AN ECOLOGICAL PERSPECTIVE ON SONG LEARNING IN THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*): THE ROLE OF SOCIAL INFLUENCES

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Otilia Menyhart
May, 2013

AN ECOLOGICAL PERSPECTIVE ON SONG LEARNING IN THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*): THE ROLE OF SOCIAL INFLUENCES

Otilia Menyhart, Ph. D.

Cornell University 2013

Communication is a social act, therefore learning to communicate cannot develop in a social vacuum. This thesis investigates vocal development in a songbird species, the Zebra finch, from an ecological perspective to understand the principles by which the naturalistic environment facilitates song learning. I focus my thesis on two such principles: the ways in which song learning is dependent on social affordances of the naturalistic social environment, and the ways in which structural regularities within the communication signal can be exploited to facilitate vocal learning. In *Chapter 2* of the dissertation, I investigate what and how juvenile songbirds learn when provided with opportunities to interact with family members. *Chapter 3* reveals that based on the structural variability in their father's song, juvenile males learn to sing statistically coherent grammatical structures. This is the first description of a new parallel between birdsong and language. In *Chapter 4,* I explore the contribution of social feedback in a strictly controlled vocal environment, by using contingent song playback on juvenile vocalizations to mimic naturalistic interactions. Overall, this thesis provides evidence that there is great potential for social interactions to guide the song learning process in the zebra finch, and song acquisition can only fully understood when it is studied in the context of a communication system.

BIOGRAPHICAL SKETCH

Otilia Menyhart was born in 1980 in the former Czechoslovakia as a member of a Hungarian speaking minority. She graduated in 2004 from the Veterinary School of Hungary (now St. Istvan University) in Budapest, and earned her degree in Applied Zoology, majoring in Conservation Biology. She developed a lifelong passion for songbirds when she conducted her first field studies in elementary school about species diversity at bird feeders. She gained her first research experience as an undergraduate while she was at the Veterinary School. While working with Dr. Tamas Szekely, she studied singing behavior and its relation to the breeding ecology in Penduline Tits in Szeged, Hungary for many summers. In 2005, she and her husband moved to Ithaca and she began her graduate studies by continuing her work on songbirds in the field. The most influential factor in determining her career path was the birth of her older daughter, Vica, which turned her interest toward developmental psychobiology and vocal learning. Under the encouraging guidance of Dr. Timothy DeVoogd, she started to collaborate with Dr. Michael Goldstein, and the successful merging of her interest in songbirds and development resulted in the following dissertation. Being a loyal member of the B.A.B.Y. lab, she produced her second offspring in 2010, who became a valuable contributor to the lab's data base. Now, as a family of four, they take their next steps together and will move to Berlin, where she will start her post-doctoral research position at the Freie Universitat.

To Miki, Vica and Kiki, with love

ACKNOWLEDGMENTS

I have been fantastically fortunate to be surrounded with great people during my graduate life. Without their support, respect, encouragement, and advice it would not have been possible to accomplish the research presented here.

First, and foremost I want to thank my graduate advisor, Timothy DeVoogd, who encouraged me from the very beginning. His confidence in my abilities helped me through the seven-year long process. As a great advisor, he always knew how to provide both professional and personal help in challenging circumstances.

I want to thank Mike Goldstein, from whom I learned an integrationist perspective, creative scientific thinking, braveness to jump into unknown territories and academic rigor, and also what it takes to be an effective mentor for students, and Jen Schwade who provided a fantastic scientific role model. My committee also played a central role in the completion of my dissertation. James Cutting, Elizabeth Adkins-Regan and Sandra Vehrencamp always asked the best questions to steer my research into the right direction.

I am eternally grateful for my graduate friends, whose emotional support was integral to my success and who became my family in Ithaca. I want to express sincere gratitude to the Olsen family, Larry and Janet who became the stand-in grandparents of our children. Without them balancing work and family would have been more difficult.

I also need to thank all the fantastic student and research assistants for their contribution to the success of the projects, and Timothy Van Deusen for his patience and help in the animal facility.

Finally, I want to thank my family: my parents, who provided every opportunity to me to grow and learn, and my sister, who always tried to unload as much burden as possible from everyone’s life. I want to thank particularly my husband, Miklos Beke, whose absolutely selfless support of my academic interest, his patience and positive approach proved to be the strong

foundation of our family, and my two children, Vica and Kiki who became the greatest joy, inspiration, and accomplishment of my life.

TABLE OF CONTENTS

Biographical Sketch iii

Dedication iv

Acknowledgment v

Table of Contents vii

Chapter 1: Introduction 1

Chapter 2: Socially guided song learning in the zebra finch: effects at multiple time scales 13

Chapter 3: Juvenile zebra finch learn the underlying statistical regularities of their fathers' song 66

Chapter 4: Contingent playback upon juvenile vocalizations shapes final song production in the zebra finch 90

Chapter 5: Conclusions 148

CHAPTER 1

INTRODUCTION

The central goal of my thesis is to study song development from an ecological perspective to understand the principles by which the naturalistic environment facilitates vocal ontogeny. I focus my thesis on two such principles: the ways in which song learning is dependent on social affordances of the naturalistic social environment, and the ways in which structural regularities within the communication signal can be exploited to facilitate vocal learning. My aim is to reveal learning mechanisms that could generate testable hypotheses about the means of song acquisition when it is studied in the context of a communication system.

Affordances originally refer to properties of physical objects that allow an individual to perform an action (Gibson, 1977). Ecological affordances refer to the opportunities the environment offers for a young learner (Gibson & Pick, 2000). More than 4600 species of songbirds exist, and they exhibit incredible diversity in their song learning strategies (Beecher & Brenowitz, 2005). To understand the mechanisms that govern song learning, we need to consider how the social environment provides opportunities for interaction that the developing individuals can exploit to facilitate learning.

Communication signals like bird song are frequently organized by species specific rules. The principles organizing individual units of a signal, such as notes in a bird song, are called the song syntax. Juveniles need to acquire species typical songs, including the syntax. Probabilistic regularities present in the naturalistic vocal environment may allow young learners to segment the information in ways that could assist signal perception and production. Thus, communication signals in naturalistic environments may provide structural organization that facilitates vocal development.

The dominant theoretical model of song learning ignores this integrative ecological approach. Isolation rearing combined with passive exposure to pre-recorded, invariant- and unisensory stimuli that exhibit great experimental control and poor ecological validity have characterized the field from the start. This traditional approach has segmented song into the innate “blueprint” that then is augmented by the learned components of species typical vocalizations (e.g. Marler & Sherman, 1983; Marler, 2004). It has described two temporally distinct learning phases. During the first, instructive, phase young songbirds passively memorize a song model. Then months or weeks later, birds start practicing to match the “song template” in the second, sensorimotor, phase (Konishi, 1965, Marler, 1970, Marler & Peters, 1982). Juveniles use auditory feedback to compare immature vocalizations with the stored template, and use internal trial and error to improve gradually. The learning process is usually described as imitation.

Isolation rearing with tape tutoring does not reveal the full learning potential of a species, only what route development takes in a specific environment (Lehrman, 1953). The question is then, *from what* are juveniles deprived, when studied out of their social context?

Not surprisingly, live tutoring experiments came to very different conclusions about learning mechanisms compared to passive tape tutoring, in such qualities as the timing of sensitive periods, speed and accuracy of song acquisition, tutor song selection and the role of the social environment on the learning process (for review see Baptista & Gaunt, 1997). In a more naturalistic context, juveniles are exposed to a fundamentally different learning environment. Information can be multisensory compared to unimodal presentation of the tutor song. The structure of stimuli in nature may be variable compared to repeated exposure to the same songs. Immature vocalizations of the juveniles may gain functional significance during learning if they elicit reaction from social companions. Lastly, reinforcement may come from the social environment rather than only internally.

The importance of the natural learning environment has been intensely studied in the gregarious brown-headed cowbirds. Plastic song of juvenile males elicits contingent and rapid visual gestures from adult females, termed wing strokes that influence the potency of the final song outcome (West & King, 1988). Juvenile males attend to the female reactions, frequently approaching them, and change the content of their song after wing strokes (West & King, 1988). Females exhibit choosiness, as they differ in their selectivity to respond to high versus low potency songs (Hamilton, King, Sengelaub & West, 1997), and their preference changes with post-natal experience (West, King, White, Gros-Louis & Freed-Brown, 2006). Additionally, both males and females possess corresponding neural components adapted to learning by external reinforcement (Hamilton et al., 1997; Hamilton, King, Sengelaub & West, 1998), which suggests that songbirds integrate multisensory information during learning. Learning in the cowbird *cannot* be described as imitation based, as juveniles do not imitate non-singing females. Parallel mechanisms have been described in human infants, where contingent non-vocal reactions (touching, smiling) from caregivers prompts more developmentally advanced babbling (Goldstein, King & West, 2003; Goldstein & Schwade, 2008).

My dissertation focuses on song learning in the zebra finch, an ideal subject for investigating learning mechanisms in social context. Zebra finches are highly gregarious. They breed readily in the laboratory. Song learning in zebra finches has usually been described as imitation based (Zann, 1996). However, it also cannot be explained by mere exposure to the song model, as juveniles require social interactions with live tutors for optimal learning (Eales, 1989). Juvenile males prefer to learn from their fathers (Bohner, 1983). Even in the wild, the majority of juveniles sing similar copies of the father's song (Zann, 1990). Importantly, the song memorization and practice phases substantially overlap, providing opportunities for influence from social interactions.

What mechanisms are endorsed by live interactions? The superior learning from live tutors has been attributed to contingency between the juvenile's actions and the tutor's song (ten Cate, 1991). Operant training paradigms, in which key pecks result in a tutor song reward, became a standard way to study the course of song development (e.g. Adret, 1993; Tchernichovski, Lints, Mitra, & Nottebohm, 1999). Thus, the focus shifted from the zebra finch as a social learner toward using isolated juveniles in a Skinner box to trace the song learning process (e.g. Tchernichovski & Mitra, 2002; Tchernichovski, Lints, Deregnaucourt, Cimenser & Mitra, 2004; Ravbar, Lipkind, Parra & Tchernichovski, 2012). Such an isolate-operant approach offers a high degree of control over the experience of the young bird, and allows constant monitoring of vocal changes. Moreover, a specialized software package traces and analyzes the entire progress automatically, and allows visualization of vocal changes (Tchernichovski, Nottebohm, Ho, Pesaran & Mitra, 2000; Tchernichovski, Swigger & Partha, 2004). The technology resulted in a great amount of valuable knowledge about acoustic changes in song during development.

However, the isolate-operant approach has serious limitations, in that it reveals patterns of song development only when a juvenile is alone in a specific environment, and does not inform about mechanisms characterizing learning when social opportunities are available. Development may take an abnormal route when it occurs in social isolation. In isolate-operant training paradigms, the presentation of the tutor song contingent on key pecks decreases similarity between juvenile and tutor song if the tutor song presentation exceeds 40 motif playbacks, or 30 seconds of total presentation per a day (Tchernikovski et al., 1999), and even when exposed to limited song presentation, juveniles exhibit high individual variability in learning (e.g. Deregnaucourt, Mitra, Feher, Pytte & Tchernichovski, 2005). When juveniles are housed with a single adult male, they frequently produce a “carbon copy” of the tutor song, suggesting that the normal developmental mechanism is different. Thus, operant training limits predictions about the learning outcome in naturalistic settings.

A previous study (Houx & ten Cate, 1998) examined effects of contingency between tutor and juvenile behavior on the outcome of song learning in natural family settings. While they did not find any indication of socially guided learning, the study was based on live observations of ongoing social interaction, and did not focus on covert behaviors, such as female visual signals described in the cowbirds.

In *Chapter 2* of the dissertation, I investigate vocal development to reveal what and how juvenile songbirds learn when provided with opportunities to interact with family members. Inspired by the importance of early vocal precursors in cowbirds and human infants (West & King, 1988; Goldstein, et al., 2003), my study focuses on the potential importance of juvenile vocalizations in the learning process, whether juvenile vocalizations are associated with parental responses, and whether these interactions are associated with attributes of subsequent mature song. I analyze the microstructure of social interactions from video recordings at the frame level, as covert behaviors contingent on juvenile vocalizations can be only revealed when analyzed with high resolution. My goal is to generate hypotheses about potential influences of social interactions on vocal learning that then can be tested in subsequent experimental investigations.

I found that learning is influenced by multisensory stimulation, organized around juvenile immature song, and involves both adult males and females (fathers and mothers). The more that juvenile singing elicited contingent adult song, the better the juvenile learned its father's song sequence. In contrast, the more the juvenile sang after his father, the less accurately he learned the father's song. Female-directed male song, although not directed toward juveniles, appeared to be a more salient learning cue compared to undirected singing. The total number of female fluff-ups correlated with learning the sequence of the father's song, and female wing strokes were associated with closer approximation of the tutor song pitch in the plastic song.

In *Chapter 3* I examine the variability present in adult zebra finch song to reveal whether structure in the communication signal is used in the song learning process. Zebra finch song contains a sequence of syllables, called a song motif. While the literature frequently treats finch song as having one stereotyped motif, several studies report that syllable sequences within motifs vary substantially (Sturdy, Phillmore & Weisman, 1999; Helekar, Marsh, Viswanath & Rosenfield, 2000). These alternate motifs are usually considered to be production errors. We treat them as potentially meaningful and quantify their form and frequency. The functional significance of this song variability has not been previously investigated.

Structure in natural behaviors, including human speech and bird song (e.g. Goodsitt, Morgan & Kuhl, 1993; Okanoya, 2004) can be exploited by segmentation based on co-occurrence and transitional probabilities among building units of the information stream (Saffran, Aslin & Newport, 1996). Probabilistic regularities in the input aid perception by allowing "chunking" of the information into clusters, and help subsequent learning, a process called *statistical learning*. Perhaps young birds learn how often and in what ways their fathers vary their song motifs, and incorporate these statistical regularities into their own songs.

I compared the mature song of zebra finch juveniles and their fathers. I looked for structural regularities using a language acquisition model (in collaboration with Oren Kolodny and Shimon Edelman) that allows the comparison of multiple individuals without a need for syllable-level correspondence among different subjects. My aim was to search for grammatical regularities in the variation found across song renditions, determine the structural hierarchy, if any, over these regularities, and then compare the regularities between fathers and their offspring across multiple families. I found that coherent patterns present in the father's song were reflected in the song of his offspring, indicating that song learning in the zebra finch is more than imitation. The results suggest that birds use a statistical learning mechanism to extract the hierarchical pattern of syllables from the tutor's song that allows building an internal representation of the grammar.

The first two chapters of the dissertation explore learning mechanisms linked to regularities accessible for juveniles from their natural environment. Both studies are based on observation of unmanipulated zebra finch families. In *Chapter 4* I explore the contribution of social feedback, using contingent song playback. I place birds into a highly controlled vocal environment, designed to be similar to interactions observed in Chapter 2, that may be used by the juvenile to learn from environmental regularities associated with his immature vocalization. I offer tutor song feedback contingent on the plastic song of the juveniles. In the contingent treatment group, juveniles receive tutor song playback immediately following plastic song vocalizations. Yoked control subjects received tutor song playback time locked to their contingent partner. If the timing of feedback on immature vocalizations is important in learning, I expect learning differences between the two treatment groups.

Although contingent and yoked subject learned the tutor song model similarly well, I found that contingent subjects developed a significantly more stable song pattern and a smaller proportion of song variants compared to the yoked birds.

Overall significance

I take an approach that re-integrates juvenile vocal learners with social context, and investigates developmental mechanisms in a species specific, variable, multisensory environment. My approach allows creating testable predictions for subsequent experimental analysis in controlled environments, by taking into account developmental mechanisms characteristics of the particular species of interest. I found that this approach reveals unexpected complexities in song learning, complexities that help explain naturalistic learning as well as inconsistencies in prior experimental observations.

REFERENCES

Adret, P. (1993). Operant-conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour, 46*(1), 149-159. doi: 10.1006/anbe.1993.1170

Baptista, L. F., & Gaunt, S. L. L. (1997). Social interaction and vocal development in birds. *Social influences on vocal development*, 23-40. doi: 10.1017/cbo9780511758843.003

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20*(3), 143-149. doi: 10.1016/j.tree.2005.01.004

Bohner, J. (1983). Song learning in the zebra finch (taeniopygia-guttata) - selectivity in the choice of a tutor and accuracy of song copies. *Animal Behaviour, 31*(1), 231-237. doi: 10.1016/s0003-3472(83)80193-6

Deregnaucourt, S., Mitra, P. P., Feher, O., Pytte, C., & Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. *Nature, 433*(7027), 710-716. doi: 10.1038/nature03275

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour, 37*, 507-508. doi: 10.1016/0003-3472(89)90097-3

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*: Oxford University Press, USA.

Gibson, J. (1977). The concept of affordances. *Perceiving, acting, and knowing*, 67-82.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515-523.

Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language, 20*(2), 229-252.

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1997). A brain of her own: A neural correlate of song assessment in a female songbird. *Neurobiology of Learning and Memory, 68*(3), 325-332. doi: 10.1006/nlme.1997.3781

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1998). Visual and song nuclei correlate with courtship skills in brown-headed cowbirds. *Animal Behaviour, 56*, 973-982. doi: 10.1006/anbe.1998.0848

Helekar, S. A., Marsh, S., Viswanath, N. S., & Rosenfield, D. B. (2000). Acoustic pattern variations in the female-directed birdsongs of a colony of laboratory-bred zebra finches. *Behavioural Processes, 49*(2), 99-110. doi: 10.1016/s0376-6357(00)00081-4

Houx, B. B., & ten Cate, C. (1998). Do contingencies with tutor behaviour influence song learning in zebra finches? *Behaviour, 135*, 599-614.

Konishi, M. (1965). The Role of Auditory Feedback in the Control of Vocalization in the White-Crowned Sparrow1. *Zeitschrift fur Tierpsychologie, 22*(7), 770-783.

Lehrman, D. S. (1953). A critique of Lorenz, Konrad theory of instinctive behavior. *Quarterly Review of Biology, 28*(4), 337-363. doi: 10.1086/399858

Marler, P. (1970). A comparative approach to vocal learning: song development in White-crowned Sparrows. *Journal of Comparative and Physiological Psychology, 71*(2p2), 1.

Marler, P., & Peters, S. (1982). Long-term storage of learned birdsongs prior to production. *Animal Behaviour, 30*(2), 479-482.

Marler, P., & Sherman, V. (1983). Song structure without auditory-feedback - emendations of the auditory template hypothesis. *Journal of Neuroscience, 3*(3), 517-531.

Okanoya, K. (2004). The bengalese finch - A window on the behavioral neurobiology of birdsong syntax. In H. P. Zeigler & P. Marler (Eds.), *Behavioral Neurobiology of Birdsong* (Vol. 1016, pp. 724-735).

Ravbar, P., Lipkind, D., Parra, L. C., & Tchernichovski, O. (2012). Vocal Exploration Is Locally Regulated during Song Learning. *Journal of Neuroscience, 32*(10), 3422-3432. doi: 10.1523/jneurosci.3740-11.2012

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (1999). Note types, harmonic structure, and note order in the songs of zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology, 113*(2), 194-203. doi: 10.1037/0735-7036.113.2.194

Tchernichovski, O., Lints, T., Mitra, P. P., & Nottebohm, F. (1999). Vocal imitation in zebra finches is inversely related to model abundance. *Proceedings of the National Academy of Sciences of the United States of America, 96*(22), 12901-12904. doi: 10.1073/pnas.96.22.12901

Tchernichovski, O., Lints, T. J., Deregnaucourt, S., Cimenser, A., & Mitra, P. P. (2004). Studying the song development process rationale and methods. In H. P. Zeigler & P. Marler (Eds.), *Behavioral Neurobiology of Birdsong* (Vol. 1016, pp. 348-363).

