.

Kosten TA, Kim JJ, Lee HJ. Early life manipulations alter learning and memory in rats. *Neurosci. Biobehav. Rev.* 36: 1985–2006, 2012.

Kozhevnikov AA, **Fee MS**. Singing-related activity of identified HVC neurons in the zebra finch. *J Neurophysiol* 97: 4271–4283, 2007.

Krings T, Töpper R, Foltys H, Erberich S, Sparing R, Willmes K, Thron A. Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett* 278: 189–193, 2000.

Kubikova L, Bosikova E, Cvikova M, Lukacova K, Scharff C, Jarvis ED. Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. *Sci Rep* 4: 6590, 2014.

Kuhl PK. Brain mechanisms in early language acquisition. *Neuron* 67: 713–727, 2010.

Labov W. *Sociolinguistic patterns*. University of Pennsylvania Press, 1972.

Lachlan RF, **van Heijningen CAA**, **ter Haar SM**, **ten Cate C.** Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front Psychol* 7: 980, 2016.

Laland K, **Galef B**. *The question of animal culture*. Harvard University Press, 2009.

Landi SM, Baguear F, Della-Maggiore V. One week of motor adaptation induces structural changes in primary motor cortex that predict long-term memory one year later. *J*

Neurosci 31: 11808–11813, 2011.

Lane H, **Webster JW**. Speech deterioration in postlingually deafened adults. *J Acoust Soc Am* 89: 859–866, 1991.

Langus A, Mehler J, Nespor M. Rhythm in language acquisition. *Neurosci Biobehav Rev*, 2016.

Larsen ON, **Goller F**. Role of syringeal vibrations in bird vocalizations. *Proc R Soc London Ser B Biol Sci* 266: 1609–1615, 1999.

Larsen ON, **Goller F**. Direct observation of syringeal muscle function in songbirds and a parrot. *J Exp Biol* 205: 25–35, 2002.

Leblois A, Perkel DJ. Striatal dopamine modulates song spectral but not temporal features through D1 receptors. *Eur J Neurosci* 35: 1771–1781, 2012.

Leblois A, Wendel BJ, Perkel DJ. Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *J Neurosci* 30: 5730–5743, 2010.

Liaw A, **Wiener M**. Classification and Regression by randomForest. *R news* 2: 18–22, 2002.

Liberman AM, **Cooper FS**, **Shankweiler DP**, **Studdert-Kennedy M**. Perception of the speech code. *Psychol Rev* 74: 431–461, 1967.

Liberman AM, Prince A. On stress and linguistic rhythm. *Linguist Inq* 8: 249–336, 1977.

Liberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition* 21: 1–36, 1985.

Lipkind D, **Marcus GF**, **Bemis DK**, **Sasahara K**, **Jacoby N**, **Takahasi M**, **Suzuki K**, **Feher O**, **Ravbar P**, **Okanoya K**, **Tchernichovski O**. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–108, 2013.

Lipkind D, Zai AT, Hanuschkin A, Marcus GF, Tchernichovski O, Hahnloser RHR. Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nat. Commun.* 8: 1247, 2017.

Liu WC, **Wada K**, **Nottebohm F**. Variable food begging calls are harbingers of vocal learning. *PLoS One* 4, 2009.

Livingston FS, White SA, Mooney R. Slow NMDA-EPSCs at synapses critical for song development are not required for song learning in zebra finches. *Nat Neurosci* 3: 482–488, 2000.

Lombardino AJ, Nottebohm F. Age at deafening affects the stability of learned song in adult male zebra finches. *J Neurosci* 20: 5054–5064, 2000.

Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456: 189–194, 2008.

Long MA, Jin DZ, Fee MS. Support for a synaptic chain model of neuronal sequence generation. *Nature* 468: 394–399, 2010.

Lotze M, Scheler G, Tan H-RM, Braun C, Birbaumer N. The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20: 1817–1829, 2003.

Lumaca M, **Baggio G**. Cultural transmission and evolution of melodic structures in multi-generational signaling games. *Artif Life* 23: 406–423, 2017.