Tchernichovski, O., & Mitra, P. P. (2002). Towards quantification of vocal imitation in the zebra finch. *Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology, 188*(11-2), 867-878. doi: 10.1007/s00359-002-0352-4

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour, 59*, 1167-1176. doi: 10.1006/anbe.1999.1416

Tchernichovski , O., Swigger, D., & Partha, M. (2004). Sound Analysis Pro User Manual. CCNY, New York.

ten Cate, C. (1991). Behavior-contingent exposure to taped song and zebra finch song learning. *Animal Behaviour, 42,* 857-859.

West, M. J., & King, A. P. (1988). Female visual-displays affect the development of male song in the cowbird. *Nature, 334*(6179), 244-246. doi: 10.1038/334244a0

West, M. J., King, A. P., White, D. J., Gros-Louis, J., & Freed-Brown, G. (2006). The development of local song preferences in female cowbirds (Molothrus ater): Flock living stimulates learning. *Ethology, 112*(11), 1095-1107. doi: 10.1111/j.1439-0310.2006.01264.x

Zann, R. (1990). Song and call learning in wild zebra finches in south-east Australia. *Animal Behaviour, 40*(5), 811-828.

Zann, R. A. (1996). *The Zebra Finch: Synthesis of Field and Laboratory Studies*. New York: Oxford University Press.

CHAPTER 2

SOCIALLY GUIDED SONG LEARNING IN THE ZEBRA FINCH: EFFECTS AT MULTIPLE TIME SCALES

Abstract

Vocal development of juvenile songbirds and human infants requires interactions with adults to develop its proper form and function. Learning in many songbird species requires more than mere exposure to the song model, as the timing, speed, and accuracy of learning typically differ when exposed to live versus taped tutors. What mechanisms drive the superior learning performance associated with live interactions? Young learners can elicit contingent responses from adult social companions (e.g., in brown-headed cowbirds and human infants) in different modalities (e.g., wing strokes and social smiling) that reinforce learning and promote more advanced vocalizations. We studied vocal development in zebra finches to assess relations between vocal precursors in juveniles and contingent responses from the parents. Eight families were recorded between 30-75 days, and social interactions surrounding juvenile or adult song were analyzed. Juvenile and father song were compared when juveniles reached song maturation.

We find that both parents contribute to juvenile song learning. Juveniles whose immature song received more contingent song from their fathers copied their father's song sequence more accurately. Conversely, adult song preceding juvenile song predicted less accurate learning. The courtship song of the adult male appears to be an especially powerful signal. Non-vocal feedback from mothers also predicts song learning. The number of maternal fluff-ups around and especially before juvenile song positively correlated with later sequential similarity between

adult and juvenile males, and female wing strokes were associated with-real time changes in the dominant frequency of the juvenile song syllables. Our results offer an alternative to the dominant model of song learning, and our findings imply that the mechanisms underlying the development of vocal communication may be best understood when studied within a social system.

Introduction

Numerous analogies have been proposed between bird song and human speech learning (Doupe & Kuhl, 1999; Brainard & Doupe, 2002; Goldstein & Schwade, 2010). One similarity is dependence on the social environment: young birds and babies need interaction with adults to develop vocalizations with proper form and function. As altricial taxa that depend on parental care for survival during the period of vocal development, the typical developmental niche of both songbirds and human infants includes structured social stimulation.

Paradoxically, our knowledge about mechanisms of song learning is fundamentally influenced by the outcome of experiments using isolate rearing. Since the pioneering studies of Thorpe on chaffinches (1958), numerous songbird species have been brought into the laboratory and reared in a highly controlled acoustic environment in social isolation, resulting in the prevalent model of song learning (Konishi, 1965). This model proposes a two-phase learning process. During a sensory or “instructive” phase, young birds memorize an adult song in the form of a “song template”. During the second sensorimotor or practice phase, juveniles try to match their vocal output to the memory trace of the tutor song. The learning process is frequently referred to as “imitation”, based on the similarity between the tutor song and the crystallized song outcome of the pupil. As immature vocalizations are thought to reach their final form via practice and maturation of the vocal production system, the learning model assumes a passive role of the juveniles in the developmental process. Mere exposure to the tutor song is thought to be enough to drive memorization and practice.

However, the principles of learning revealed by isolate rearing studies are restricted to very specific circumstances, and may not generalize to the actual learning pathways of birds in nature. More naturalistic approaches using live tutoring paradigms reach different conclusions about many aspects of vocal development than do studies employing tape tutoring in social isolation (for review see Baptista & Gaunt, 1997). For example, live interactions affect tutor choice and may even override interspecific boundaries (Baptista & Petrovich, 1984, 1986; Immelmann, 1969; Bohner, 1983), prolong the length of the sensitive phase (Baptista & Petrovich, 1984; Nordby, Campbell & Beecher, 2001), and accelerate the learning process (Baptista, Bell & Trail, 1993a). Consequently, live tutoring is far more effective at teaching song than passive tape exposure for many songbird species (Thielcke, 1970; Baptista & Petrinovich, 1986; Chaiken, Bohner & Marler, 1993; Waser & Marler, 1977; Derégnaucourt, Poirier, Kant, Linden & Gahr, in press).

The superior results of live tutoring challenge assumptions about learning mechanisms grounded in isolate rearing studies. In live interactions, the developing juveniles are active participants in the development of their communication skills. For example, matched counter-singing interactions between juveniles and “mimicked” territorial rivals in white-crowned sparrows will consolidate the young bird’s vocal repertoire (Nelson & Marler, 1994). If during song practice juveniles become both senders and receivers, then there is the potential that interactions elicited by juvenile vocalizations shape vocal ontogeny. The plastic song of juvenile male brown-headed cowbirds elicits wing strokes from adult females, after which the singers modify the content of their song (West & King, 1988). Thus, to understand development, it is crucial to explore the learning opportunities juveniles can exploit from their social environment. An ecological approach, which emphasizes the fit between a learner's sensory and motor capacities and the opportunities afforded by the immediate environment, may yield insight into the socially-embedded learning process.

As songbirds exhibit enormous diversity both in their natural ecology and song learning strategies (Beecher & Brenowitz, 2005), understanding their learning opportunities will depend on the natural ecology of the species of interest. In numerous territorial North-American sparrow species, memorization and song practice are temporally separate processes (Marler, 1970; Marler & Peters, 1981; Marler & Peters, 1987). In territorial species, neighbors often share repertoire elements. If song learning is solely based on the memorization process, song sharing is possible if juveniles memorize the repertoire elements of their territorial neighbors. Alternatively, in species where juveniles disperse from their natal territory after the ability to acquire novel songs, immature vocalizations may gain an important role during the learning process. Juveniles will retain elements from their plastic song in their final repertoire as a function of experience with territorial neighbors, a process called *action based learning* (Nelson, 1992; Nelson & Marler, 1994; Marler, 1997).

In socially gregarious species, juvenile vocalizations may elicit different learning opportunities compared to territorial songbirds. Learning cues might come from a variety of stimuli involving multiple sensory channels. For example, in the gregarious brown-headed cowbirds immature juvenile song elicits wing strokes from females, and the males use the females' feedback to adjust their repertoire during vocal ontogeny (West & King, 1988). Although social interactions organized around the immature vocalizations of juveniles have the potential to play an active role during development, instances of *socially guided learning* have not been clearly identified in bird species other than the cowbird.

The highly gregarious zebra finch is the most popular model species for studies of the neurobiology of song learning due to its fast maturation and prolific nature in captivity. Despite the large amount of knowledge accumulated about its neurobiology, the exact mechanisms of song leaning have not yet been identified. Here we take an ecological approach to song learning in the zebra finch. There exists strong indirect evidence that social interactions may have an important role in song learning. Only male zebra finches sing, and their songs consist of a single song type. The song is composed of variable syllable types organized into a consistent, but not stereotyped, pattern called a motif. The motif is initially preceded by a train of repetitive introductory elements. A single song bout contains 1-8 motif repetitions (Price, 1979). In the wild, about 65 % of juvenile males learn their song from their fathers even if other males are available as song tutors, and about 80% copy the contact call of their father (Zann, 1990). In laboratory settings, juveniles also learn the choreography associated with courtship song from their father (Williams, 2001). Bohner (1983) separated juveniles from their fathers at the age of 40 days and reared them in visual and acoustic contact to both the father and an unrelated male; juveniles chose to learn their father's song. Moreover, siblings from the same clutch develop distinct songs with a wide range of similarity to the father's song, but when housed in isolation with a single adult male, juveniles frequently produce an exact imitation of the song model (Tchernichovski & Nottebohm, 1998). All of the above indicate that factors beyond mere exposure are driving learning.

The aspects of social interactions that predict song tutor selection have been investigated extensively in the zebra finch. Song learning is better from live tutors than from tapes or noninteractive playback paradigms (e.g. Derégnaucourt et al., in press). For example, Eales (1989) broadcasted live adult song through speakers although the juveniles were unable to interact with the tutors. As a consequence none of the juveniles learned the tutor song. In an aviary environment containing 12 song tutors, juveniles learned their songs primarily from two of the available tutors. An adult male's salience as a song tutor was mainly based on the amount of parental care delivered to fledglings (Williams, 1990). Another aviary study found that physical proximity between tutors and juveniles predicted song tutor choice, and fathers were more likely to be chosen as tutors than unrelated males (Mann & Slater, 1995). When given a choice between two adult males, juveniles preferred to learn from the one who was more aggressive toward them (Clayton, 1987b; Jones & Slater, 1996). Pair bonds between adult males and females may direct juvenile learning, as juveniles preferentially learn from the tutor paired with their mother rather than a single male, even if the single male is their father (Eales, 1987b; Mann & Slater, 1994). Visual cues are important in tutor selection. Juveniles preferred to choose tutors with the same color morph as their rearing parent (Mann, Slater, Eales & Richards, 1991; Mann & Slater, 1995). Auditory information also regulates tutor choice. Juveniles removed from the father at 35 days of age learned from a male with a song similar to the father's (Clayton, 1987b).

Importantly, there is an extensive overlap between the memorization and acquisition phases of song learning in the zebra finch (Slater, Eales, & Clayton, 1988). Juveniles may begin to memorize the tutor's song as early as 25 days post-hatching (Roper & Zann, 2006). Males start producing subsong around 30 days of age and develop a recognizable pattern by 65 days. Until then young males are still capable of modifying their song based on external cues (Slater, Eales & Clayton, 1988). Thus, overlap between memory formation and the practice phase provides potential opportunities for social interaction to influence song learning. Once adulthood is reached at around 3 months of age, the song of the young bird will finally be stable and resistant to change for the rest of the bird's life (e.g. Brainard & Doupe, 2002).

We focused on the developmental period when memorization and practice overlap to investigate mechanisms responsible for superior learning from live tutors. Social interactions are continuous and rapid. To capture these dynamic changes, we applied an ecological perspective to map the microstructure of social interactions that occur around the production of immature vocal behavior. We collected videos of large amounts of social information that we then analyzed at the frame-level. This method helps to identify infrequent behaviors that may have a strong impact on the learning process. In brown-headed cowbirds, females display wing strokes on average once for every 100 juvenile song renditions. However the rare wing strokes reliably shape the potency of the male song (West & King, 1988). Our approach provides an account of how sporadic behaviors are important for canalizing the developmental trajectory. Such an approach has proven useful for investigating the dynamic structure of multimodal parent-infant interactions on early speech and language learning (e.g. Goldstein, King & West, 2003; Yu, Smith, Shen, Pereira & Smith, 2009; Smith, Yu & Pereira, 2011).

What properties of live interactions are responsible for enhanced learning compared to passive tape exposure? One possibility is response contingency, by which a contingent feedback on a juvenile action may increase learning due to heightened arousal or attention (ten Cate, 1991). This hypothesis led to the development of an operant procedure, in which tutor song was provided for key pecks. Experiments using operant conditioning have returned mixed results. Although in some studies the paradigm resulted in higher copying fidelity of the tutor song compared to passive tape exposure (Adret, 1993; Derégnaucourt et al., in press), operant training frequently resulted in high learning variability among subjects (Derégnaucourt, Mitra, Feher, Pytte & Tchernichovski 2005; Phan, Pytte & Vicario, 2006). Other studies applying operant contingencies and stimulus-stimulus contingencies using audiovisual compound stimuli did not find song learning differences between contingent and control subjects (Bolhuis, Van Mil & Houx, 1999; Houx & ten Cate, 1999 a,b). Moreover, across all studies, subjects failed to reach the level of similarity to the tutor's song shown by males raised with live tutors. This suggests that the stimuli were not adequate for modeling real time social interactions. To date, only a single observational study attempted to find contingencies in interactions between tutors and pupils in natural family settings to reveal their possible influence on the song learning process (Houx & ten Cate, 1998). The authors did not find any clear relations between social interactions and learning, however, they did not include female visual signals in the interactions they assessed. Moreover, the method of observing live interactions may not have been able to catch the subtle and rapid changes in the behavior of individual birds.

What forms of typically-available contingent interaction might support real-time song learning? Juveniles may interact vocally with their fathers, which may promote retention of elements from the father's vocal repertoire, as described during the process of action-based learning (Nelson & Marler, 1994). As listeners, non-singing mothers also might contribute to the learning process and might use visual, and/or tactile modalities to provide feedback. Zebra finch females differentially react to features of the male's song. Directed courtship song is strongly preferred by females over undirected song (Woolley & Doupe, 2008). Moreover, females favor males with more elaborate songs (for review see Riebel, 2009). Three types of female visual signals have been linked to female preference in songbirds: copulation-solicitation displays in numerous songbird species, e.g. canaries (Vallet & Kreutzer, 1995), song sparrows (Searcy & Marler, 1981; O'Loghlen & Beecher, 1997), swamp sparrows (Anderson, 2009), brown-headed cowbirds (King & West, 1977); wing strokes in brown-headed cowbirds (West & King, 1988), and shakes or fluff-ups in the zebra finch (Vyas, Harding, Borg & Bogdan, 2009). Zebra finch females demonstrate copulation-solicitation displays only occasionally. However, wing strokes and fluff ups may be signals of female song preference that they could use to exert selective pressure on male vocal development.

A number of studies attempted to reveal females' contribution to song acquisition in the zebra finch, although none investigated direct behavioral interactions between adult females and juvenile males. Jones and Slater (1993) attempted to reveal the female shaping effects on eventual song repertoire formation, in an experiment in which juveniles were given a choice to learn from two unrelated adult males simultaneously. Subsequently the juvenile was placed with a female that was a daughter of one of the tutors. The authors found a weak preference for learning from the female's father. Moreover, juveniles housed with a deaf female sang more frequently and more atypical songs than subjects housed with females with intact hearing (Williams, 2004).

Taken together, previous research on zebra finch suggests the possibility that social interactions involving both adult males and females are important during song learning. However, none of these studies focused on fine-grained analysis of contingent interactions in a social context. The goal of the present research was to investigate the nature and significance of different social interactions as organized by immature vocalizations of the juveniles in unmanipulated families. Our study focused on a developmental period during which song memorization and production occur as parallel processes. First, we explored the effects of directed and undirected adult male song along with female visual signals (wing strokes and fluff ups) on juvenile learning. We predicted that parents respond consistently to juvenile vocalizations, and that juveniles can use social feedback to modify their song outcome.

Second, we aimed to determine whether specific types of social interactions cause changes in song outcome of the juveniles on multiple time scales. The majority of previous studies focused only at the end of the learning process and analyzed the final repertoire. However this approach does not give an opportunity to track real-time changes associated with facilitative social interactions during vocal ontogeny. We compared juvenile plastic song occurring before and after social feedback to determine whether any vocal changes after social feedback move in the direction of the mature song. Because the dominant frequency of zebra finch song is learned, precisely controlled and sensitive to reinforcement (Tchernichovski, Mitra, Lints & Nottebohm, 2001; Tumer & Brainard, 2007), we predicted shifts in the dominant frequency of the juvenile syllables toward the fathers' syllables as a function of female visual signals compared to songs that did not get feedback. We focused on wing strokes based of their prior efficacy on the song learning process in brown-headed cowbirds (West & King, 1988).

Methods

Breeding

Nine pairs of zebra finch were set up in single cages (46 x 44 x 36 cm, n = 6 pairs, and 62 x 44 x 36 cm, n = 3 pairs) with a nest box and nest material, in a communal room in both visual and acoustic contact with each other and with birds in other aviaries. The birds were kept on a 14:10 light cycle and were provided with mixed foreign finch seed and water *ad libitum*.

Each pair nested, laid eggs, hatched and fledged 2-5 young (*M* = 4.0, S*D* = 1.22). Juveniles were color banded at about 3 weeks of age (*M* = 21.56, range 16 - 26 days; age always refers to the hatching date of the oldest juvenile). Families living in the smaller cages were transferred to larger cages when the oldest chick was about 24 days old. All families were recorded in identical cages. Nest boxes were removed at about 35 days in order to prevent re-nesting of the parents during the study.

During preliminary behavioral annotations it proved unmanageable to assess all the social permutations caused by having 4 juvenile brothers within a single family, and so we limited our analysis to families containing fewer than 4 sons. Thus, data from eight families containing 14 juvenile males and 15 juvenile females total were included into the final analysis.

Behavioral recordings

Video recordings of family interactions began between 19 and 35 days after hatching (*M* = 26, *SD* = 5.4) and continued at approximately 3-day intervals until 70 - 95 days after hatching. In this study, we focus on analysis of video recordings made between 30 and 75 days of age. Within this developmental window each family was recorded for about 16 hours (*M* = 16.01, S*D* = 2.46), in sessions occurring about twice a week. The entire cage containing the parents and offspring was placed overnight in a sound attenuating enclosure (140 x 98 x 80 cm) made of PVC pipe and sound attenuating foam glued onto plank boards. The enclosures were open in the front and the bottom and were used to enhance sound quality of recordings by eliminating reverberation.

Family interactions were recorded for one hour during the following morning using a Sennheiser K6/ME64 shotgun microphone connected to a Canon MiniDV ZR930 camcorder with Fujifilm DVCassette miniDVs. After the recordings, family cages were returned to the communal housing room and left undisturbed until the subsequent recording. Families were kept together until about 99 days (*SD* = 9 days). The juveniles were then transferred to single-sex aviaries with other juvenile birds.

Behavioral analysis

Sound files were created from video recordings using Soundtrack Pro (version 3.0.1). Elan: EUDICO Linguistic Annotator Version 4.0.0. (Copyright © 2001 - 2012 Max- Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands) was used to synchronize the video and sound files and to code the behavior of juvenile males and their parents. Elan permits annotations of multiple categories of behavior from individual animals with a frame-by-frame accuracy onto different fields associated with the video record.

The focus of the initial analysis was to obtain relations between songs of the fathers and juvenile males. Thus in an initial pass through the records, all adult male and juvenile song bouts were annotated with a single frame accuracy. Song in zebra finches occurs in bouts, which we defined in behavioral annotations as singing separated by less than 1000 msec of silence. Songs were coded as *directed* whenever the singer initiated singing while clearly facing another individual. Directed songs often included courtship dances (Williams, 2001). If the bird initiated his song without directing it toward another individual, the entire song was coded as *undirected*. After finding every occurrence of song in the hour-long recordings, 15-second time windows before and after the songs were marked in Elan for every target individual. When songs were produced in quick succession, thus we could not create separate after and before time windows, the time between songs was labeled as “between” (Figure 2.1).

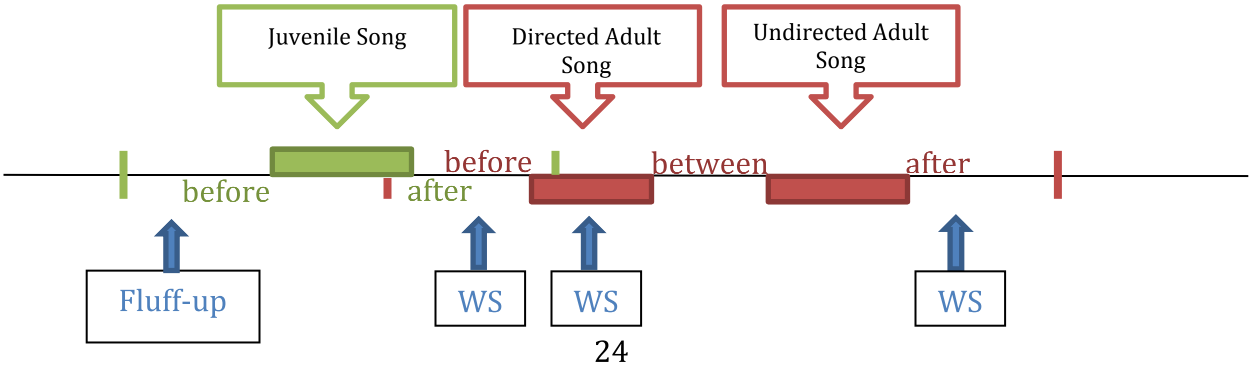

Figure 2.1 15-sec time windows were created before and after every adult and juvenile male song. When juvenile or adult songs followed in quick succession and we could not create separate before and after windows, the time period between songs was labeled as “between”. Female gestures (fluff-ups and wing strokes, WS) were not coded outside of the time windows. Contingent interactions were assessed based on temporal relations between behaviors (e.g. directed adult male song after juvenile song, female fluff-up before juvenile song, etc.).

Two behaviors by the mother, wing strokes and fluff-ups, were coded during songs as well as before, after, and between songs. Wing strokes are a single fast movement of one (unilateral) or both (bilateral) wings, made independently of large body movements. Wing strokes were coded as unilateral or bilateral, but summarized for subsequent data analysis as the number of “total” wing strokes. Because birds tend to adjust their feathers once they land after flying, wing strokes were not coded within 1 s after landing. Wing strokes were coded after hopping and jumping, but only after the female was in contact with the surface. Wing strokes were not coded during preening, or when the movement of the tail appeared to cause the movement of the wing. Tiny wing adjustments during sudden bending were not coded as wing strokes. As these movements can be brief and subtle, wing strokes were independently coded by two individuals. A third person then checked both coders’ annotations and resolved any discrepancies. Annotations from this last assessment were used in data analysis.

Fluff ups are brief feather fluffs accompanied with a quick shaking movement of the entire body from side to side followed by smoothing the feathers back down (Vyas et al., 2009). Initially, two independent coders annotated the presence of fluff ups. As they were highly consistent (87% intercoder reliability), annotations of only one coder were used in subsequent analyses.

Behavioral contingencies

From the annotations, we summarized the occurrence of directed, undirected, and total (directed + undirected) adult male songs, juvenile songs, and female wing strokes and fluff ups for the entire developmental period for each family. Because juveniles sang directed song rarely around male song during the entire observation period (*M* = 2, *SD* = 1.77), we used the total juvenile male song in subsequent analysis. We also analyzed the total number of wing strokes by pooling uni- and bilateral wing strokes.

Behaviors were assigned to categories based on their temporal relationship, namely how many behaviors occurred before, during, after, and between adult male or juvenile songs, as illustrated in Figure 2.1.

Mature song recordings

Mature song was recorded from the sons (n = 11) in the fourth month after hatching (*M* = 104, *SD* = 9.59 days) and from their fathers (n = 8). Three sons did not sing during these initial recording times, so their songs were recorded in the fifth month after hatching, between the ages of 142-153 days.

For all of the recordings, males were placed in a sound-proof room overnight in a 46 x 44 x 36 cm cage. The following morning, an adult female zebra finch was placed in an identical cage next to the male's cage and recording began. If the male did not sing in 60 minutes, further recordings were attempted on subsequent days until we obtained at least 10 song bouts in each recording.

All recordings were made with a Sennheiser shotgun microphone attached to a Canon MiniDV ZR930 camcorder on Fujifilm DVCassette miniDVs. MiniDV tapes were ingested with a JVC Super VHS ET Professional deck at 44.1 KHz. Sound files were created using Soundtrack Pro (version 3.0.1) and uncompressed sound files were saved as separate wav files.

Mature sound analysis

Zebra finch song has been considered highly stereotyped. In fact, although males sing a relatively consistent song motif, syllable ordering shows substantial variability (Scharff & Nottebohm, 1991; Sturdy, Phillmore & Weisman, 1999; Helekar, Marsh, Viswanath & Rosenfield, 2000). About 60% of motifs belong to the most common motif variant, frequently referred to as the “dominant” or “canonical motif” (Sturdy et al., 1999; Brainard & Doupe, 2001; Zann, 1996). For the purpose of sound analysis, we randomly choose 10 renditions of the dominant motif from each recording (both sons and fathers). When selecting for motifs, motifs containing background noise, female calls or cage noise were omitted. We also skipped the first motif of a song bout as it tends to deviate from subsequent motifs. To assess song learning, we then compared 10 motifs from the mature song of sons to 10 motifs from their fathers, performing 10 x 10 comparisons within the similarity batch feature of Sound Analysis Pro 2A.04 (SAP2) software (Tchernichovski, Swigger & Partha, 2004). For our analyses, we used the default settings of the similarity module and derived the three different measurements of similarity between compared songs that this software package provides:

Percent similarity measures the likelihood that two songs are related to each other. Percent similarity score is computed over longer intervals of sound (typically 50 - 70 msec), and refers to the percentage of tutor’s sounds included in *final sections* (as defined in the SAP2 manual), which are superior similarity sections that passed a final stage of filtering in the analysis process by omitting similarity sections that explain very little similarity (and are likely to be noise).

Accuracy is computed across shorter time windows (5 - 10 msec) and indicates across the similar segments the accuracy of the vocal match between the two songs.

Sequential match accounts for the temporal order (syntax) of sounds. Sequential match is calculated by sorting the *final sections* according to their temporal order in reference to sound 1, and then examining their corresponding order in sound 2.

Statistical analysis

To assess the contribution of behavioral interactions during development on song learning, we correlated behaviors occurring in the 15 sec time windows with similarity scores to father obtained from the mature song. To control for the effect of juvenile song activity and the number of observations per family, we partialed out juvenile song rate per hour from the correlations. Due to multiple comparisons, we controlled for the false discovery rate by adjusting the p-values (Benjamini & Hochberg, 1995).

Some families contained more than one son, and so the 14 juvenile males could not be treated as independent data points. To analyze the effects of family on the three song outcome measures we conducted separate intercept-only linear mixed models with family as the random factor and song similarity measures as the dependent variable in each analysis. For accuracy between juvenile and father, the percent of variance attributed to family was 4.7 %, for percent similarity 24%, and for sequential match 73%. Because family accounted for large proportion of variance on the latter two outcome measures, family identity was included in all subsequent data analyses.

Due to the limited sample size it was not possible to include all behavioral variables as predictors in a single multilevel regression model. Moreover, numerous behaviors were strongly correlated with each other. Therefore, a factor analysis was used to determine how many independent factors could be extracted from the behavioral interactions. To satisfy the assumptions of the analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy needs to be above 0.5 (Field, 2009). Given the limited sample size, we satisfied this assumption by performing four factor analyses.

First, three different factor analyses were conducted (all with Varimax rotation): the first set analyzed adult male singing behavior relative to juvenile song, the second set analyzed wing strokes and fluff ups relative to juvenile song, and the third set analyzed juvenile song relative to their fathers' song. We performed a separate linear mixed model between every factor we extracted from the three factor analyses and the three similarity measures between juvenile and father, with song rate/hour as a fixed effect and family as random effect.

Next, we conducted a factor analysis by combining the paternal and maternal factors that predicted song learning of the juvenile. This generated two components. We completed another linear mixed model analysis to determine which of the two components contributed to song similarity measures by using song rate/hour as the fixed and family as the random factor. Finally, behavioral variables with the strongest contribution to the component predicting song learning were included into separate backward multiple linear regressions to assess their involvement in similarity measures between the tutor and juvenile song.

Characterizing real-time changes in plastic song

To characterize potential real-time changes in juvenile song caused by social feedback during development, immature juvenile vocalizations were analyzed between 60 and 75 days of age from 12 juveniles from seven of the families studied. One family from the original data set was excluded due to insufficient sound quality of the recordings. This developmental period was chosen because plastic song typically begins 40 - 60 days of age and ends at 80 - 100 days of age, and contains discrete, recognizable syllables (Johnson, Soderstrom, & Whitney, 2002). Song bout production also increases dramatically during plastic song from 49 to 70 days of age (Johnson et al., 2002). From each family about 5 recordings were included in the analysis.

Motif and syllable selection

We first created catalogs for each juvenile male of all syllable types that he used in his mature song. During an initial pass, we determined the maximum number of available adjacent motif pairs associated with female wing strokes and not occluded by background noise from the recordings. For each juvenile, up to twelve songs were randomly selected for each of the 5 recordings that had been produced between 60-75 days, during the plastic phase of song learning, and had been associated with female wing strokes. These songs were broken down into constituent syllables identified in the crystallized song recordings using Syrinx software (John Burt, www.syrinxpc.com). To obtain fine-grained resolution of the songs around wing strokes, song bouts here were defined as strings of syllables with no silent intervals longer than 500 msec. Syllables occluded by noise or calls, that appeared to be precursors of multiple syllables in the crystallized song, or that could not be identified by visual inspection when compared to the crystallized song, were excluded from analysis.

The timing of each wing stroke was marked within the plastic song, and song motifs were grouped based on their timing relative to the wing stroke. If a wing stroke occurred during or within a 500 msec silent interval after a motif, then we defined that motif as eliciting a wing stroke. If there was more than one "eliciting” motif in a song bout, only the first was analyzed. Motifs that preceded the “eliciting” motif in the same song were assigned into a “before” category, and motifs following it into an “after” motif category (Figure 2.2). Ideally, the three motifs before, during, and after the female wing stroke would be adjacent to each other in the same song. Due to the activity of other family members, syllables were frequently occluded by noise. Therefore we frequently found only pairs of adjacent motifs suitable for acoustic analysis (before - eliciting, or eliciting - after). We extracted every syllable from the pairs of selected song motifs. If a syllable was repeated more than once in a row within a motif, e.g. D in ABCDDE, the first example of the syllable that was not occluded by noise was used for the subsequent analysis.

We wished to determine whether we could measure real time changes in song that could be related to the presence of wing strokes (wing stroke songs)—in other words, whether acoustic features of pairs of identical syllables extracted from adjacent motifs categories change due to the intervention of a female wing stroke. Syllables were grouped based on corresponding motif categories, and the mean dominant frequency was calculated for each category. To serve as a control, an equal number of syllable pairs were selected from adjacent motifs from songs that did not receive a wing stroke (control songs). We compared the dominant frequency values of syllables from wing stroke songs to frequency values of control syllable pairs that were not associated with wing strokes.

In this data set, juveniles produced 4.83 (*SD* = 2.12) different syllable types in song associated with wing strokes that could be identified either as a syllable of the juvenile's mature song or a version of a syllable in the father's song, based on visual inspection of the waveform of the syllable's spectrogram. Each syllable type occurred in 4.33 (*SD* = 1.67) recordings within the data set.

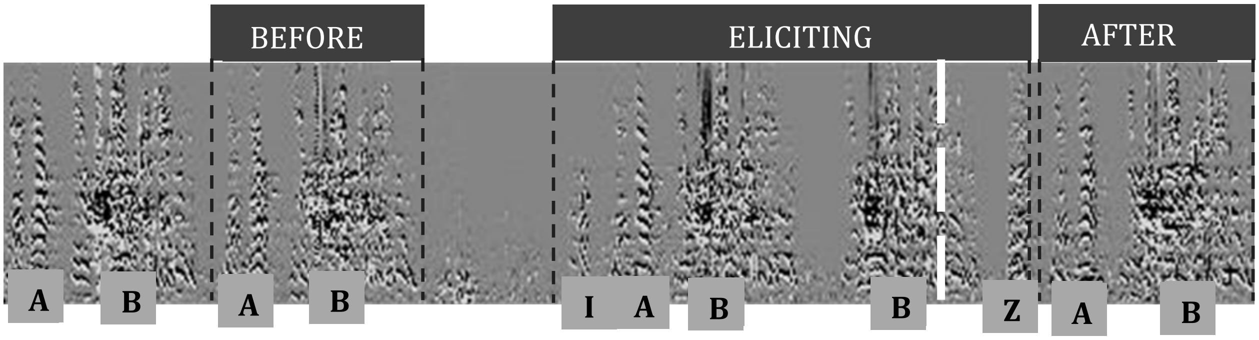

Figure 2.2 Spectrogram of a song that received a wing stroke. The motif in which a wing stroke (white dashed line) occurs is labeled as "eliciting", adjacent motifs are labeled as "before" and "after".

Characterizing real-time changes in dominant frequency

The dominant frequency of each selected syllable from both wing stroke and control songs was determined using SAP2 using the Explore & Score feature's auto-segmentation tool (entropy: -9.5; dB: 44; FFT data window: 9.27 msec; advance window: 1.00 msec; frequency range: 8000 Hz; contour threshold: 10). This program calculates dominant frequency using two different methods depending on the structure of the syllable (Tchernichovski et al., 2004). For harmonically complex syllables composed of related harmonic frequencies, the dominant frequency is calculated as an estimate of the fundamental frequency, which is equal to the frequency difference between successive harmonics. For syllables without well-defined harmonic structure, dominant frequency is calculated as mean frequency, a measure of the center of the distribution of power across frequencies. When analyzing a syllable, SAP2 first computes a cepstrum (spectrum of log spectrum) to detect the presence of a harmonic dominant frequency. If not detected, SAP2 automatically switches to an estimate of dominant frequency using mean frequency.

We first assessed the average dominant frequency for each syllable type per motif category for each recording. We then calculated the mean value for each syllable type in each motif category for the entire developmental period. The resulting mean dominant frequency values were used to calculate frequency shifts that occurred within adjacent motif pairs. The frequency change before the female wing stroke is the difference between "before" and "eliciting" motifs ("before wing stroke"), and the frequency change after the female wing stroke is the difference between the "eliciting" and "after" motifs ("after wing stroke"). We also calculated the mean syllable dominant frequency change between control syllable pairs.

To determine if the juveniles' real-time changes in dominant frequency were related to their father's song, ten examples of the father's dominant motif were randomly selected from song recordings. The dominant frequency of each selected syllable in the father's song was determined using the Explore & Score feature's auto-segmentation tool of SAP2.

Statistical analysis

Because contingent feedback has been shown to affect target syllables in real-time without affecting adjacent syllables (Tumer & Brainard, 2007), syllables were analyzed as independent units. In order to determine whether syllables showed predictable pitch changes, separate Wilcoxon sign-rank tests were used to calculate the accuracy of dominant frequency and the direction of frequency change. Accuracy of dominant frequency was higher when the juvenile's syllable frequency was closer to his father's after receiving a wing stroke. Direction of frequency change assessed whether the dominant frequency of the syllable in juvenile song shifted toward his father's before a wing stroke, after receiving a wing stroke, and in control songs. Only syllable types with data from more than one recording were used in the analysis.

Results

Description of behavioral observations

Singing activity of juvenile and adult males and behavioral activity of females calculated from the 15-sec time windows around song are summarized in Table 2.1. To assess what percentage of juvenile vocalizations elicited adult male responses, we calculated the proportion of juvenile song occurring immediately before the father's song compared to the total juvenile song produced. About 4 % (*M* = 4.26, *SD* = 2.86, min: 0.44 - max: 9.12) of juvenile songs received contingent feedback, which translated into *M* = 44, *SD* = 37.43 (min: 4, max: 102) songs for each juvenile over the entire 16 hours of observations per family. Juveniles sang about 72 songs per hour (*M* = 72.2, *SD* = 50.39). Thus, on average, 3 juvenile songs received contingent feedback within an hour (*M* = 2.89, *SD* = 2.53), or about every 51st song received feedback from the father.

To determine the adult male's responsiveness to juvenile song, we calculated the number of adult male songs produced during, after, and between juvenile songs compared to the total number of male songs produced. We also calculated the number of juvenile songs produced during, after, and between male songs compared to the total number of juvenile songs produced (Table 2.2). Distributions of the adult male singing activity relative to the juvenile song, and juvenile song relative to adult male song, are illustrated in Figure 2.3.

We also calculated average female responsiveness to adult and juvenile male songs. We assessed the average number of wing strokes performed during, after, and between adult and juvenile songs out of the total number of songs. We also calculated the number of female fluff-ups per adult and juvenile male song. On average, females responded with wing strokes twice as often to the song of their mate than to juvenile song. Fluff-ups were rare events compared to wing strokes. About every 14th adult male song and every 25th juvenile song elicited a fluff-up from the adult female (Table 2.3). The distributions of the female wing strokes and fluff-ups around juvenile song are illustrated in Figure 2.4.

Table 2.1 Behavioral activity during development, described as mean number of behaviors per bird over the entire recording period, and mean number of behaviors per hour.

	Activity over recording period		Activity per hour	
	Mean ± SD	min - max	Mean ± SD	min - max
Juvenile song activity	1147.07 ± 659.61	358 - 2682	72.2 ± 50.39	24.06 - 216.29
Adult male song activity	456.63 ± 283.5	251 - 1042	30.32 ± 23.58	13.49 - 84.03

Female WS	1716.88 ± 996.49	614 - 3287	108.05 ± 65.44	45.71 - 227.9
Female Fluff	139.5 ± 94.87	54 - 306	8.5 ± 5.28	4.35 - 17.44

Table 2.2 Percentage of total male song occurring during, after, and between juvenile song, and percentage of total juvenile song occurring during, after, and between adult male song

	Mean ± SD
Male song associated with juvenile song	10.4 ± 8.04
Juvenile song associated with male song	4.12 ± 3.08

Table 2.3 Mean number of total wing strokes and fluff-ups occurring during, after ,and between singing events, per adult and juvenile male song

	Mean ± SD
WS per adult song	1.02± 0.37
WS per juvenile song	0.51 ± 0.27
Fluff-ups per adult song	0.07 ± 0.05
Fluff-ups per juvenile song	0.04 ± 0.02

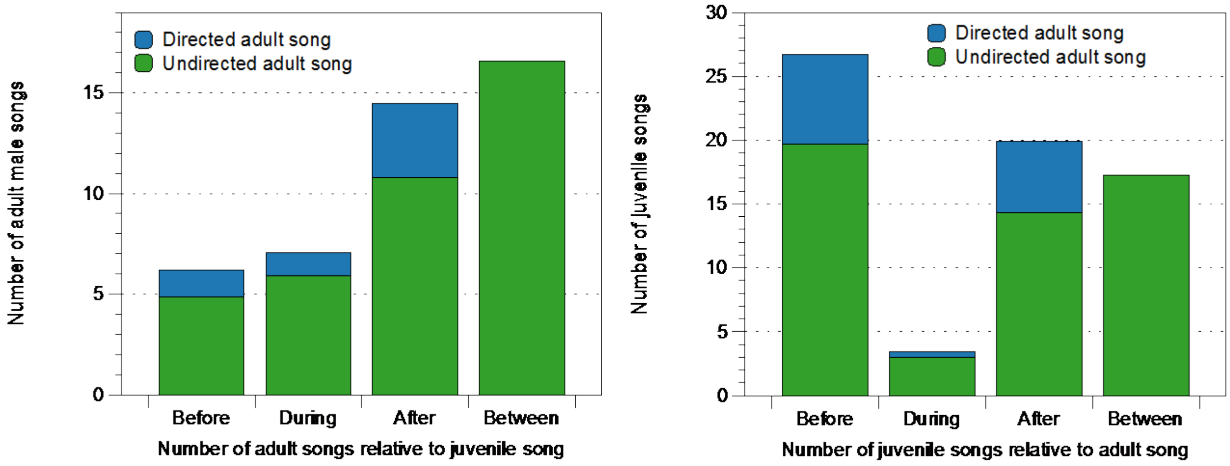

Figure 2.3 Mean number of adult male and juvenile male songs and their temporal relation. Adult male directed songs are labeled with blue. Number of directed songs was not assessed for the between time period**.** A: number of (directed or undirected) male songs occurring before, during, after, and between juvenile songs; B: number of juvenile songs occurring before, during, after, and between (directed and undirected) male songs.