MacNeilage PF. The frame/content theory of evolution of speech production. *Behav. Brain Sci.* 21: 499–546, 1998.

MacNeilage PF, **Davis BL**, **Kinney A**, **Matyear CL**. The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Dev* 71: 153–163, 2000.

Madison G, Karampela O, Ullén F, Holm L. Effects of practice on variability in an isochronous serial interval production task: Asymptotical levels of tapping variability after training are similar to those of musicians. *Acta Psychologica* 143: 119–128, 2013.

Magill RA. *Motor learning and control: concepts and applications* (7th ed.). New York: McGraw-Hill, 2004.

Makino H, Hwang EJ, Hedrick NG, Komiyama T. Circuit mechanisms of sensorimotor learning. *Neuron* 92: 705–721, 2016.

Manly B. *Randomization, bootstrap and Monte Carlo methods in biology.* London, UK: Chapman & Hall, 1997.

Marler P. A comparative approach to vocal learning: Song development in white-crowned sparrows. *J Comp Physiol Psychol* 71: 1–25, 1970a.

Marler P. Birdsong and human speech: could there be parallels? *Am Sci* 58: 669–674, 1970b.

Marler P. Three models of song learning: evidence from behavior. *J Neurobiol* 33: 501–516, 1997.

Marler P. Bird calls: Their potential for behavioral neurobiology. In: *Annals of the New York Academy of Sciences,* 31–44, 2004.

Marler P, **Peters S**. Selective vocal learning in a sparrow. *Science* 198: 519–521, 1977.

Marler P, **Sherman V**. Song structure without auditory feedback: emendations of the auditory template hypothesis. *J Neurosci* 3: 517–531, 1983.

Marler P, **Sherman V**. Innate differences in singing behaviour of sparrows reared in isolation from adult conspecific song. *Anim Behav* 33: 57–71, 1985.

Marler P, **Tamura M**. Song "dialects" in three populations of white-crowned sparrows. *Condor* 64: 368–377, 1962.

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

Matheson LE, Sakata JT. Catecholaminergic contributions to vocal communication signals. *Eur J Neurosci* 14: 1180–1194, 2015.

Matheson LE, Sun H, Sakata JT. Forebrain circuits underlying the social modulation of vocal communication signals. *Dev Neurobiol* 76, 47–63, 2016.

Mehr SA, **Singh M**, **York H**, **Glowacki L**, **Krasnow MM**. Form and function in human song. *Curr Biol* 28: 356-368.e5, 2018.

Mello CV, **Velho TAF**, **Pinaud R**. Song-induced gene expression: A window on song auditory processing and perception. *Ann N Y Acad Sci* 1016: 263–281, 2004.

Mets DG, **Brainard MS**. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proc Natl Acad Sci* 115: 421–426, 2018.

Mol C, Chen A, Kager RWJ, ter Haar SM. Prosody in birdsong: A review and perspective. *Neurosci Biobehav Rev*, 2017.

Mooney R. Sensitive periods and circuits for learned birdsong. *Curr Opin Neurobiol* 9: 121–127, 1999.

Mooney R. Neural mechanisms for learned birdsong. *Learn Mem* 16: 655–669, 2009.

Mooney R, Prather JF. The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *J Neurosci* 8: 1952–1964, 2005.

Mori C, **Liu W**, **Wada K**. Recurrent development of song idiosyncrasy without auditory inputs in the canary, an open-ended vocal learner. *Sci Rep* 8: 8732, 2018.

Mori C, **Wada K**. Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *J Neurosci* 35: 878–889, 2015.

Morris D. The reproductive behaviour of the zebra finch, with special reference to pseudofemale behaviour and displacement activities. *Behaviour* 6: 271–322, 1954.

Morrison RG, Nottebohm F. Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *J Neurobiol* 24: 1045–1064, 1993.

Müller H, Sternad D. Motor learning: changes in the structure of variability in a redundant task In: *Progress in Motor Control*, Springer US, 2009, 439–456.