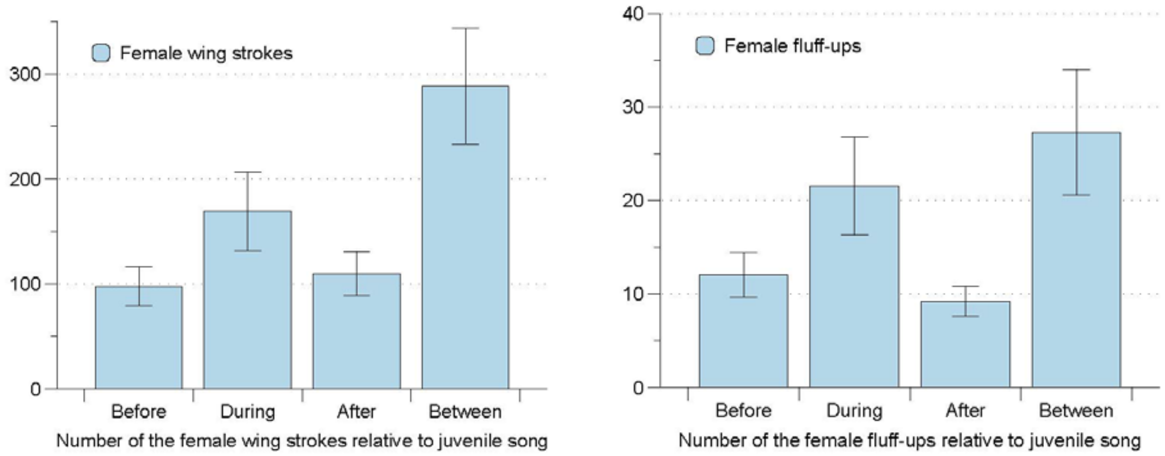

Figure 2.4 Distribution of female activity around juvenile song in the 15 seconds time window: A: number of wing strokes before, during, after, and between juvenile songs; B: number of fluff-ups before, during, after, and between juvenile songs.

Table 2.4 Song similarity between juveniles and fathers measured from crystallized song.

	Mean ± SD	*min – max*
Percent similarity to tutor	73.06 ± 11.16	49.44 – 95.34
Accuracy	75.3 ± 2.09	71.4 – 79.11
Sequential match	59.41 ± 14.53	42.24 – 91.53

Prediction of behavioral contingencies for song similarity measures

Juveniles copied the tutor song well (similarity measures are summarized in Table 2.4), as expected. In addition, behavioral interactions predicted the level of song learning.

To obtain the influence of behavioral interactions on song learning, we correlated behavioral contingencies occurring in the 15 seconds time windows with the three measures of song similarity obtained from the crystallized song. Correlations are summarized in Table 2.5. The results suggest that a higher incidence of directed songs by the father just after the juvenile male sings is associated with a song sequence in the later mature song of the juvenile that is closer to that of his father (r = 0.750, *p* = 0.018) (Figure 2.5A). Consistently, the more the juvenile male sings before directed adult male song, the more accurately they learn his song sequence (r = 0.711, *p* = 0.032) (Figure 2.5B).In contrast, a higher incidence of the father singing before the juvenile is associated with the juveniles coping the song less accurately (r = -0.661, *p* = 0.071) (Figure 2.5C). Female fluff-ups correlated with higher sequence similarity between juveniles and their fathers, regardless whether fluff-ups were before, during, after or between juvenile male song. The total number of female fluff ups around juvenile song was highly correlated with sequential similarity between fathers and juveniles (r = 0.873, *p* = 0.001) (Figure 2.5D).

Factor analysis

We conducted factor analysis to determine how many independent factors could be obtained from the behavioral interactions. Adult male song relative to juvenile, juvenile song relative to adult song, and female behaviors relative to juvenile song were included in three separate factor analyses. The resulting components are listed in Table 2.6.

A linear mixed model revealed that the components obtained from the factor analyses were related to song learning. The number of directed male songs around juvenile song predicted accuracy of learning, and female fluff ups predicted the sequential match between juveniles and fathers (Table 2.7).

To investigate the joint effects of adult male and female components, these variables were included in a new factor analysis. The analysis resulted in two components (Table 2.8). Their contributions to the three song outcome measures were analyzed with separate linear mixed models. The results confirmed that both directed male song and female fluff ups contribute to song learning (Table 2.9).

We conducted another multiple linear regression to assess the individual contribution of these behavioral variables to song learning. The results were similar to the findings of the correlations: juveniles who received more directed male song immediately after singing learned more similar songs to the tutor. When directed song was before juvenile song, juveniles copied less accurately. Juveniles who received more interruptions by directed male song also learned less similar songs to their fathers. Birds who experienced more female fluff ups immediately before singing learned a more accurate copy of the father song, and reached a higher sequential similarity, but female fluff ups between juvenile songs tended to decrease the accuracy score between juvenile and tutor (Table 2.10).

Table 2.5 Partial correlations between the number of contingent behaviors obtained from 15 sec time windows and song similarity measures to tutor (percent similarity, accuracy of learning and sequential match), with juvenile songrate/hour partialed out (MS - adult male song, JS - juvenile song). r values are listed when adjusted *p* < 0.1.

	PCT similarity	**Accuracy**	**Sequential match**
directed MS before JS	ns	r = -0.762, *p* = 0.068	ns
undirected MS before JS	ns	ns	ns
total MS before JS	ns	r = -0.661, *p* = 0.071	ns
directed MS during JS	ns	ns	ns
undirected MS during JS	ns	ns	ns
total MS during JS	ns	ns	ns
total MS between JS	ns	ns	ns
directed MS after JS	ns	ns	r = 0.750 , *p* = 0.018
undirected MS after JS	ns	ns	ns
Total MS after JS	ns	ns	ns
JS before directed MS	ns	ns	r = 0.711, *p* = 0.032
JS before undirected MS	ns	ns	ns
JS before total MS	ns	ns	ns
JS during directed MS	ns	ns	ns
JS during undirected MS	ns	ns	ns
JS during total MS	ns	ns	ns
JS after directed MS	ns	r = -0.648, *p* = 0.075	ns
JS after undirected MS	ns	ns	ns
JS after total MS	ns	r = -0.662, *p* = 0.071	ns
JS between MS	ns	r = -0.682, *p* = 0.071	ns
WS before JS	ns	ns	ns

WS during JS	ns	ns	ns
WS after JS	ns	ns	ns
WS between JS	ns	ns	ns
Fluff-ups before JS	ns	ns	r = 0.917, *p* < 0.001
Fluff-ups during JS	ns	ns	r = 0.797, *p* = 0.007
Fluff-ups after JS	ns	ns	r =0 .688, *p* = 0.042
Fluff-ups between JS	ns	ns	r =0 .878, *p* = 0.001
Total Fluff-ups around JS	ns	ns	r =0 .873, *p* = 0.001

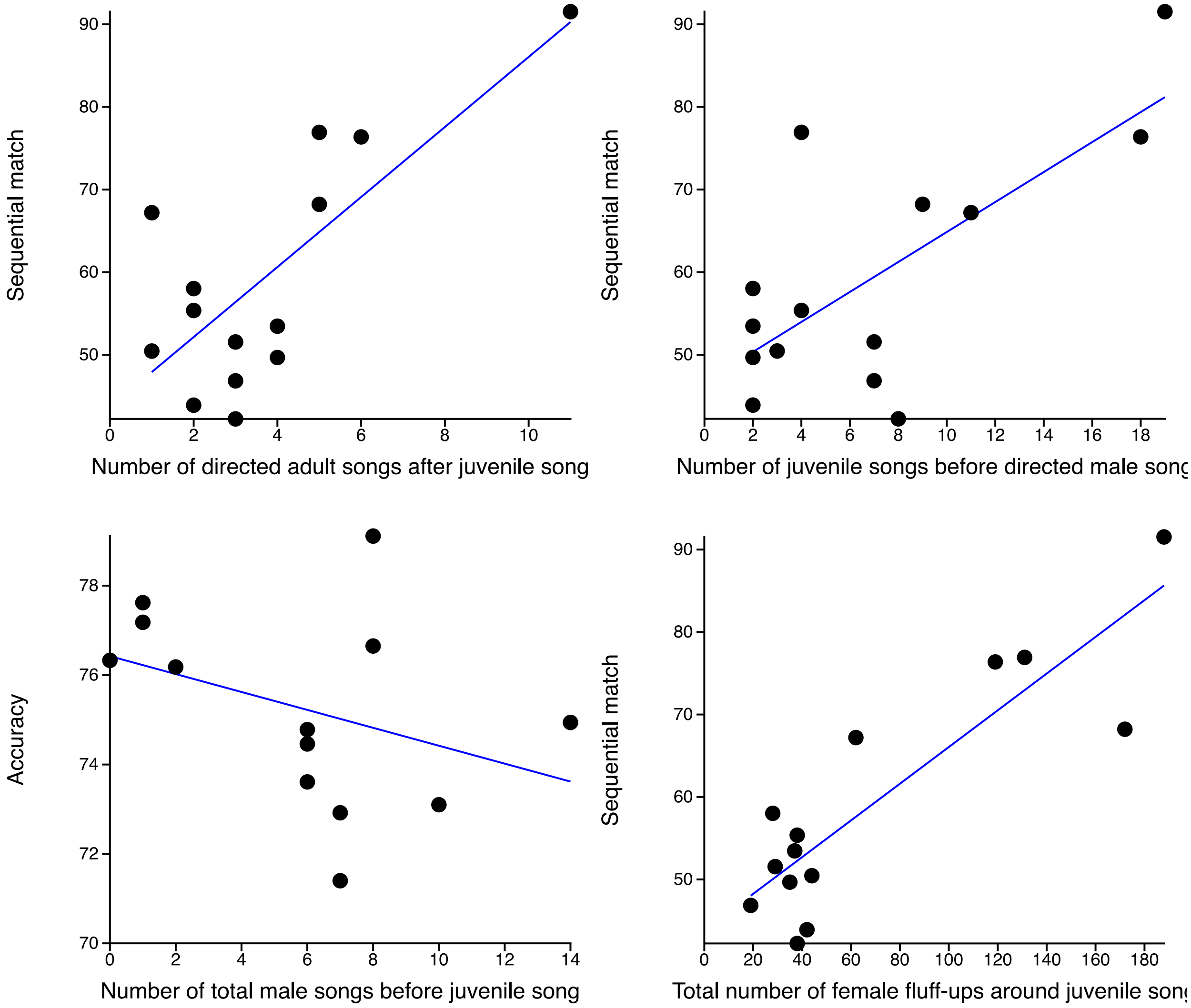

Figure 2.5 The number of behavioral interactions during development is related to song similarity measures between juvenile and tutor assessed from the mature song. A: the number of directed adult male song after juvenile song predicts sequential match between juvenile and tutor; B: the number of juvenile songs before directed adult male song predicts sequential match between juvenile and tutor; C: the number of total male songs before juvenile song tend to predict accuracy of learning D: the total number of female fluff-ups around juvenile song predict sequential match between juvenile and tutor.

Table 2.6 Components derived from the three separate factor analyses on male song (MS) relative to juvenile song (JS), juvenile song relative to male song, female wing strokes and fluff ups relative to the juvenile song. Behaviors with the strongest contribution to each component are marked in boldface.

Factor analysis #1: MS relative to JS	Components derived	
	1	2
Number of Undirected MS before JS	**0.904**	-0.139
Number of Undirected MS during JS	**0.809**	0.033
Number of MS between JS	**0.804**	**0.461**
Number of Undirected MS after JS	**0.72**	0.106
Number of Directed MS during JS	0.057	**0.903**
Number of Directed MS before JS	-0.1	**0.781**
Number of Directed MS after JS	**0.442**	**0.679**
% of variance	40.62	30.44

Factor analysis #2: JS relative to MS	Components derived		
	1	2	3
Number of JS after undirected MS	**0.937**	0.117	0.116
Number of JS between total MS	**0.877**	0.13	**0.388**
Number of JS before undirected MS	**0.828**	**0.505**	0.085
Number of JS after directed MS	0.068	**0.884**	0.225
Number of JS before directed MS	**0.378**	**0.825**	-0.055
Number of JS during undirected MS	0.188	-0.062	**0.927**
Number of JS during directed MS	0.18	**0.452**	**0.749**
% of variance	36.4	27.93	23.51

Factor analysis #3: WS and Fluff-ups relative to JS	Components derived	
	1	2
Number of WS between JS	**0.964**	0.12
Number of WS before JS	**0.959**	0.243
Number of WS during JS	**0.949**	0.255
Number of WS after JS	**0.934**	0.239
Number of Fluff –ups between JS	0.248	**0.942**
Number of Fluff –ups before JS	0.188	**0.898**
Number of Fluff –ups after JS	0.069	**0.896**
Number of Fluff –ups during JS	**0.404**	**0.884**
% of variance	48.59	43.44

Table 2.7 Results of linear mixed models investigating individual effects of male, female, and juvenile behavior on song similarity measures using factors (from Table 2.6, song rate/hour as fixed factors and family as a random factor. F-values are listed when $p < 0.1$.

	Percent similarity	Accuracy	Sequential match
MS relative to JS:			
Component 1	ns	F(1,10) = 3.78 $p = 0.081$	ns
Component 2	ns	F(1,10) = 4.82 $p = 0.053$	ns
Songrate/hour	ns	F(1,10) = 3.91 $p = 0.074$	ns
JS relative to MS			
Component 1	ns	ns	ns
Component 2	ns	ns	ns
Component 3	ns	ns	ns
Songrate/hour	ns	ns	ns
WS and Fluff-ups relative to JS			
Component 1	ns	ns	ns
Component 2	ns	ns	F(1,10) = 33.71 $p < 0.001$
Songrate/hour	ns	ns	ns

Table 2.8 Components derived from a factor analysis combining maternal and paternal behaviors that predicted the song learning of juvenile males.

	Components derived	
	1	2
Number of Fluff-ups between JS	**0.967**	0.148
Number of Fluff-ups during JS	**0.962**	0.077
Number of Fluff-ups before	**0.871**	0.218
Number of Fluff-ups after	**0.849**	0.174
Number of Directed MS after JS	**0.772**	0.495
Number of Directed MS during JS	0.218	**0.857**
Number of Directed MS before JS	-0.01	**0.806**
Number of MS between JS	0.291	**0.616**
% of variance	50.84	26.43

Table 2.9 Results of linear mixed models testing the influence of joint paternal and maternal behaviors on juvenile song (components of factor analysis from Table 2.8). F-values are listed when $p < 0.1$.

	Percent similarity	**Accuracy**	**Sequential match**
Component 1	F(1,6.21) = 4.13 p = 0.087	ns	F(1,10) = 31.66 **$p < 0.001$**
Component 2	ns	F(1,7.87) = 9.98 **$p = 0.014$**	ns
Songrate/hour	ns	ns	ns

Table 2.10 Models obtained from multiple linear regressions testing the joint effects of directed male song and female fluff-ups (behaviors contributing most heavily to components derived from the factor analysis, investigating joint paternal and maternal effects) on juvenile song learning.

Percent similarity	**B**	**SE B**	**β**	**Sig.**	**95%CI**	
					Lower Bound	Upper Bound
(Constant)	66.578	3.788		0.000	58.241	74.915
Directed MS after JS	3.579	1.007	0.829	0.005	1.363	5.795
Directed MS during JS	-5.957	1.850	-0.751	0.008	-10.030	-1.885

Accuracy	**B**	**SE B**	**β**	**Sig.**	**95%CI**	
					Lower Bound	Upper Bound
(Constant)	77.033	0.605		0.000	75.685	78.382
Directed MS before JS	-1.579	0.293	-0.867	0.000	-2.232	-0.926
Fluff before JS	0.191	0.080	0.814	0.038	0.013	0.370
Fluff between JS	-0.070	0.028	-0.835	0.031	-0.132	-0.008

Sequential match	**B**	**SE B**	**β**	**Sig.**	**95%CI**	
					Lower Bound	Upper Bound
(Constant)	41.514	2.909		0.000	35.176	47.851
Fluff before JS	1.482	0.196	0.909	0.000	1.055	1.910

Real-time influence of wing strokes on song development

Twelve juveniles from 7 families produced 58 syllable types in plastic song that could be identified as a version of the same syllable in the father's song. These 58 syllables were included in subsequent analysis.

To assess the real-time contribution of female feedback on juvenile song learning, we measured the accuracy of syllable dominant frequency after receiving a wing stroke compared to a syllable from a song motif on which no wing stroke was given. A Wilcoxon sign-rank test revealed that female feedback affected the accuracy of the dominant frequency of the juvenile syllables: the dominant frequency of syllables was significantly closer to the father's after receiving a female wing stroke than when no wing stroke was given, z = -2.67, *p* = 0.008 (Figure 2.6).

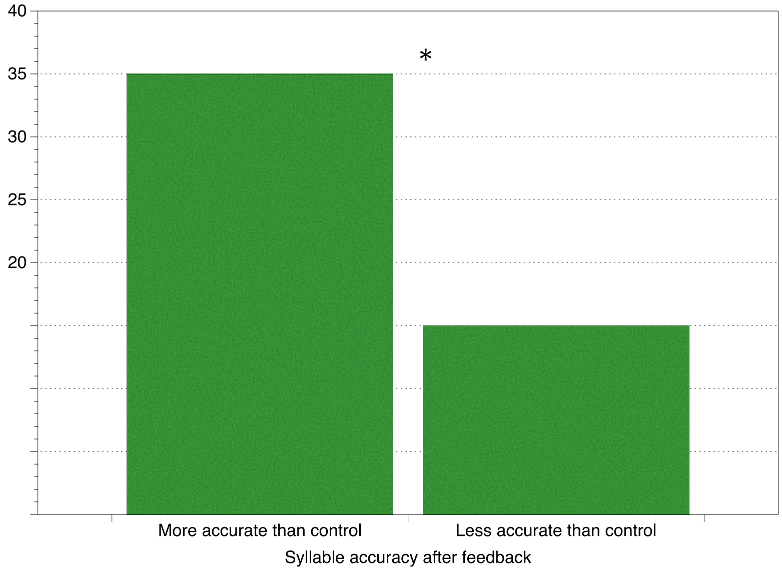

Figure 2.6 More syllables are closer to the father's dominant frequency after female wing strokes compared to control songs.

Female feedback also influenced the direction of change in dominant frequency: juveniles exhibited significant frequency shifts toward the father's syllables during development both before (Figure 2.7.) and after (Figure 2.8.) receiving a wing stroke, but not in control songs (before wing strokes: z = -2.143, *p* = 0.032; after wing strokes: z = -2.81, *p* = 0.005; between control song motifs: z = -0.742, *p* = 0.458).

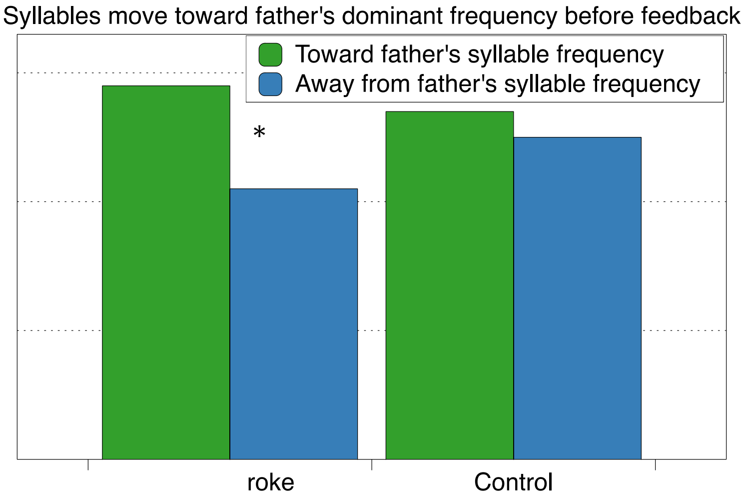

Figure 2.7 Dominant frequency in juvenile syllables moves toward father’s before female wing strokes, but none of these changes are present in control song regions.