Murphy K, James LS, Sakata JT, Prather JF. Advantages of comparative studies in songbirds to understand the neural basis of sensorimotor integration. *J. Neurophysiol.* 118: 800–816, 2017.

Nelson DA, **Marler P**. Innate recognition of song in white-crowned sparrows: a role in selective vocal learning? *Anim Behav* 46: 806–808, 1993.

Newport E. Maturational constraints on language learning. *Cogn Sci* 14: 11–28, 1990.

Nick TA, Konishi M. Neural song preference during vocal learning in the zebra finch depends on age and state. *J Neurobiol* 62: 231–242, 2005.

Nishikawa J, Okanoya K. Dynamical neural representation of song syntax in Bengalese Finch: a model study. *Ornitholog Science* 5: 95–103, 2006.

Nixdorf-Bergweiler BE, Wallhäusser-Franke E, DeVoogd TJ. Regressive development in neuronal structure during song learning in birds. *J Neurobiol* 27: 204–15, 1995.

Nordeen KW, **Nordeen EJ**. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57: 58–66, 1992.

Okanoya K. The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann NY Acad Sci* 1016: 724–735, 2004.

Okanoya K. Song syntax in Bengalese finches: proximate and ultimate analyses, In: *Advances in the Study of Behavior*. New York: Academic Press: 297–346, 2004.

Okanoya K. Behavioural factors governing song complexity in Bengalese finches. *Intl J Comp Psychol* 25: 44–59, 2012.

Okanoya K, Yamaguchi A. Adult Bengalese finches (*Lonchura striata* var. *domestica*) require real time auditory feedback to produce normal song syntax. *J Neurobiol* 33: 343–356, 1997.

Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS. Growth and splitting of neural sequences in songbird vocal development. *Nature* 528: 352–357, 2015.

Oller DK. The effect of position in utterance on speech segment duration in English. *J Acoust Soc Am* 54: 1235–1247, 1973.

Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3: e153, 2005.

Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol* 106:386–397, 2011.

Osmanski MS, Dooling RJ. The effect of altered auditory feedback on control of vocal production in budgerigars (*Melopsittacus undulatus*). *J Acoust Soc Am* 126: 911–919, 2009.

Ota N, **Soma M**. Age-dependent song changes in a closed-ended vocal learner: elevation of song performance after song crystallization. *J Avian Biol* 45: 566–573, 2014.

Patel A. *Music, language, and the brain*. Oxford, UK: Oxford University Press, 2010.

Petersen SE, van Mier H, Fiez JA, Raichle ME. The effects of practice on the functional anatomy of task performance. *Proc Natl Acad Sci USA* 95: 853–860, 1998.

Perani D, Abutalebi J, Paulesu E, Brambati S, Scifo P, Cappa SF, Fazio F. The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum Brain Mapp* 19: 170–182, 2003.

Phan ML, Pytte CL, Vicario DS. Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proc Natl Acad Sci USA* 103: 1088–1093, 2006.

Plamondon SL, Goller F, Rose GJ. Tutor model syntax influences the syntactical and phonological structure of crystallized songs of white-crowned sparrows. *Anim Behav* 76, 1815–1827, 2008.

Plamondon SL, **Rose GJ**, **Goller F**. Roles of syntax information in directing song development in white-crowned sparrows (Zonotrichia leucophrys). *J Comp Psychol* 124: 117–32, 2010.

Podos J. Motor constraints on vocal development in a songbird. *Anim Behav* 51: 1061–1070, 1996.

Podos J, Lahti DC, Moseley DL. Vocal performance and sensorimotor learning in songbirds. *Adv Study Behav* 40: 159–195, 2009.

Poldrack RA. Imaging brain plasticity: conceptual and methodological issues--a theoretical review. *Neuroimage* 12: 1–13, 2000.

Prather JF. Auditory signal processing in communication: perception and performance of vocal sounds. *Hear Res* 305: 144–155, 2013.

Prather J, Okanoya K, Bolhuis JJ. Brains for birds and babies: neural parallels between birdsong and speech acquisition. *Neurosci Biobehav Rev*, 2017.

Prather JF, Peters S, Nowicki S, Mooney R. Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature* 451: 305–310, 2008.

Price PH. Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93: 260–277, 1979.

Piristine HC, **Choetso T**, **Gobes SMH**. A sensorimotor area in the songbird brain is required for production of vocalizations in the song learning period of development. *Dev Neurobiol* 76: 1213–1225, 2016.

Pytte CL, **Gerson M**, **Miller J**, **Kirn JR**. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev Neurobiol* 67: 1699–1720, 2007.

Rajan R, Doupe AJ. Behavioral and neural signatures of readiness to initiate a learned motor sequence. *Curr Biol* 23: 87–93, 2013.

Ravbar P, Lipkind D, Parra LC, Tchernichovski O. Vocal exploration is locally regulated during song learning. *J Neurosci* 32: 3422–3432, 2012.

Ravignani A, **Delgado T**, **Kirby S**. Musical evolution in the lab exhibits rhythmic universals. *Nat Hum Behav* 1: 0007, 2016.

Reader SM. Animal social learning: associations and adaptations. *F1000Research* 5: 2120, 2016.

Reed JL, Kaas JH. Statistical analysis of large-scale neuronal recording data. *Neural Netw* 23: 673–684, 2010.

Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Medina L, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Güntürkün O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, Smulders TV, Wada K, Dugas-Ford J, Husband S, Yamamoto K, Yu J, Siang C, and Jarvis ED. Revised nomenclature for avian telencephalon and some related brainstem nuclei. *J Comp Neurol* 473: 377–414, 2004.

Riebel K. Song and female mate choice in zebra finches: a review. In: *Advances in the Study of Behavior*. Elsevier Inc., 2009, 197–238.

Riede T, **Goller F**. Morphological basis for the evolution of acoustic diversity in oscine songbirds. *Proc R Soc B Biol Sci* 281: 20132306–20132306, 2014.

Riters L V., Eens M, Pinxten R, Duffy DL, Balthazart J, Ball GF. Seasonal changes in courtship song and the medial preoptic area in male European starlings (*Sturnus vulgaris*). *Horm Behav* 38: 250–261, 2000.

Rivera-Gutierrez HF, **Pinxten R**, **Eens M**. Multiple signals for multiple messages: great tit, Parus major, song signals age and survival. *Anim Behav* 80: 451–459, 2010.

Rivera-Gutierrez HF, **Pinxten R**, **Eens M**. Tuning and fading voices in songbirds: age-dependent changes in two acoustic traits across the life span. *Anim Behav* 83: 1279–1283, 2012.

Roberts TF, Hisey E, Tanaka M, Kearney MG, Chattree G, Yang CF, Shah NM, Mooney R. Identification of a motor-to-auditory pathway important for vocal learning. *Nat. Neurosci.* 20: 978–986, 2017.

Roberts TF, Tschida KA, Klein ME, Mooney R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* 463: 948–952, 2010.

Roper A, Zann R. The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* 112: 458–470, 2006.

Sakai K, Hikosaka O, Nakamura K. Emergence of rhythm during motor learning. *Trends Cogn Sci* 8: 547–553, 2004.

Sakata JT, Crews D. Embryonic temperature shapes behavioural change following social experience in male leopard geckos, *Eublepharis macularius*. *Anim Behav* 66: 839–846, 2003.

Sakata JT, Crews D. Developmental sculpting of social phenotype and plasticity. *Neurosci Biobehav Rev* 28: 95–112, 2004.

Sakata JT, **Brainard MS**. Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26: 9619–9628, 2006.

Sakata JT, Brainard MS. Online contributions of auditory feedback to neural activity in avian song control circuitry. *J Neurosci* 28: 11378–11390, 2008.

Sakata JT, Brainard MS. Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *J Neurophysiol* 102: 2485–2497, 2009.

Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99: 1700–1711, 2008.

Sakata JT, Vehrencamp SL. Integrating perspectives on vocal performance and consistency. *J Exp Biol* 215: 201–209, 2012.

Salvante KG, Racke DM, Campbell CR, Sockman KW. Plasticity in singing effort and its relationship with monoamine metabolism in the songbird telencephalon. *Dev Neurobiol* 14, 2009.

Samuels M, Witmer J. *Statistics for the Life Sciences*. 3rd edition. New Jersey: Prentice Hall. 2002.

Sanes JN, Donoghue JP. Plasticity and primary motor cortex. *Annu Rev Neurosci* 23: 393–415, 2000.

Sankoff G. Language change across the lifespan. *Annu Rev Linguist* 4: 297–316, 2018.

Savage PE, Brown S, Sakai E, Currie TE. Statistical universals reveal the structures and functions of human music. *Proc Natl Acad Sci* 112: 8987–8992, 2015.

Schaefer H, **Hess H**. Color preferences in imprinting objects. *Z Tierpsychol* 16: 161–172, 1959.

Scharff C, Nottebohm F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11: 2896–2913, 1991.

Schmidt MF. Pattern of interhemispheric synchronization in HVc during singing correlates with key transitions in the song pattern. *J Neurophysiol* 90: 3931–3949, 2003.

Schmidt MF, **Goller F**. Breathtaking songs: coordinating the neural circuits for breathing and singing. *Physiology* 31: 442–451, 2016.

Schmidt RA, Lee T. *Motor control and learning*. Champaign, IL: Human kinetics, 1998.

Senghas A, **Coppola M**. Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychol Sci* 12: 323–328, 2001.

Shadmehr R, Krakauer JW. A computational neuroanatomy for motor control. *Exp Brain Res* 185: 359–381, 2008.

Shors TJ, Miesegaes G. Testosterone in utero and at birth dictates how stressful experience will affect learning in adulthood. *Proc Natl Acad Sci* 99: 13955–13960, 2002.

Sluckin W, **Salzen EA**. Imprinting and perceptual learning. *Q J Exp Psychol* 13: 65–77, 1961.

Smith BL, **Kenney MK**, **Hussain S**. A longitudinal investigation of duration and temporal variability in children's speech production. *J Acoust Soc Am* 99: 2344–2349, 1996.

Smith GT, Brenowitz EA, Beecher MD, Wingfield JC. Seasonal changes in testosterone, neural attributes of song control nuclei, and song structure in wild songbirds. *J Neurosci* 17: 6001–10, 1997.

Smith K, **Perfors A**, **Fehér O**, **Samara A**, **Swoboda K**, **Wonnacott E**. Language learning, language use and the evolution of linguistic variation. *Philos Trans R Soc B Biol Sci* 372: 20160051, 2017.

Sober SJ, **Brainard MS**. Adult birdsong is actively maintained by error correction. *Nat Neurosci* 12: 927–931, 2009.

Soha JA, **Marler P**. Vocal syntax development in the white-crowned sparrow (*Zonotrichia leucophrys*). *J Comp Psychol* 115: 172–180, 2001.

Solis MM, Doupe AJ. Compromised neural selectivity for song in birds with impaired sensorimotor learning. *Neuron* 25: 109–121, 2000.

Sossinka R, **Böhner J**. Song types in the zebra finch. *Z Tierpsychol* 53: 123–132, 1980.

Spierings MJ, ten Cate C. Zebra finches are sensitive to prosodic features of human speech. *Proc R Soc B Biol Sci* 281: 20140480, 2014.

Spierings M, Hubert J, ten Cate C. Selective auditory grouping by zebra finches: testing the iambic–trochaic law. *Anim Cogn* 20: 665–675, 2017.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104: 2474–2486, 2010.

Sturdy CB, Phillmore LS, Weisman RG. Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 113: 194–203, 1999.