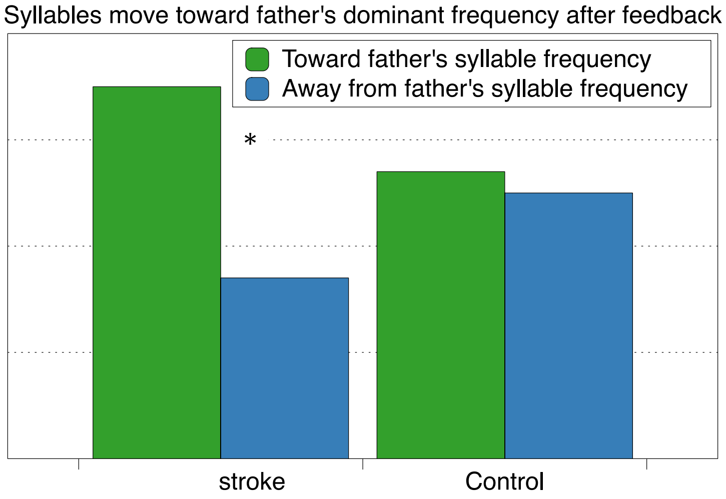

Figure 2.8 Dominant frequency in juvenile syllables moves toward father’s after female wing strokes, but none of these changes are present in control song regions.

Discussion

Inspired by demonstrations of socially guided learning in brown-headed cowbirds (West & King, 1988), we investigated the type and amount of social feedback occurring around subsong and plastic song of juvenile zebra finches by exploring the microstructure of social interactions. We found that contingent interactions with parents have the potential to shape the song learning process and influence the final song outcome. Both paternal and maternal activity around juvenile song predicted fidelity of learning. Juveniles whose immature song received more contingent directed song from their fathers learned more similar copies of their father's song, and learned his sequence more correctly. Conversely, increased juvenile singing after tutor song predicted less accurate learning.

Female-directed adult song was an especially powerful signal, as juvenile learning outcome was predicted by directed tutor song rather than undirected singing. Maternal feedback also influenced song learning, as juveniles experiencing more fluff-ups around and especially before their singing learned the fathers' song sequence more accurately.

Contingent paternal and maternal stimuli occurred infrequently. About 4 % of juvenile song elicited feedback from adult males, and directed father song produced contingently on juvenile song accounted for less than 2% of total adult male song. About every second juvenile song elicited a wing stroke (7 % of songs) and about every 25th song (4 % of songs) elicited a fluff up. Our results corroborate the importance of microanalytic approaches to quantifying the form and timing of behavior. Analyzing the microstructure of social interaction revealed the organization of behavior around immature singing by juveniles, even when the response behavior was rare. Wing strokes in cowbird females are extremely rare behaviors, as on average 1.1 wing strokes occur to every 100 juvenile songs (West & King, 1988). Nevertheless, wing strokes robustly shape juvenile song, and change its content as well as its developmental trajectory (Smith, King &West, 2000).

How could infrequent parental responses gain strong predictive value during learning? Observation of contingent interactions when other juveniles receive parental feedback may promote song acquisition and perceptual learning in the zebra finch. There is evidence from other avian species to support this idea. Following Pepperberg’s social observation theory, witnessing social interactions among others may facilitate vocal development (Pepperberg, 1985). Female brown-headed cowbirds eavesdrop on interactions related to other females and use the information available in the social group for preference formation (Gros-Louis, White, King, & West, 2003), and attend to other female’s vocalizations to modify their preferences (Freed-Brown & White, 2009).

Parental responsiveness may vary over developmental time, such that there may be periods of more frequent feedback. We have assessed the overall impact of paternal feedback on the final song outcome of the juvenile males. To uncover development in progress, we need to observe changes in juvenile singing activity and in the associated parental feedback over time. Moreover, combined paternal and maternal influences may further increase each other’s influence on song learning.

Examining social interactions at finer time-scales further supported our hypotheses regarding the role of socially guided learning in the zebra finch. Assessing real-time changes in dominant frequency in plastic song revealed that juvenile syllable frequency moved toward the father’s after wing strokes compared to instances of juvenile song that did not elicit wing strokes. The dominant frequency is a learned and precisely controlled feature of the song (Tchernichovski et al., 2001; Kao, Doupe & Brainard, 2005). Sensory feedback can rapidly modulate the fundamental frequency of individual syllables in the song of adult Bengalese finch. Negative reinforcement in the form of a short burst of white noise delivered contingently on the fundamental frequency of a syllable induced adaptive frequency changes of targeted syllables (Tumer & Brainard, 2007). In the present study, the subjects shifted the dominant syllable frequency in the correct direction to approximate father's frequency when female wing strokes occurred, which indicates that female feedback is likely to serve as reinforcement for the development of this song feature. Our results suggest that juveniles can evaluate their song production based on contingent social feedback consisting of auditory behavior from adult males as well as visual behavior from adult females.

Female-directed song by the adult male had a stronger impact on juvenile song outcome compared to undirected song, despite being produced about three times less frequently in conjunction with immature song. Courtship song may be a more salient communication signal than undirected singing for a number of reasons: it is frequently accompanied by dynamic choreography aimed at the female and is delivered with a faster tempo and less spectral variability than undirected song. Moreover, the song of adult males elicited twice as many visual signals from adult females than were elicited by juvenile song.

There is a striking similarity in the courtship dance of fathers and their juveniles, suggesting that juvenile attention is drawn to the father's courtship display (Williams, 2001). The adult male's dance choreography and song are correlated (Williams, 2001), and this redundancy in perceptual experience may facilitate attention. Moreover, males change the vocal quality of their song when it is produced toward females, which may reflect a "performance" state compared to undirected song, that may function as motor exploration (Jarvis, Scharff, Grossman, Ramos, & Nottebohm, 1998). When exposed to auditory information, females are able to perceive subtle context-dependent differences in male song, which is also reflected in immediate early gene expression in auditory telencephalic areas, and show a strong preference for directed singing (Woolley & Doupe, 2008). Taken together, simultaneous exposure to multisensory information (song of the adult male, possibly coupled with visual signals from the adult female) may organize juvenile perception and enhance learning via increased attention or arousal. Exposure to concurrent auditory and visual stimuli has been shown to enhance learning in nightingales, as juveniles exposed to stroboscope flashes during tutoring developed both larger repertoires and produced better copies of the song model than controls exposed only to the tutor songs (Hultsch, Schleuss, & Todt, 1999). Studies on human infants suggest a similar mechanism at work. Infants aged 7.5 months use visual information that is synchronized with the speech stream to aid perceptual segmentation (Hollich, Newman & Jusczyk, 2005).

Why do juveniles learn from contingent feedback provided by the father? Several possible mechanisms may explain the learning outcome: the act of singing may alter the inner state of the singer, or contingent feedback may promote preferential retention of particular repertoire elements (*action based learning,* Nelson & Marler, 1994). Because singing behavior is regulated by brain areas involved in vocal behavior as well as those involved in social behavior, the act of singing may be a result of altered physiological state or, alternatively, may alter the inner state of the singer. The precise role of these brain areas in song depends on social context. Directed singing is characterized by altered neural activity in the anterior forebrain pathway (AFP, basal ganglia-forebrain pathway) (Jarvis, Scharff, Grossman, Ramos & Nottebohm, 1998; Hessler & Doupe, 1999), and higher levels of dopamine in the Area X (basal ganglia, striatum) of the AFP (Sasaki, Sotnikova, Gainetdinov & Jarvis, 2006). Dopamine functions in the pursuit of social reward (for review see Riters, 2012). Unfortunately, the current knowledge about the role of dopamine related reinforcement during song development is limited, therefore we do not know whether it functions similarly in juveniles as in adults. Moreover, juvenile birds sing directed song infrequently during vocal development (Zann, 1996). In our sample, it was produced on average twice by each juvenile within the 15 sec time window around the male song during the entire observation period.

Feedback after singing may coincide with a physiological state in which juveniles are more aroused due to self-stimulation. A self-stimulating mechanism has been described in the ring dove, where the female's own "nest coo" vocalizations trigger an endocrine cascade leading to egg-laying (Cheng, 1992, 2003). Undirected singing in adult songbirds may be regulated by opioids, which would indicate that the act of singing is intrinsically rewarding (Khurshid, Hameed, Mohanasundaram & Iyengar, 2010; Riters et al., 2005) Confirmation of the role of opioids in developing juveniles awaits further research.

Alternatively, according to the action based learning model, juveniles may adjust their repertoire depending on vocal interactions and keep elements that elicit reactions from the father. Juveniles removed from their father at the age of 35 days prefer to learn from a tutor who shares high proportion of elements with their father (Clayton, 1987a), which indicates that action based learning may be present in this species. The influence of vocal interactions on real-time changes in juvenile song needs further investigation in the zebra finch.

Our results offer an alternative to the dominant model of song learning, which describes song memorization as an instructive phase, followed by a later practice phase based on recalled memories of the tutor song (Konishi, 1965, Marler, 1976). In territorial songbirds, the two phases are temporally distinct. Immature vocalizations gain significance when compared to vocalizations of territorial neighbors. Dissimilar vocalizations are pruned from the juvenile's repertoire (Nelson & Marler, 1994). Not only vocal interactions, but also passive exposure to or "eavesdropping" on vocal duels between territorial neighbors facilitate form of learning (Beecher, Burt, O'Loghlen, Templeton & Campbell, 2007) .

In the zebra finch, juvenile vocalizations elicit the feedback of conspecifics, and come in response to actions of conspecifics. Therefore, developing individuals play an active role in the communication network. As the zebra finch is socially gregarious, close proximity to conspecifics allows for rapid interchanges involving multiple modalities, which may explain why passive tape tutoring results in impoverished learning compared to live interactions (Deregnaucort et al., in press). Moreover, zebra finches are motivated to interact socially, as conspecifics have positive valence in contrast to related territorial finch species (Goodson, Evans, Lindberg & Allen, 2005; Goodson & Thompson, 2010). Therefore, learning mechanisms described in the zebra finch may be a property of a social species. Thus, a single model of song acquisition may describe one kind of bird with a specific type of ecology, but likely will not generalize across species with alternative life histories (Beecher & Brenowitz, 2005). Consequently, the biology of a particular species and its specific social environment must always be considered to uncover mechanisms of song acquisition.

Additionally, current knowledge of song learning mechanisms has been influenced by experiments based on isolation rearing and tape tutoring. The isolation paradigm aims to control song exposure and reveal innate components of species typical song without interference from environmental components. However, nature and nurture cannot be separated, as stated by Lehrman: "The interaction out of which the organism develops is *not* one, as is so often said, between heredity and environment. It is between *organism* and environment! And the organism is different at each different stage of its development" (Lehrman, 1953, p.345). Isolate rearing reveals only what route development takes in a specific environment. Moreover, an impoverished social environment alters not only song production but also auditory perceptual processes (Cousillas et al., 2006). To make informed prediction about mechanisms at work, it is crucial to study developing individuals in a naturalistic social setting.

When studied in social context, both zebra finches and brown-headed cowbirds show evidence of socially guided learning. Why have these effects not been shown in other songbird species? The lessons of the last 30 years of research on cowbirds have not been generalized to research programs on other songbirds. As a result, the crucial role of female feedback on song learning has generally gone under-appreciated. Nevertheless, there are indications from several species that females influence the vocal outcome. When canaries were hybridized with chaffinches and juveniles were raised by their mothers, the young birds learned syllables based on the syntax of the maternal species (Guettinger, Wolffgramm & Thimm, 1978). Females in these species do not sing unless undergoing external hormonal manipulation. The authors concluded that syntax is genetically encoded. The alternative explanation is that females' reactions to the song of the juveniles shaped the syntax to their preference. In cowbirds, females from a local versus a distant population influence the rate of development and song potency differently (King, West & Goldstein, 2005), and individual females exhibit diverse responsiveness to high- vs. low-potency song (Hamilton, King, Sengelaub & West, 1997).

The evolved complex communication system based on multisensory integration requires underlying neural adaptations in both sexes. Preference for high-potency songs assessed on the basis of female copulatory responses correlates with the volume and neuron number of lMAN (lateral magnocellular nucleus of the anterior neostriatum) of the anterior-forebrain pathway (Hamilton et al., 1997). The function of lMAN is crucial during song learning and in modulating social context-sensitive adaptive plasticity in male songbirds (e.g. Scharff & Nottebohm, 1991; Jarvis et al., 1998; Hessler & Doupe, 1999; Kao & Brainard, 2006; Stepanek & Doupe, 2010). Female cowbirds do not sing, thus lMAN may function in perceptual processes (Hamilton et al., 1997). Furthermore, courtship success, measured as female-directed singing and song potency, correlates positively with the volume of a visual thalamic nucleus (nucleus rotundus) in male cowbirds (Hamilton, King, Sengelaub & West, 1998). Thus, specialized learning mechanisms have corresponding neural components in both sexes. The systems-based approach, in which learning mechanisms were characterized at the social as well as the neural levels of organization, made the cowbird research program particularly informative, and should inspire subsequent research on neural correlates of adaptive behaviors.

Socially guided learning is not equivalent to explicit, purposeful teaching. Wing strokes are believed to be precursors of the copulation-solicitation preference displays in cowbirds. Thus, juveniles are able to take advantage of female signals that evolved for a different function, a process called “inadvertent coaching” (Hoppitt & Laland, 2008). The developmental process bears striking similarity to vocal learning in human infants. Infant babbling elicits social response from caregivers (Goldstein, Schwade & Bornstein, 2009), and the social environment predicts individual differences in communicative development, as maternal responsiveness to babbling predicts vocal learning (Goldstein, Schwade & Vollmer, submitted). In a strong parallel with the songbird findings, contingent parental feedback can induce more developmentally-advanced vocalizations even when it consists only of non-vocal responses as touching, moving closer or smiling at the infants (Goldstein, et al., 2003), and learning of specific phonological patterns present in the caregiver’s feedback (Goldstein & Schwade, 2008).

There are limitations to our conclusions. In our study, families were kept in cages, therefore we cannot assess whether the same interactions are important during song acquisition in different social settings or in the wild. Subject density in aviary studies has generated different conclusions about tutor choice in the zebra finch (e.g. Williams, 1990; Mann & Slater, 1995). In addition, up to now we focused on the relation between female wing strokes and dominant frequency of juvenile syllables in real time, but we also need to observe the association between real time changes in juvenile song with female fluff-ups and contingent adult male song. Moreover, our results are based on observations, and correlation between behavioral interactions and song outcome does not equal causation. Our study is only the first step toward identifying social mechanisms of vocal development in the zebra finch. Experimental manipulation will be essential to separate and quantify the contribution of paternal and maternal interactions during song acquisition. By using a training paradigm, in which feedback is controlled, we have started to explore the significance of contingent vocal feedback during song learning in the zebra finch (see Chapter 3 of dissertation). Furthermore, although we evaluated female contributions to learning from visual gestures around juvenile song, we do not know the information value of her reaction for learning. To assess social signal-to-noise ratio in contingent visual gestures of adult females, thus the information value (or a signal strength) of a given behavior compared to baseline activity, we need to explore levels of female wing strokes and fluff-ups at times when these do not coincide with vocal activity of the males. Family identity accounted for a large proportion of variance in the data. It is possible that differences in the behavior of the mothers explain the observed effects.

Finally, male song plays a crucial role in mating (for review see Riebel, 2009). Thus we need to obtain potency of the final song outcome, and differences in song learning based on acoustic analysis need to be validated through a behavioral assay of female choice.

Our observational study revealed that vocal development involves contingent and salient multisensory feedback organized around juvenile vocalizations. At these times, the receptive state of the singer may make the feedback particularly influential. As such, our findings imply that developmental processes may be best understood when studied on a social systems level. In nature, juvenile songbirds do not exist in social isolation. Their altricial state requires the presence of caregivers. Thus young songbirds are born into a structured social environment; an essential developmental task is to extract information from it.

Our results support the existence of a dynamic learning system, in which senders and receivers alternate roles as they participate in multisensory information exchange. Depending on the natural ecology of the species, non-singing conspecifics contribute to the process. Our microanalytic approach permits the testing of predictions about the contribution of individual social behaviors to the developmental outcome. Thus, to reveal development in action, we took the bird outside of the usual laboratory environment – the isolation box – and placed it back in its social context.

REFERENCES

Adret, P. (1993). Operant-conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour, 46*(1), 149–59. Doi: 10.1006/anbe.1993.1170

Anderson, R. C. (2009). Operant conditioning and copulation solicitation display assays reveal a stable preference for local song by female swamp sparrows Melospiza georgiana. *Behavioral Ecology and Sociobiology, 64*(2), 215-223. Doi: 10.1007/s00265-009-0838-y

Baptista, F., Bell, D. A., & Trail, P. W. (1993). Song learning and production in the white-crowned sparrow - parallels with sexual imprinting. *Netherlands Journal of Zoology, 43*(1-2), 17-33.

Baptista, L. F., & Gaunt, S. L. L. (1997). Social interaction and vocal development in birds. *Social influences on vocal development*, 23-40. Doi: 10.1017/cbo9780511758843.003

Baptista, L. F., & Petrinovich, L. (1984). Social-interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Animal Behaviour, 32*(1), 172-181. Doi: 10.1016/s0003-3472(84)80335-8

Baptista, L. F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow - social-factors and sex-differences. *Animal Behaviour, 34,* 1359-1371. Doi: 10.1016/s0003-3472(86)80207-x

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20*(3), 143-149. doi: 10.1016/j.tree.2005.01.004

Beecher, M. D., Burt, J. M., O'Loghlen, A. L., Templeton, C. N., & Campbell, S. E. (2007). Bird song learning in an eavesdropping context. *Animal Behaviour, 73,* 929-935. doi: 10.1016/j.anbehav.2006.10.013

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Bohner, J. (1983). Song learning in the zebra finch (taeniopygia-guttata) - selectivity in the choice of a tutor and accuracy of song copies. *Animal Behaviour, 31*(1), 231-237. doi: 10.1016/s0003-3472(83)80193-6

Bolhuis, J. J., Van Mil, D. P., & Houx, B. B. (1999). Song learning with audiovisual compound stimuli in zebra finches. *Animal Behaviour*, *58*, 1285-1292.

Brainard, M. S., & Doupe, A. J. (2001). Postlearning consolidation of birdsong: Stabilizing effects of age and anterior forebrain lesions. *Journal of Neuroscience*, *21*(7), 2501-2517.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature, 417*(6886). doi: 10.1038/417351a

Chaiken, M., Bohner, J., & Marler, P. (1993). Song acquisition in european starlings, sturnus-vulgaris - a comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. *Animal Behaviour, 46*(6), 1079-1090. doi: 10.1006/anbe.1993.1298

Cheng, M. F. (1992). For whom does the female dove coo - a case for the role of vocal self-stimulation. *Animal Behaviour, 43*(6), 1035-1044. doi: 10.1016/s0003-3472(06)80016-3

Cheng, M. F. (2003). Vocal self-stimulation: From the ring dove story to emotion-based vocal communication. *Advances in the Study of Behavior, Vol 33, 33*, 309-353. doi: 10.1016/s0065-3454(03)33007-4

Clayton, N. S. (1987b). Song tutor choice in zebra finches. *Animal Behaviour, 35*(3), 714-721. doi: 10.1016/s0003-3472(87)80107-0

Cousillas, H., George, I., Mathelier, M., Richard, J.-P., Henry, L., & Hausberger, M. (2006). Social experience influences the development of a central auditory area. *Naturwissenschaften, 93*(12), 588-596. doi: 10.1007/s00114-006-0148-4

Derégnaucourt, S., Mitra, P. P., Feher, O., Pytte, C., & Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. *Nature, 433*(7027), 710-716. doi: 10.1038/nature03275

Derégnaucourt, S., Poirier, C., Kant, A. V. d., Linden, A. V. d., & Gahr, M. Comparisons of different methods to train a young zebra finch (Taeniopygia guttata) to learn a song. *Journal of Physiology-Paris*(0). doi: http://dx.doi.org/10.1016/j.jphysparis.2012.08.003

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*(1), 567-631. doi: 10.1146/annurev.neuro.22.1.567

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour, 37*, 507-508. doi: 10.1016/0003-3472(89)90097-3

Field, A. P. (2009) *Discovering statistics using SPSS*. Los Angeles: SAGE.