Suthers R, **Goller F**, **Pytte C**. The neuromuscular control of birdsong. *Philos Trans R Soc London Ser B Biol Sci* 354: 927–939, 1999.

Suthers RA, **Goller F**. Motor correlates of vocal diversity in songbirds. In: *Current*

Ornithology. Springer US, 1997, 235–288.

Suthers RA, **Goller F**, **Wild JM.** Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proc Natl Acad Sci* 99: 5680–5685, 2002.

Tabachnick B, **Fidell L**, **Ullman J**. *Using multivariate statistics*. Boston, MA: Pearson, 2007.

Taborsky B, **Oliveira RF**. Social competence: An evolutionary approach. *Trends Ecol Evol* 27: 679–688, 2012.

Tanaka M, Singh Alvarado J, Murugan M, Mooney R. Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. *Proc Natl Acad Sci* 113: E1720–E1727, 2016.

Tchernichovski O, Lints T, Mitra PP, Nottebohm F. Vocal imitation in zebra finches is inversely related to model abundance. *Proc Natl Acad Sci* 96: 12901–12904, 1999.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291: 2564–2569, 2001.

Tchernichovski O, Nottebohm F. Social inhibition of song imitation among sibling male zebra finches. *Proc Natl Acad Sci* 95: 8951–8956, 1998.

Tchernichovski O, Nottebohm F, Ho C, Pesaran B, Mitra P. A procedure for an automated measurement of song similarity. *Anim Behav* 59: 1167–1176, 2000.

Teramitsu I, White SA. FoxP2 regulation during undirected singing in adult songbirds. *J Neurosci* 26: 7390–7394, 2006.

Thomason S, **Kaufman T**. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press, 1988.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F. Dual pre-motor contribution to songbird syllable variation. *J Neurosci* 31: 322–330, 2011.

Tierney AT, **Russo FA**, **Patel AD**. The motor origins of human and avian song structure. *Proc Natl Acad Sci* 108: 15510–15515, 2011.

Tomassini V, Jbabdi S, Kincses ZT, Bosnell R, Douaud G, Pozzilli C, Matthews PM, Johansen-Berg H. Structural and functional bases for individual differences in motor learning. *Hum Brain Mapp* 32: 494–508, 2011.

Tomback DF, **Baker MC**. Assortative mating by white-crowned sparrows at song dialect boundaries. *Anim Behav* 32: 465–469, 1984.

Trainor LJ. Are there critical periods for musical development? *Dev Psychobiol* 46: 262–

278, 2005.

Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J Neurophysiol* 84: 1204–1223, 2000a.

Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning II. Temporal hierarchies and the learning of song sequence. *J Neurophysiol* 3: 1224–1239, 2000b.

Tschida K, Mooney R. The role of auditory feedback in vocal learning and maintenance. *Curr Opin Neurobiol* 22: 320–327, 2012.

Tumer EC, **Brainard MS**. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* 450: 1240–1244, 2007.

Ungerleider LG, Doyon J, Karni A. Imaging brain plasticity during motor skill learning. *Neurobiol Learn Mem* 78: 553–564, 2002.

Vaissière J. Language-independent prosodic features. In *Prosody: Models and Measurements,* 1983, 53–66.

Veit L, Aronov D, Fee MS. Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds. *J Neurophysiol* 106: 1747–1765, 2011.

Vihman MM. *Phonological development: the origins of language in the child. Applied language studies.* Oxford: Blackwell Publishing, 1996.

Vingerhoets G, Van Borsel J, Tesink C, van den Noort M, Deblaere K, Seurinck R, Vandemaele P, Achten E. Multilingualism: an fMRI study. *Neuroimage* 20: 2181–2196, 2003.

Volman SF, Khanna H. Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 109: 211–221, 1995.

Voorhuis TAM, De Kloet ER, De Wied D. Effect of a vasotocin analog on singing behavior in the canary. *Horm Behav* 25: 549–559, 1991.

Vu ET, Mazurek ME, Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14: 6924–6934, 1994.

Wang CZ, Herbst JA, Keller GB, Hahnloser RH. Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biol* 6: e250, 2008.

Warren TL, Tumer EC, Charlesworth JD, Brainard MS. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J Neurophysiol* 106: 1806–1821, 2011.

Warren TL, **Charlesworth JD**, **Tumer EC**, **Brainard MS**. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci* 32: 15414–15425, 2012.

Werker JF, **Hensch TK**. Critical periods in speech perception: new directions. *Annu Rev Psychol* 66: 173–196, 2015.

Werker JF, **Pegg JE**, **McLeod PJ**. A cross-language investigation of infant preference for infant-directed communication. *Infant Behav Dev* 17: 323–333, 1994.

Whaling CS, **Solis MM**, **Doupe AJ**, **Soha JA**, **Marler P**. Acoustic and neural bases for innate recognition of song. *Proc Natl Acad Sci* 94: 12694–12698, 1997.

Wheatcroft D, **Qvarnström A**. A blueprint for vocal learning: auditory predispositions from brains to genomes. *Biol Lett* 11: 20150155, 2015.

Whitney O, Soderstrom K, Johnson F. Post-transcriptional regulation of zenk expression associated with zebra finch vocal development. *Brain Res Mol Brain Res* 80: 279–290, 2000.

Williams H, Mehta N. Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *J Neurobiol* 39: 14–28, 1999.

Willingham DB. A neuropsychological theory of motor skill learning. *Psychol Rev* 105: 558–584, 1998.

Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci* 358: 593–602, 2003.

Woolley SC. Dopaminergic regulation of vocal-motor plasticity and performance. *Curr Opin Neurobiol* 54: 127–133, 2019.

Woolley SC, **Doupe AJ**. Social context-induced song variation affects female behavior and gene expression. *PLoS Biol* 6: 0525–0537, 2008.

Woolley SC, Kao MH. Variability in action: Contributions of a songbird cortical-basal ganglia circuit to vocal motor learning and control. *Neuroscience* 296: 39–47, 2015.

Wohlgemuth MJ, Sober SJ, Brainard MS. Linked control of syllable sequence and phonology in birdsong. *J Neurosci* 30: 12936–12949, 2010.

Woolley SMN, **Rubel EW**. Bengalese finches (*Lonchura striata domestica*) depend upon auditory feedback for the maintenance of adult song. *J Neurosci* 17: 6380–6390, 1997.

Xiao L, **Chattree G**, **Oscos FG**, **Cao M**, **Wanat MJ**, **Roberts TF**. A basal ganglia circuit sufficient to guide birdsong learning. *Neuron* 98: 208-221.e5, 2018.

Yanagihara S, **Yazaki-Sugiyama Y**. Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. *Nat Commun* 7: 11946, 2016.

Yang C, **Crain S**, **Berwick RC**, **Chomsky N**, **Bolhuis JJ**. The growth of language: Universal Grammar, experience, and principles of computation. *Neurosci Biobehav Rev*, 2017.

Yu AC, Margoliash D. Temporal hierarchical control of singing in birds. *Science* 273: 1871–1875, 1996.

Zann RA. *The zebra finch: A synthesis of field and laboratory studies*. Oxford, UK: Oxford University Press, 1996.

Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 8: 547–558, 2007.

Zatorre RJ, Delhommeau K, Zarate JM. Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front Psychol* 3: 544, 2012.

Zevin JD, **Seidenberg MS**, **Bottjer SW**. Limits on reacquisition of song in adult zebra finches exposed to white noise. *J Neurosci* 24: 5849–5862, 2004.

Zhang YS, **Wittenbach JD**, **Jin DZ**, **Kozhevnikov AA**. Temperature manipulation in songbird brain implicates the premotor nucleus HVC in birdsong syntax. *J Neurosci* 37: 2600–2611, 2017.

Zollinger SA, **Suthers RA**. Motor mechanisms of a vocal mimic: implications for birdsong production. *Proc R Soc B Biol Sci* 271: 483–491, 2004.