Freed-Brown, G., & White, D. J. (2009). Acoustic mate copying: female cowbirds attend to other females' vocalizations to modify their song preferences. *Proceedings of the Royal Society B-Biological Sciences, 276*(1671), 3319-3325. doi: 10.1098/rspb.2009.0580

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 8030-8035. doi: 10.1073/pnas.1332441100

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515-523. doi: 10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience.* Oxford University Press.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The Value of Vocalizing: Five-Month-Old Infants Associate Their Own Noncry Vocalizations With Responses From Caregivers. *Child Development, 80*(3), 636-644. doi: 10.1111/j.1467-8624.2009.01287.x

Goldstein, M. H., Schwade, J. A., & Vollmer, L. (submitted). Developmental changes in infants' learning about the social efficacy of their prelinguistic vocalizations.

Goodson, J. L., Evans, A. K., Lindberg, L., & Allen, C. D. (2005). Neuro-evolutionary patterning of sociality. *Proceedings of the Royal Society B-Biological Sciences, 272*(1560). doi: 10.1098/rspb.2004.2892

Goodson, J. L., & Thompson, R. R. (2010). Nonapeptide mechanisms of social cognition, behavior and species-specific social systems. *Current Opinion in Neurobiology, 20*(6), 784-794. doi: 10.1016/j.conb.2010.08.020

Gros-Louis, J., White, D. J., King, A. P., & West, M. J. (2003). Female brown-headed cowbirds' (Molothrus ater) social assortment changes in response to male song: a potential source of public information. *Behavioral Ecology and Sociobiology, 53*(3), 163-173. doi: 10.1007/s00265-002-0560-5

Guettinger, H. R., Wolffgramm, J., & Thimm, F. (1978). The relationship between species specific song programs and individual learning in song birds a study of individual variation in songs of canaries greenfinches and hybrids between the 2 species. *Behaviour, 65*(3-4), 241-262.

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1997). A brain of her own: A neural correlate of song assessment in a female songbird. *Neurobiology of Learning and Memory, 68*(3), 325-332. doi: 10.1006/nlme.1997.3781

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1998). Visual and song nuclei correlate with courtship skills in brown-headed cowbirds. *Animal Behaviour, 56,* 973-982. doi: 10.1006/anbe.1998.0848

Helekar, S. A., Marsh, S., Viswanath, N. S., & Rosenfield, D. B. (2000). Acoustic pattern variations in the female-directed birdsongs of a colony of laboratory-bred zebra finches. *Behavioural Processes, 49*(2), 99-110. doi: 10.1016/s0376-6357(00)00081-4

Hessler, N. A., & Doupe, A. J. (1999). Social context modulates singing-related neural activity in the songbird forebrain. *Nature Neuroscience, 2*(3), 209-211. doi: 10.1038/6306

Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development, 76*(3), 598-613. doi: 10.1111/j.1467-8624.2005.00866.x

Hoppitt, W., & Laland, K. N. (2008). Social processes influencing learning in animals: A review of the evidence. In H. J. Brockmann, T. J. Roper, M. Naguib, K. E. WynneEdwards, C. Barnard & J. C. Mitani (Eds.), *Advances in the Study of Behavior, Vol 38* (Vol. 38, pp. 105-165).

Houx, B. B., & ten Cate, C. (1998). Do contingencies with tutor behaviour influence song learning in zebra finches? *Behaviour, 135*, 599-614.

Houx, B. B., & ten Cate, C. (1999a). Do stimulus-stimulus contingencies affect song learning in zebra finches (Taeniopygia guttata)? *Journal of Comparative Psychology, 113*(3), 235-242.

Houx, B. B., & ten Cate, C. (1999b). Song learning from playback in zebra finches: is there an effect of operant contingency? *Animal Behaviour, 57*, 837-845.

Hultsch, H., Schleuss, F., & Todt, D. (1999). Auditory-visual stimulus airing enhances perceptual learning in a songbird. *Animal Behaviour, 58*(1), 143-149. doi: 10.1006/anbe.1999.1120

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In Hinde, R. A. Ed. *Bird Vocalizations*, Cambridge Univ. Press, London. Pp. 61–74.

Jarvis, E. D., Scharff, C., Grossman, M. R., Ramos, J. A., & Nottebohm, F. (1998). For whom the bird sings: Context-dependent gene expression. *Neuron, 21*(4), 775-788. doi: 10.1016/s0896-6273(00)80594-2

Johnson, F., Soderstrom, K., & Whitney, O. (2002). Quantifying song bout production during zebra finch sensory-motor learning suggests a sensitive period for vocal practice. *Behavioural Brain Research, 131*(1-2), 57-65. doi: 10.1016/s0166-4328(01)00374-6

Jones, A. E., & Slater, P. J. B. (1993). Do young male zebra finches prefer to learn songs that are familiar to females with which they are housed. *Animal Behaviour, 46*(3), 616-617.

Jones, A. E., & Slater, P. J. B. (1996). The role of aggression in song tutor choice in the zebra finch: Cause or effect? *Behaviour,133*, 103-115. doi: 10.1163/156853996x00053

Kao, M. H., & Brainard, M. S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *Journal of Neurophysiology, 96*(3), 1441-1455. doi: 10.1152/jn.01138.2005

Kao, M. H., Doupe, A. J., & Brainard, M. S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature, 433*(7026). doi: 10.1038/nature03127

Khurshid, N., Jayaprakash, N., Hameed, L. S., Mohanasundaram, S., & Iyengar, S. (2010). Opioid modulation of song in male zebra finches (Taenopygia guttata). *Behavioural Brain Research, 208*(2), 359-370. doi: 10.1016/j.bbr.2009.12.003

King, A. P., & West, M. J. (1977). Species identification in north-american cowbird - appropriate responses to abnormal song. *Science, 195*(4282), 1002-1004. doi: 10.1126/science.841321

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: Parallels to human speech development. *Ethology, 111*(1), 101-117. doi: 10.1111/j.1439-0310.2004.01039.x

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift fur Tierpsychologie, 22*(7), 770-783.

Lehrman, D. S. (1953). A critique of Lorenz, Konrad theory of instinctive behavior. *Quarterly Review of Biology, 28*(4), 337-363. doi: 10.1086/399858

Mann, N. I., & Slater, P. J. B. (1994). What causes young male zebra finches, Taeniopygia-guttata, to choose their father as song tutor. *Animal Behaviour, 47*(3), 671-677. doi: 10.1006/anbe.1994.1091

Mann, N. I., & Slater, P. J. B. (1995). Song tutor choice by zebra finches in aviaries. *Animal Behaviour, 49*(3), 811-820. doi: 10.1016/0003-3472(95)90054-3

Mann, N. I., Slater, P. J. B., Eales, L. A., & Richards, C. (1991). The influence of visual-stimuli on song tutor choice in the zebra finch, taeniopygia-guttata. *Animal Behaviour, 42*, 285-293. doi: 10.1016/s0003-3472(05)80560-3

Marler, P. (1970). A comparative approach to vocal learning - song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology, 71*(2p2), 1. doi: 10.1037/h0029144

Marler, P. (1976). Sensory templates in species specific behavior. *Fentress, John C. (Ed.). Simpler Networks and Behavior. X+403p. Illus. Sinauer Associates, Inc.: Sunderland, Mass., USA. Isbn 0-87893-182-1*, 314-329.

Marler, P. (1997). Three models of song learning: Evidence from behavior. *Journal of Neurobiology, 33*(5), 501-516. doi: 10.1002/(sici)1097-4695(19971105)33:5<501::aid-neu2>3.0.co;2-8

Marler, P., & Peters, S. (1981). Sparrows learn adult song and more from memory. *Science, 213*(4509), 780-782. doi: 10.1126/science.213.4509.780

Marler, P., & Peters, S. (1987). A sensitive period for song acquisition in the song sparrow, melospiza-melodia - a case of age-limited learning. *Ethology, 76*(2), 89-100.

Nelson, D. A. (1992). Song overproduction and selective attrition lead to song sharing in the field sparrow (spizella-pusilla). *Behavioral Ecology and Sociobiology, 30*(6), 415-424.

Nelson, D. A., & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences of the United States of America, 91*(22), 10498-10501. doi: 10.1073/pnas.91.22.10498

Nordby, J. C., Campbell, S. E., & Beecher, M. D. (2001). Late song learning in song sparrows. *Animal Behaviour, 61*, 835-846. doi: 10.1006/anbe.2000.1673

O'Loghlen, A. L., & Beecher, M. D. (1997). Sexual preferences for mate song types in female song sparrows. *Animal Behaviour, 53*, 835-841. doi: 10.1006/anbe.1996.0348

Pepperberg, I. M. (1985). Social modeling theory - a possible framework for understanding avian vocal learning. *Auk, 102*(4), 854-864.

Phan, M. L., Pytte, C. L., & Vicario, D. S. (2006). Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proceedings of the National Academy of Sciences of the United States of America, 103*(4), 1088-1093. doi: 10.1073/pna5.0510136103

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology, 93*(2), 260-277. doi: 10.1037/h0077553

Riebel, K. (2009). Song and Female Mate Choice in Zebra Finches: A Review. In M. Naguib, K. Zuberbuhler, N. S. Clayton & V. M. Janik (Eds.), *Advances in the Study of Behavior, Vol 40*, 197-238.

Riters, L. V. (2012). The role of motivation and reward neural systems in vocal communication in songbirds. *Frontiers in Neuroendocrinology, 33*(2), 194-209. doi: 10.1016/j.yfrne.2012.04.002

Riters, L. V., Schroeder, M. B., Auger, C. J., Eens, M., Pinxten, R., & Ball, G. F. (2005). Evidence for opioid involvement in the regulation of song production in male European starlings (Sturnus vulgaris). *Behavioral Neuroscience*, *119*(1), 245. doi: 10.1037/0735-7044.119.1.245

Roper, A., & Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology, 112*(5), 458-470. doi: 10.1111/j.1439-0310.2005.01169.x

Sasaki, A., Sotnikova, T. D., Gainetdinov, R. R., & Jarvis, E. D. (2006). Social context-dependent singing-regulated dopamine. *Journal of Neuroscience, 26*(35), 9010-9014. doi: 10.1523/jneurosci.1335-06.2006

Scharff, C., & Nottebohm, F. (1991). A comparative-study of the behavioral deficits following lesions of various parts of the zebra finch song system - implications for vocal learning. *Journal of Neuroscience, 11*(9), 2896-2913.

Searcy, W. A., & Marler, P. (1981). A test for responsiveness to song structure and programming in female sparrows. *Science, 213*(4510), 926-928. doi: 10.1126/science.213.4510.926

Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches (Taeniopygia guttata): progress and prospects. *Advances in the Study of Behavior, 18*,1.

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: the dynamics of toddler visual experience. *Developmental Science, 14*(1), 9-17. doi: 10.1111/j.1467-7687.2009.00947.x

Smith, V. A., King, A. P., & West, M. J. (2000). A role of her own: female cowbirds, Molothrus ater, influence the development and outcome of song learning. *Animal Behaviour, 60*, 599-609. doi: 10.1006/anbe.2000.1531

Stepanek, L., & Doupe, A. J. (2010). Activity in a Cortical-Basal Ganglia Circuit for Song Is Required for Social Context-Dependent Vocal Variability. *Journal of Neurophysiology, 104*(5), 2474-2486. doi: 10.1152/jn.00977.2009

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (1999). Note types, harmonic structure, and note order in the songs of zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology, 113*(2), 194-203. doi: 10.1037/0735-7036.113.2.194

Tchernichovski, O., Mitra, P. P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science, 291*(5513). doi: 10.1126/science.1058522

Tchernichovski, O., & Nottebohm, F. (1998). Social inhibition of song imitation among sibling male zebra finches. *Proceedings of the National Academy of Sciences of the United States of America, 95*(15), 8951-8956. doi: 10.1073/pnas.95.15.8951

Tchernichovski , O., Swigger, D., & Partha, M. (2004). Sound Analysis Pro User Manual. CCNY, New York.

ten Cate, C. (1991). Behavior-contingent exposure to taped song and zebra finch song learning. *Animal Behaviour, 42,* 857-859.

Thielcke, G. (1970). Learning of song as a possible pacemaker of evolution. *Zeitschrift fuer Zoologische Systematik und Evolutionsforschung, 8*(4), 309-320.

Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch Fringilla coelebs. *Ibis, 100*(4), 535-570. doi: 10.1111/j.1474-919X.1958.tb07960.x

Tumer, E. C., & Brainard, M. S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature, 450*(7173). doi: 10.1038/nature06390

Vallet, E., & Kreutzer, M. (1995). Female canaries are sexually responsive to special song phrases. *Animal Behaviour, 49*(6), 1603-1610. doi: 10.1016/0003-3472(95)90082-9

Vyas, A., Harding, C., Borg, L., & Bogdan, D. (2009). Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Hormones and Behavior, 55*(1), 50-59. doi: 10.1016/j.yhbeh.2008.08.005

Waser, M. S., & Marler, P. (1977). Song learning in canaries. *Journal of Comparative and Physiological Psychology, 91*(1), 1. doi: 10.1037/h0077299

West, M. J., & King, A. P. (1988). Female visual-displays affect the development of male song in the cowbird. *Nature, 334*(6179), 244-246. doi: 10.1038/334244a0

Williams, H. (1990). Models for song learning in the zebra finch - fathers or others. *Animal Behaviour, 39*(4),745-757. doi: 10.1016/s0003-3472(05)80386-0

Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (Taeniopygia guttata). *Journal of Experimental Biology, 204*(20)3497-3506.

Williams, H. (2004). Birdsong and singing behavior. In H. P. Zeigler & P. Marler (Eds.), *Behavioral Neurobiology of Birdsong* (Vol. 1016).

Woolley, S. C., & Doupe, A. J. (2008). Social context - induced song variation affects female behavior and gene expression. *Plos Biology, 6*(3), e62. doi: 10.1371/journal.pbio.0060062

Yu, C., Smith, L. B., Shen, H., Pereira, A. F., & Smith, T. (2009). Active Information Selection: Visual Attention Through the Hands. *Ieee Transactions on Autonomous Mental Development, 1*(2), 141-151. doi: 10.1109/tamd.2009.2031513

Zann, R. (1990). Song and call learning in wild zebra finches in south-east australia. *Animal Behaviour, 40,* 811-828. doi: 10.1016/s0003-3472(05)80982-0

Zann, R. A. (1996). *The Zebra Finch: Synthesis of Field and Laboratory Studies*. New York: Oxford University Press.

CHAPTER 3

JUVENILE ZEBRA FINCH LEARN THE UNDERLYING STATISTICAL REGULARITIES OF THEIR FATHERS' SONG

Abstract

Natural behaviors, such as foraging, tool use, social interaction, birdsong, and language, exhibit sequential, and in some cases hierarchical, structure. If the statistics of early experience reveal this structure, it is in principle learnable. We report that juvenile zebra finches learn sequential structure in song. The zebra finch is the most common species in which to study song learning. It is commonly held that young males acquire song by imitating tutors. Variability in the order of elements in mature song occurs, but is commonly dismissed as production error. Thus the degree to which zebra finch song may be characterized by statistical regularities has not been previously quantified. We hypothesized that variation in mature song may include non-random, learned structure. We used a novel approach, deriving graph-structured finite-state grammars from juveniles' songs, to show that variation in learned song can be described by syntax-like patterns. We found that juveniles learn statistically coherent grammatical structures from their tutors, drawing a new parallel between birdsong and language. Importantly, juveniles' grammars show higher structural similarity to those of their fathers than to those of unrelated individuals. Our graph-theoretic method does not assume syllable-level correspondence between individuals, and can be applied across species and contexts to analyze learned behaviors as distinct as foraging, tool use, and language.

Introduction

Developing individuals are surrounded by a highly structured environment. Successful functioning among conspecifics requires appropriate responses to actions of others, such as turn taking during human conversations (Sacks, Scheglof & Jefferso, 1974), or replying with a proper song type during avian territorial encounters (e.g. Beecher & Campbell, 2005). Thus learning the governing rules of such interactions are crucial. To use the structure provided by the rearing environment, developing cognitive systems must be able to recognize regularities and segment the continuous stream of information.

Mechanisms of human speech learning are frequently studied in the context of structure provided by the linguistic environment (Saffran, Aslin & Newport, 1996). Human speech, unlike written language, does not contain explicit physical information about word boundaries. However, children master the rules of spoken language with ease and within a short time. According to the poverty of the stimulus argument, insufficient information is provided within speech to permit language acquisition – de novo language learning is not possible, therefore humans must be born with an innate linguistic capacity (Chomsky 1975, Fodor, 1983). However, recent research on statistical learning of artificial grammars leads to different conclusions (Gomez & Gerken, 2000). Statistical learning refers to a process during which observers are able to extract probabilistic patterns from the environment. Regularities in the input permit acquisition of statistically coherent and possibly hierarchically structured units of perception and production. Spoken language has rich statistical structure, characterized by consistent, predictable co-occurrences and transitional probabilities among sounds in the speech stream (e.g. Goodsitt, Morgan & Kuhl, 1993). Transitional probabilities are high among syllables that belong to the same word, but low across word boundaries. For example, in case of the utterance "pretty baby”, the segment “pre” is frequently followed by “ty” in natural spoken English, therefore it has high transitional probability, but there is a small probability that ”ty” will be followed by “ba” within an English word. Human infants can segment a continuous stream of speech based on the underlying statistical regularities (Saffran et al., 1996). When exposed to structured input in the form of short sequences from an artificial language, infants extract abstract information analogous to syntax, then generalize that structure to novel strings in which the same grammatical rules are applied to new syllables (Gomez & Gerken, 1999). Moreover, mapping the statistical regularities of ambient speech plays a crucial role in the development of several aspects of language, including phonotactics, phonology, morphology, and syntax (Tomasello, 1998; Edelman & Waterfall, 2007; Goldstein & Schwade, 2008). Sensitivity to statistical regularities has been demonstrated for information presented outside of the speech domain (e.g. information in a visual stream, Kirkham, Slemmer & Johnson, 2002; non-speech sound patterns, Saffran, Johnson, Aslin & Newport, 1999), thus, statistical learning could be a powerful domain general tool.

Co-occurrence probabilities may also help non-human species to segment information. Animals segment available stimuli into smaller units. Pigeons are able to cluster discrete items into chunks during sequence learning tasks (Terrace, 1987), and common marmosets can learn the sequential order of visual stimuli (Koba, Takemoto, Miwa, & Nakamura, 2012). Moreover, many types of natural behaviors, such as foraging in rodents, tool use in corvids, social interaction in monkeys, and courtship-related vocalization in songbirds exhibit sequential, and in some cases hierarchical, structure (Lashley, 1951), which may be exploited to assist action segmentation and subsequent learning.

Songbirds are one of the more frequently studied systems of complex vocal learning. Numerous parallels between speech acquisition in humans and song learning in birds exist at multiple levels of organization, from neural to social (Brainard & Doupe, 2002). Songbirds, like human infants, learn to produce complex, hierarchically organized communicative signals within a short time interval. They develop in an acoustic environment rich in underlying structural regularities. Like in humans, the continuous stream of information presented in the vocal environment is often segmented into shorter sequences, or chunks, and chunks are frequently recombined from multiple tutors (e.g. Hughes, Nowicki, Searcy & Peters, 1998; Williams, 1990; Williams & Staples, 1992; Nelson, Marler & Palleroni, 1995; Takahashi, Yamada & Okanoya, 2010). Durations of silent intervals among song elements explain rules of song segmentation in some species (Hultsch & Todt, 2004). In other species, sequencing rules aid subsequent learning (Rose et al., 2004). Song syntax can be described by the set of probabilistic rules that define sequences across a finite repertoire of note clusters. In many Estrildid finch species, song is hierarchically constructed. Individual note types are combined into syllables, and syllable sequences are assembled to form song motifs. Finch species vary in the degree to which their song is characterized by probabilistic dependencies across syllables. For example, the song of the Bengalese finch is characterized by high levels of syntactic variability (Honda & Okanoya, 1999; Okanoya, 2004), and there is evidence that statistical regularities aid its song development (Takahashi et al., 2010).

To be able to compare statistical regularities in behaviors among multiple subjects of a non-human species, one needs computational tools capable of (i) detecting and describing the structure of behavior and (ii) comparing the results across individuals and groups. When used together, these tools can reveal common patterns, quantify individual differences, and, for acquired behaviors, help elucidate the mechanisms of learning. We used two such tools – a model of grammar acquisition that is being developed for natural language applications (Kolodny, Lotem & Edelman, 2011) and a family of procedures for quantifying graph similarity (Wilson & Zhu, 2008) – to study the grammar-like structure in the most frequently studied songbird species, the zebra finch.

The zebra finch is a close relative of the Bengalese finch. Song learning in the zebra finch is frequently described as imitation, however the learning mechanisms driving song development are not well understood. According to the standard view, when adults, zebra finch develop highly stereotyped songs, with a single "canonical" motif, which is the most frequently produced sequence of syllables (Zann, 1996, Brainard & Doupe, 2001) Acoustic variability is generally attributed to production error (Sturdy, Phillmore & Weisman, 1999). Zebra finch song has not been examined for the presence of probabilistic syntax-like patterns. However, individuals exhibit substantial diversity at the levels of syllable transitions (sequence linearity) and motif occurrences (sequence consistency) (Scharff & Nottebohm, 1991). One third of motifs are non-canonical, including ones with syllable deletions, additions or repetitions (Sturdy et al., 1999). Could song learning include acquiring underlying structural regularities in song variation? If so, the task for a young learner is to distill statistical regularities from tutor songs and incorporate these into his own production. Any such statistical regularities in zebra finch song would indicate a new parallel between avian song learning and human language.

Furthermore, the distinction between structural or grammar-like regularities and those pertaining to the individual units comprising the vocalization sequence ("lexical" regularities) has not, to our knowledge, been attempted in any non-human species. (To appreciate this distinction, consider the sentences "This bird can sing" and "That pig will fly," which are lexically disjoint but identical in their grammatical structure.) If present, structural regularities in song would dramatically change our view of the task of song learning, as probabilistic patterns shared by juveniles and their fathers would suggest a statistical learning mechanism and a complex, perhaps hierarchical, internal representation.

Our aim was to search for grammatical regularities in the full corpus of variation found across song renditions, both in temporal relations among syllables and in temporal relations among longer units, thus accounting for possible hierarchical structure. We then compared the regularities between fathers and their offspring across multiple families. If statistical learning plays an important role during song development, then statistically coherent patterns present in a tutor's song should be reflected in the song of his offspring. Thus, structural similarities should be more pronounced between fathers and sons from the same family than between males from different families.

Using techniques devised for analyzing structure in human languages, we examined zebra finch song for evidence of hierarchical statistical regularities in song motifs and compared the resulting 'grammars' across individuals. In grammar we refer to the set of syntactic rules and principles by which grammatical structures are created. Typically in such projects, one uses the corpus of song recorded from an individual singer to infer a grammar for that individual (e.g., e.g., Nishikawa & Okanoya, 2006; Jin, 2009; Jin & Kozhevnikov, 2010; Berwick, Okanoya, Beckers & Bolhuis, 2011). Although corpora can be pooled across individuals, such pooling assumes that the same basic lexicon of units (syllables) underlies song production in all the individuals — a problematic and untested assumption. In contrast, we describe a computational method that transcends this limitation, and makes no assumptions about the commensurability of the lexicons of different birds. It does so by quantifying graph similarity (graphs are network structures created on the basis of transitional probabilities between syllables, thus graphs are the finite-state grammars) — in ways that are purely structural and do not involve the labels (syllable symbols) that construct the nodes of the graphs (graph vertices).

We compared songs of fathers and sons from nine zebra finch families, each of which were housed together until adulthood. A syllable catalog was created for each of the 27 males by assigning a symbol to each syllable type. Samples of each bird's song were subsequently annotated using this catalog and processed so as to yield a graph-structured grammar. We described each male's song by three types of grammar (syllable-based, or SYL; syllable collocation-based, or COL; and one that allowed variation within collocations, or SLOT; see Methods). To assess similarity in the grammar structure of juveniles and fathers, the created grammars were compared pairwise. Every male was compared to every other, distinguishing within-family and between-family pairs of comparisons (SAME and DIFF). We tested three methods for comparing graphs to one another (see Methods), and chose to focus on a measure based on the eigenvalue spectrum of the grammar graph's adjacency matrix, which is closely correlated with graph edit distance (Wilson & Zhu, 2008).

Materials and Methods

Nine zebra finch pairs were set up in single cages with a nest box and nest material, in a community room in both visual and acoustic contact with each other and with birds in other aviaries. The birds were kept on a 14:10 light cycle and were provided with finch seed and water *ad libitum*. Each pair nested, laid eggs, hatched and fledged 2-5 young (mean = 4.0 ± 1.2). Families were kept together for more than 3 months (99 ± 9) days. The juveniles were then transferred to single sex aviaries with other juvenile birds.

Mature song recordings

Songs of all nine fathers were recorded. Mature female-directed song was recorded from15 male juveniles at the age of 108 ± 14 days. Songs from 3 additional individuals was recorded at an age of 142-153 days. For all recordings, males were placed in a sound proof room overnight in a 46*44*36 cm large cage. The following morning, an adult female zebra finch was placed in an identical cage next to the male's cage and recording began. If the male did not sing within 60 minutes, further recordings were attempted after a day back with the colony, until we obtained at least 10 song bouts in a single recording. All recordings used a Sennheiser shotgun microphone attached to a Canon MiniDV ZR930 camcorder on Fujifilm DVCassette miniDVs. MiniDV tapes were digitized with a JVC Super VHS ET Professional deck at 44.1 KHz. Uncompressed sound files were created using Soundtrack Pro 6 and were saved as separate wav files.

Song annotations

Based on Sossinka and Bohner's definition (Sossinka & Bohner, 1980), a song bout consists of introductory elements followed by one or more song motifs. Song bouts in our data set were defined as strings of syllables in which all silent intervals were shorter than 500 ms. The present research investigates probabilistic dependencies among syllables combined into stable sequences or motifs, therefore we assigned a letter to each syllable type in each individual's song (cf. Price, 1979, Eales, 1985) . Every song in the recordings was then broken down into these constituent syllables using Syrinx (John Burt, www.syrinxpc.com).

Inferring the grammar

For each individual’s song corpus, we derived three types of finite-state grammar using a variant of a biologically inspired model of language acquisition, BAGEL (Kolodny, Lotem & Edeleman, 2011).

All three types of grammar had the form of a probabilistic first-order Markov graph specifying the transition probabilities among basic units; the grammars differed in the kinds of units that formed the lexicon, and hence in their ability to capture hierarchical structure. The basic building block of the lexicon units in each case was the syllable. Grammars were based on syllables (SYL), collocations of syllables (COL), and collocations containing slots holding equivalent units (SLOT). For the first type of grammar, SYL, the units used in learning of the Markov model were the actual syllables. For the second type, COL, collocations of syllables forming longer units were allowed as well. We used the concept of collocation, borrowed from computational and corpus linguistics (e.g., Mel’ćuk, 1998, Croft, 2001, Arnon & Snider, 2010), to operationalize the idea of “dominant motif” found in the behavioral literature on birdsong analysis (e.g., Brainard & Doupe, 2001). Intuitively, a collocation is a sequence of basic units that recurs in the corpus in question more often than warranted by chance. From the normative computational standpoint, the proper criterion for finding significant collocations must be based on the Minimum Description Length principle (MDL) (Rissanen, 1987), which has been proposed as a method for grammar acquisition (Grunwald, 1994) and used with great effect for learning natural language morphology (Goldsmith, 2001). For the purposes of the present study, we approximated the MDL approach (which can be computationally problematic, see Adriaans and Vitanyi, 2007) by a heuristic greedy search procedure. Specifically, our model identified motifs with the longest recurring sequences of syllables in each song corpus. Only sequences that did not contain an inner repetition of more than two syllables and that did not end in a partial repetition of their own first syllable/s were added to the lexicon (e.g., not retained: “a b c d a b c d”, “a b c d a b”; retained: “a b c d”). Among these, only sequences that occurred more frequently than a certain threshold were added to the lexicon, which also included by default all single syllables. A range of different parameter values in the search for motifs led to similarly significant results. The third type of grammar that we considered, SLOT, allowed collocations to contain slots, populated with “filler” units (syllables or collocations) that the model deemed to be equivalent (interchangeable) in the context of the “matrix” or embedding collocation. This option adds complexity, including the possibility of self-embedding (recursion), to the learned grammar (as discussed, e.g., by Solan, Horn, Ruppin & Edelman, 2005). As with the other two types, SYL and COL, the SLOT grammar type has the form of a graph, allowing the structure of different individuals’ grammars to be compared using an appropriate measure of graph similarity.

To illustrate these concepts, we show, for one of the individuals (#423), the corpus, the lexicon (syllables + motifs) for the COL type grammar, and the actual grammar inferred by our model (Figure 3.4).

Estimating similarity of grammars

Multiple techniques exist for comparing graphs. The main constraint on the choice of graph similarity in the present case is the need to avoid using vertex labels (due to the possible incommensurability of individual lexicons). This rules out the use of obvious measures such as graph edit distance, in which the dissimilarity between two graphs is defined as the smallest number of vertex and edge deletions, insertions, and substitutions that transform one graph into the other. Among the remaining options, the most straightforward one is spectral graph distance, defined as the Euclidean distance between the lists of eigenvalues of the adjacency matrices of the two graphs (e.g., Wilson & Zhu, 2008). The two lists of eigenvalues are sorted in decreasing order; if one of them is shorter than the other (because the adjacency matrix has a lower rank), it is padded with zeros. Interestingly, this spectral distance, which we refer to as SpecAdj correlates very closely with the edit distance (Wilson & Zhu, 2008), while avoiding any use of vertex labels, and is therefore best suited a priori to the task of comparing song grammars, where deletions, insertions, and substitutions of elements are the most natural causes of song difference. We focused on this measure of similarity in the above-described exploration.

A related distance, which we refer to as SpecNormLap, is defined in terms of the eigenvalues of the normalized graph Laplacian, which is computed from the graph's adjacency matrix. Because this distance measure is known to afford a finer discrimination between similar graphs, we expected it to be less useful for the present purposes — quantifying song relatedness, not distinctions — than SpecAdj.

The third approach to defining graph distance that we explored, CNAFeat, is based on a family of graph features used in computational network analysis (CNA); the particular features we considered have been used for characterizing brain dynamics and are part of the Brain Connectivity Toolbox (BCT) (Rubinov & Sporns, 2010). Because of the diverse nature of these features, some of which are global (pertain to the entire graph) and others local (per-vertex), we employed the Mahalanobis distance, which weights individual dimensions by their variance. The composition of the graph feature vectors that we looked at is as follows (for definitions of each measure see Rubinov & Sporns, 2010): transitivity (global); clustering coefficient (per vertex); modularity index (global) and module membership (per vertex); betweenness centrality (per vertex); 3-vertex motif intensities for the 13 classical motifs (per vertex); 4-vertex motif intensities for the 199 classical motifs (per vertex).

Statistical analysis

Data from the 27 birds gave rise to 27*(27-1)/2 = 351 possible pairwise comparisons; of those, 31 pairs were defined as SAME-family (father-son or siblings) and the remaining 320 pairs as DIFFerent (unrelated) for the purposes of the analysis. Our dependent variable was the similarity between grammars. This gave rise to 3*3 = 9 cases: three types of grammar (SYL, COL, SLOT) times three types of graph distance (SpecAdj, SpecNormLap, CNAFeat).

To avoid relying on assumptions of normality, equal variance, etc., we employed a nonparametric test, the Kruskal-Wallis statistic, to estimate the significance of the difference between the similarity values in the SAME and DIFF conditions. This test was performed for each of the nine cases. In addition, for each of the nine cases, we conducted 31 Wilcoxon one-sample sign rank tests, each comparing the value of grammar similarity for one of the SAME pairs to the list of values of all 320 DIFF pairs. Each such test was performed with alpha = 0.05/31 = 0.0016, which incorporates the Bonferroni correction for multiple comparisons. Finally, with S denoting the number of those tests that came out as significant, we conducted a binomial test of the significance of having S successes out of 31 trials.

In addition to the analyses reported in the main text of the paper, we carried out similar analyses of the songs without introductory notes (cf. Tchernichovski & Nottebohm, 1998) using the same graph distance, SpecAdj, as selected before. For the SYL grammar type, which is constrained not to form larger units out of syllables, this distinction is irrelevant: because introductory notes sometimes appear in mid-song, they cannot be excluded without disrupting grammar learning. For the SLOT grammar type, the differences between SAME and DIFF conditions were not significant. However, for the COL grammar type, on which our analysis reported in the main text converged (Figures 3.1, 3.2 and graph in Figure 3.4), the outcome was similar to the results with introductory notes included: Kruskal-Wallis statistic $p < 0.078$; Bonferroni-corrected binomial tests $p < 0.035$, 21 out of 31.

Results

The three types of grammatical representations of song yielded the predicted statistical regularities across individuals. The mean grammar similarity was higher between related compared to unrelated males [SYL ($p < 0.01$), COL ($p < 0.038$), and SLOT ($p < 0.06$)]. The binomial tests for the number of Bonferroni-corrected significant pairwise outcomes were in two out of three cases consistent with this picture: for SYL, 19 out of 31 ($p = 0.14$); for COL, 23 out of 31 ($p < 0.0053$); for SLOT, 22 out of 31 ($p < 0.015$). The best discrimination between SAME and DIFF comparisons was obtained with the COL grammar (collocation-based motifs; no recursive slot structure allowed), using the SpecAdj similarity measure (spectral distance derived from the grammar graph’s adjacency matrix) (Figure 3.1, 3.2). We have also carried out these analyses on a set of data from which the introductory notes opening many of the song bouts were excluded (cf. Tchernichovski & Nottebohm, 1998), and obtained similar results (see Methods). Furthermore, as opposed to the syntactic (song structure) contrast just reported, there were no significant differences in syllable repertoire size between SAME and DIFF conditions. Thus, while the grammars of related birds were far from structurally identical, they were more alike than for unrelated individuals (Figure 3.3), an effect highlighted by the comparison method that correlates with the number of edits needed to transform one graph into another—a correspondence in pattern rather than in sound.

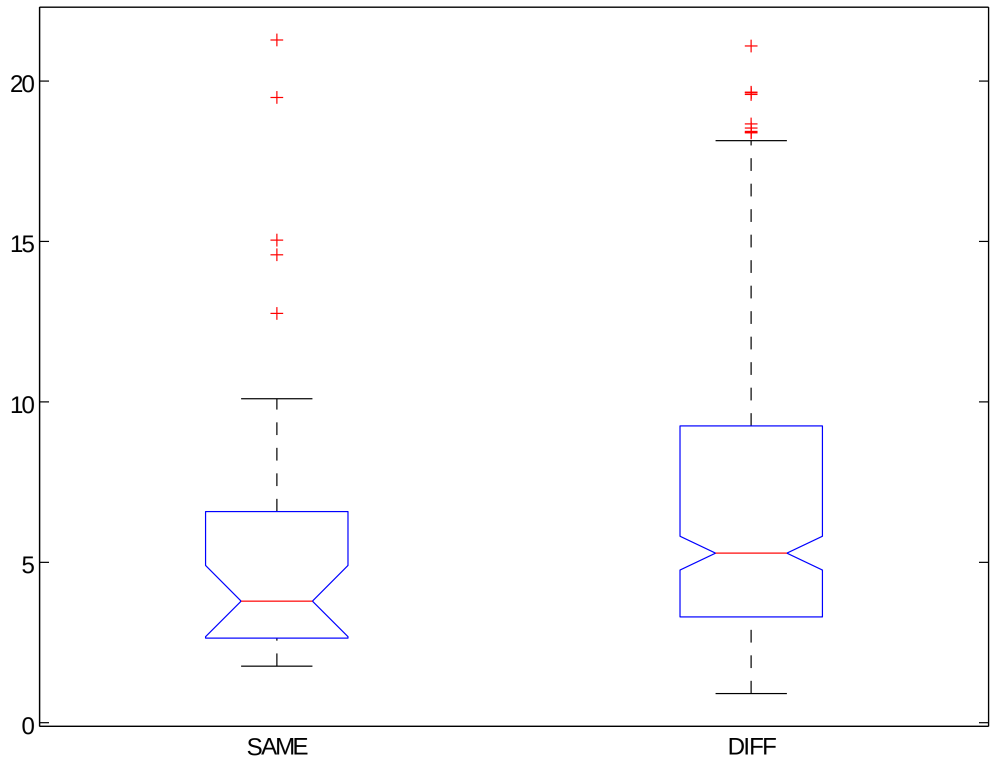

Figure 3.1 Similarity among pairs of song grammars. The distribution of grammar similarity values for the 31 pairs of related individuals (SAME family) and for the 320 pairs of unrelated ones (DIFFerent families), for the COL grammar type (based on syllables + motifs) and the SpecAdj grammar similarity measure (based on the eigenvalue spectrum of the adjacency matrix), showing medians, 1st and 3rd quartiles (box), limits of 1.5 times the inter-quartile range (whiskers) and outliers (+ symbols). The median of the SAME distribution is significantly lower than that of the DIFF distribution ($p < 0.038$, Kruskal-Wallis rank sum test), indicating greater similarity in songs of related individuals. Of the 31 pairwise similarity values for SAME birds, 23 were significantly lower than the median similarity value for DIFF birds (Wilcoxon sign rank test with Bonferroni-corrected alpha of 0.0016). A binomial test showed this pattern to be significant ($p < 0.005$).

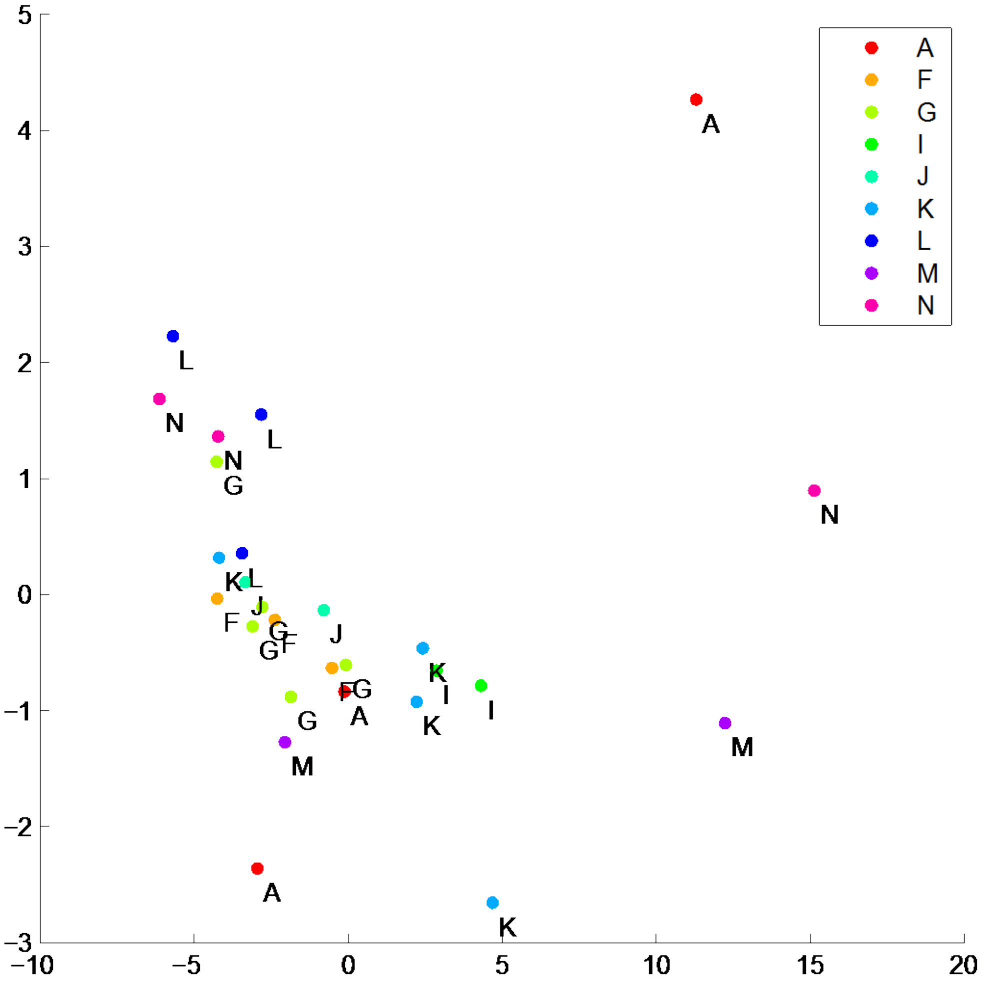

Figure 3.2 A 2D map of similarity relationships among grammars. A multidimensional scaling embedding of the 27 x 27 distance matrix into 2D, for the COL grammar type and SpecAdj graph similarity measure. We used the Matlab procedure *mdscale* with the squared stress function, which brings out the larger distances; the resulting stress was 0.024. Points corresponding to birds from the same family (coded by identical letters and colors) tend to cluster together.

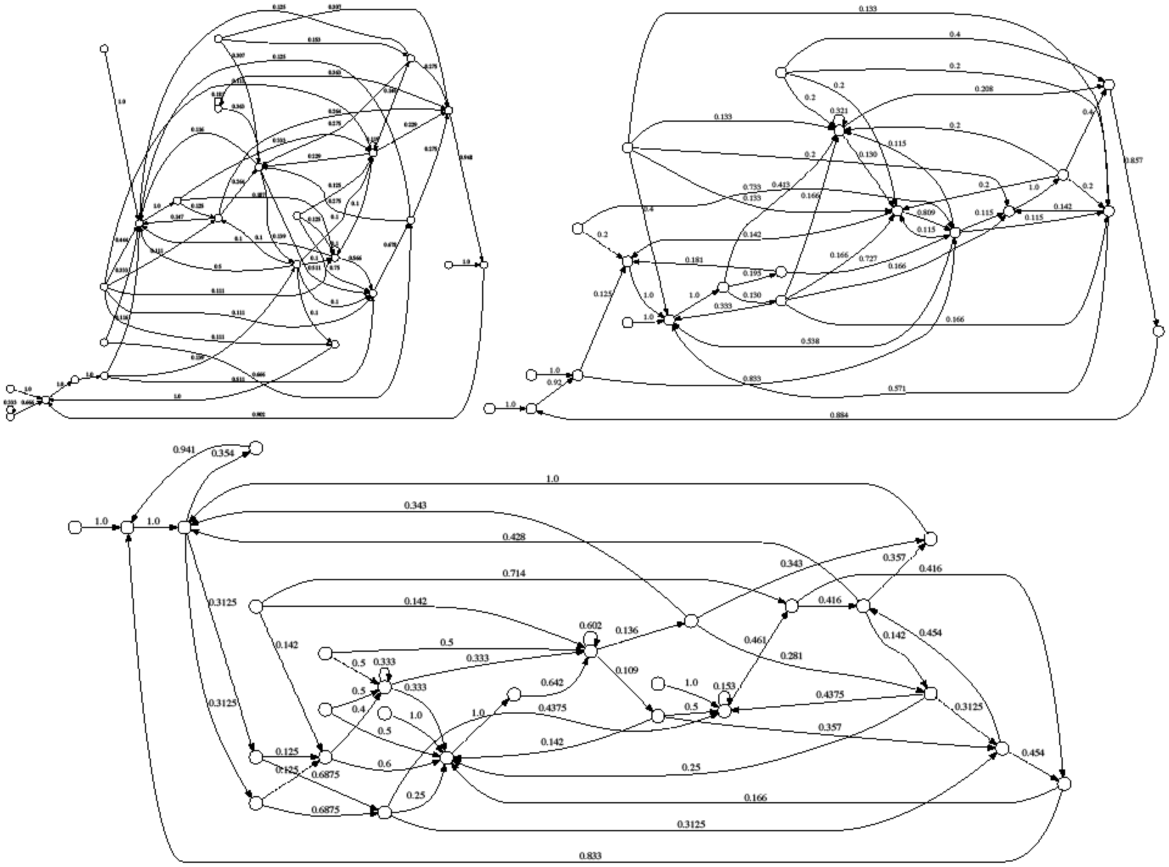

Figure 3.3 Examples of grammars of three individuals: these graphs describe the finite state grammar of type COL derived for two unrelated juveniles and the father of one of them. Clockwise from left top: father, son, unrelated juvenile. Transitions occurring at a probability less than 0.1 were omitted for the sake of clarity. Nodes appear without labels, as our analysis does not rely on syllable-level correspondence among individuals.

i i i i i a i i a b c f e
i a b c f d
i i i a b c f d
i i i a b b f d i a b c f d
i i i i i a b c f d i a b c f d
i i i a b c f d
i i i a b c f d i a b c f d
i i i a b c f d h g a b c f d
i i i i i a b c f d h g a b c f d
i i i a b c f e
i i a
i i i a b c f
i i i a b c f d
i i a b c f d h g i a b c f d
i i i a b c f e
i a b c f d
i i i a b c f d
i i a b c f d
i i a b c f d

BEGIN
END
a
g
b
c
d
e
i
i a b c f d
i i a b c f
i i a b c f d
i i i a b c f
i i i a b c f d
h
f

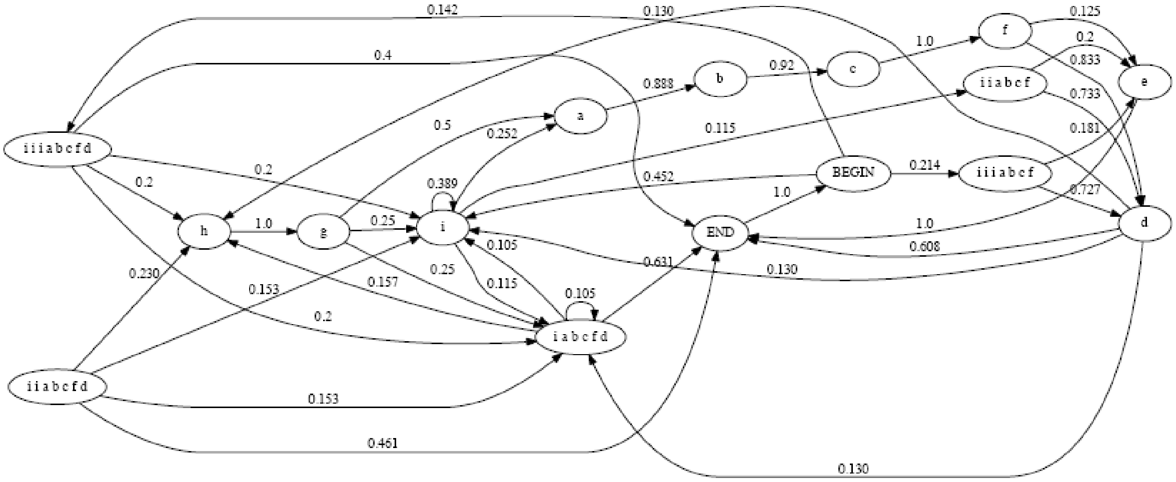

Figure3.4 Corpus, lexicon and grammar of a single individual (#423). Clockwise from left top: the entire *corpus*; the *lexicon* (syllables + motifs) for the COL type grammar (the BEGIN and END symbols are added by the grammar inference procedure); and the COL type *grammar* inferred by our model.

Discussion

The analysis of graph-based distances revealed that related birds sang similarly structured songs compared to birds from different families. We quantified song structure using graph-based methods that did not require a common labeling scheme of specific syllables across individuals. Thus just as humans learn patterns for using words in a language that are separate from the particular words used, juvenile males learn the syntactical structure of song from their fathers in addition to the form of notes and syllables.

We believe the present study is the first to compare finite-state grammars across multiple individuals in a songbird species. Past research on song learning has focused on the acoustic content of song and the extent of copying. With the zebra finch, several laboratories have tried to develop a universal classification system based on note shape, each using slightly different number and type of categories (e.g., Zann, 1993, Sturdy et al., 1999, Lachlan, Peters, Verhagen & ten Cate, 2010). However, even in the zebra finches' relatively simple song, developing a catalog for every song element across individuals has been difficult. Instead of trying to reconcile individual differences in song elements, our method allows direct comparison of the grammars of individuals without calling for a specific acoustic classification system. Furthermore, our method may be exceptionally well-suited for studying song in bird species with large song repertoires. In such species, song classification based on note types would be particularly difficult, and a grammar-based method correspondingly useful.

We find that song learning in the zebra finch is more than a process of mimicking the surface-level acoustic features of an adult male's song, as has been assumed in the past. In aviary settings, zebra finches learn hybrid songs composed by elements from multiple tutors, copied as chunks (Williams, 1990, Williams & Staples, 1992). Longer intersyllable durations, transitions between call- and non-call like elements, and locations of song-break in the tutor song mark boundaries of the copied chunks (Williams & Staples, 1992). Besides phonetic and segmentation cues, our results show that juveniles also learn the underlying statistical structure of song, separate from syllable-level correspondence between learner and tutor songs. This ability suggests that the birds use statistical learning mechanisms to map out the hierarchical organization of the tutor's song into an internally represented grammar of song production. Human adults and prelinguistic infants are sensitive to statistical regularities in segmentation tasks when learning artificial and natural languages (Gomez & Gerken, 2000; Saffran et al., 1996; Onnis, Waterfall & Edelman, 2008; Goldstein et al., 2010). Our present finding of such sensitivity to grammar in songbirds indicates a new parallel between song learning in birds and language learning in humans.

The kind of statistical learning mechanism that allows zebra finches to learn grammatical structure is applicable also to the general problem of learning structured, serially ordered behavior. For instance, statistical learning can be useful in learning the structure of foraging environments, food handling, tool use, and organizing one's behavior within a complex social group. Though these behaviors transcend domains, species, and scientific disciplines, the computational tools used here are capable of revealing their underlying grammatical structures and yield insight into the cognitive capacities necessary for learning adaptive skills.

REFERENCES

Adriaans, P., & Vitanyi, P. M. B. (2007). *The power and perils of MDL*. Nice, France.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67-82. doi: 10.1016/j.jml.2009.09.005

Beecher, M. D., & Campbell, S. E. (2005). The role of unshared songs in singing interactions between neighbouring song sparrows. *Animal Behaviour, 70,* 1297-1304. doi: 10.1016/j.anbehav.2005.03.008

Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences, 15*(3), 113-121. doi: 10.1016/j.tics.2011.01.002

Brainard, M. S., & Doupe, A. J. (2001). Postlearning consolidation of birdsong: Stabilizing effects of age and anterior forebrain lesions. *Journal of Neuroscience, 21*(7), 2501-2517.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature, 417*(6886), 351-358. doi: 10.1038/417351a

Chomsky, N. (1975b). *Reflections on Language*. New York: Pantheon Books.

Croft, W. (2001). *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Eales, L. A. (1985). Song learning in zebra finches - some effects of song model availability on what is learnt and when. *Animal Behaviour, 33*(4), 1293-1300. doi: 10.1016/s0003-3472(85)80189-5

Edelman, S., & Waterfall, H. (2007). Behavioral and computational aspects of language and its acquisition. *Physics of Life Reviews, 4*(4), 253-277. doi: 10.1016/j.plrev.2007.10.001

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational Linguistics, 27*(2), 153-198. doi: 10.1162/089120101750300490

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515-523. doi: 10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences, 14*(6), 249-258. doi: 10.1016/j.tics.2010.02.004

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109-135. doi: 10.1016/s0010-0277(99)00003-7

Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*(5), 178-186. doi: 10.1016/s1364-6613(00)01467-4

Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language, 20*(2), 229-252.

Grunwald, P. (1994). A minimum description length approach to grammar inference. In G. Scheler, S. Wernter & E. Riloff (Eds.), *Connectionist, statistical and symbolic approaches to learning for natural language* (Vol. 1004, pp. 203-216). Berlin: Springer Verlag.

Honda, E., & Okanoya, K. (1999). Acoustical and syntactical comparisons between songs of the white-backed munia (Lonchura striata) and its domesticated strain, the Bengalese finch (Lonchura striata var. domestica). *Zoological Science, 16*(2), 319-326. doi: 10.2108/zsj.16.319

Hughes, M., Nowicki, S., Searcy, W. A., & Peters, S. (1998). Song-type sharing in song sparrows: implications for repertoire function and song learning. *Behavioral Ecology and Sociobiology, 42*(6), 437-446. doi: 10.1007/s002650050458

Hultsch, H., & Todt, D. (2004). Approaches to the mechanisms of song memorization and singing provide evidence for a procedural memory. *Anais Da Academia Brasileira De Ciencias, 76*(2), 219-230.

Jin, D. Z. (2009). Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Physical Review E, 80*(5), 051902. doi: 10.1103/PhysRevE.80.051902

Jin, D. Z., & Kozhevnikov, A. A. (2011). A Compact Statistical Model of the Song Syntax in Bengalese Finch. *Plos Computational Biology, 7*(3), e1001108. doi: 10.1371/journal.pcbi.1001108

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42. doi: 10.1016/s0010-0277(02)00004-5

Koba, R., Takemoto, A., Miwa, M., & Nakamura, K. (2012). Characteristics of serial order learning in common marmosets (Callithrix jacchus). *Journal of Comparative Psychology, 126*(3), 279-287. doi: 10.1037/a0026613

Lachlan, R. F., Peters, S., Verhagen, L., & ten Cate, C. (2010). Are There Species-Universal Categories in Bird Song Phonology and Syntax? A Comparative Study of Chaffinches (Fringilla coelebs), Zebra Finches (Taenopygia guttata), and Swamp Sparrows (Melospiza georgiana). *Journal of Comparative Psychology, 124*(1), 92-108. doi: 10.1037/a0016996

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior* (pp. 112-146). New York: Wiley.

Nelson, D. A., Marler, P., & Palleroni, A. (1995). A COMPARATIVE APPROACH TO VOCAL LEARNING - INTRASPECIFIC VARIATION IN THE LEARNING-PROCESS. *Animal Behaviour, 50*, 83-97. doi: 10.1006/anbe.1995.0223

Nishikawa, J., & Okanoya, K. (2006). Distributed neural representation of song element sequence in nucleus HVC of Bengalese finches. *Neuroscience Research, 55*, S97-S97.

Okanoya, K. (2004). The bengalese finch - A window on the behavioral neurobiology of birdsong syntax. In H. P. Zeigler & P. Marler (Eds.), *Behavioral Neurobiology of Birdsong* (Vol. 1016, pp. 724-735).

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition, 109*(3), 423-430. doi: 10.1016/j.cognition.2008.10.004

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology*, *93*(2), 260-277. doi: 10.1037/h0077553

Rissanen, J. (1987). Minimum description length principle. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistic Sciences* (Vol. 5, pp. 523-527): J. Wiley and Sons.

Rose, G. J., Goller, F., Gritton, H. J., Plamondon, S. L., Baugh, A. T., & Cooper, B. G. (2004). Species-typical songs in white-crowned sparrows tutored with only phrase pairs. *Nature, 432*(7018), 753-758. doi: 10.1038/nature02992

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage, 52*(3), 1059-1069. doi: 10.1016/j.neuroimage.2009.10.003

Sacks, H., Scheglof.Ea, & Jefferso.G. (1974). Simplest systematics for organization of turn-taking for conversation. *Language*, *50*(4), 696-735. doi: 10.2307/412243

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27-52. doi: 10.1016/s0010-0277(98)00075-4

Scharff, C., & Nottebohm, F. (1991). A comparative-study of the behavioral deficits following lesions of various parts of the zebra finch song system - implications for vocal learning. *Journal of Neuroscience, 11*(9), 2896-2913.

Solan, Z., Horn, D., Ruppin, E., & Edelman, S. (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(33), 11629-11634. doi: 10.1073/pnas.0409746102

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (1999). Note types, harmonic structure, and note order in the songs of zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology, 113*(2), 194-203. doi: 10.1037/0735-7036.113.2.194

Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and Prosodic Cues for Song Segmentation Learning by Bengalese Finches (Lonchura striata var. domestica). *Ethology, 116*(6), 481-489. doi: 10.1111/j.1439-0310.2010.01772.x

Tchernichovski, O., & Nottebohm, F. (1998). Social inhibition of song imitation among sibling male zebra finches. *Proceedings of the National Academy of Sciences of the United States of America, 95*(15), 8951-8956. doi: 10.1073/pnas.95.15.8951

Terrace, H. S. (1987). Chunking by a pigeon in a serial-learning task. *Nature, 325*(6100), 149-151. doi: 10.1038/325149a0

Tomasello, M. (1998). *The new psychology of language*. Mahwah, NJ: Erlbaum.

Williams, H. (1990). Models for song learning in the zebra finch - fathers or others. *Animal Behaviour, 39,* 745-757. doi: 10.1016/s0003-3472(05)80386-0

Williams, H., & Staples, K. (1992). Syllable chunking in zebra finch (taeniopygia-guttata) song. *Journal of Comparative Psychology, 106*(3), 278-286. doi: 10.1037/0735-7036.106.3.278

Wilson, R. C., & Zhu, P. (2008). A study of graph spectra for comparing graphs and trees. *Pattern Recognition, 41*(9), 2833-2841. doi: 10.1016/j.patcog.2008.03.011

Zann, R. (1993). Structure, sequence and evolution of song elements in wild australian zebra finches. *Auk, 110*(4), 702-715.

Zann, R. A. (1996). *The Zebra Finch: Synthesis of Field and Laboratory Studies*. New York: Oxford University Press.

CHAPTER 4

CONTINGENT PLAYBACK UPON JUVENILE VOCALIZATIONS SHAPES FINAL SONG PRODUCTION IN THE ZEBRA FINCH

Abstract

The immature vocalizations of juvenile songbirds and human infants become a potent communication signal within a short time interval. However, developmental mechanisms promoting learning are often unclear. Contingent feedback on juvenile behavior is thought to be a key feature. Inspired by the observation of natural interactions in the zebra finch, we predicted that immature vocal precursors have functional significance in the learning process. In past tape-tutoring studies juvenile zebra finches were trained in paradigms where juvenile action (key pecking or perch hopping) triggered song playback. We revised the paradigm to mimic more naturalistic interactions. To study the effect of contingent playback on song learning, we arranged for immature juvenile song to trigger playback of adult song. Fourteen acoustically isolated juvenile zebra finch males were trained on six alternate days from the age of 35 days. Seven subjects were assigned to the Contingent treatment group receiving contingent playback upon vocalization. Subjects in the Yoked treatment group received playback independent from their own vocalization. From all subjects mature song was collected and compared with the tutor's to assess quality of learning. Our intervention influenced singing behavior during training and the ultimate song outcome at maturation. Contingent subjects developed significantly more consistent song with a more stable pattern and a smaller proportion of song variants compared to the Yoked birds, although birds in both treatment groups learned the tutor motifs similarly. Our

results suggest that contingent vocal feedback during development is a guiding force in shaping final song production in the zebra finch.

Introduction

Vocal communication is widely prevalent, but vocal learning has been described in only a few vertebrate taxa: in three orders of birds (songbirds, parrots, hummingbirds) and three distantly related groups of mammals (bats, cetaceans, elephants) (Jarvis, 2007). Recent studies of learning processes has led to the discovery of striking similarities between the most frequently studied subjects, songbirds and human infants (e.g., Doupe & Kuhl, 1999; Brainard & Doupe, 2002; Goldstein & Schwade, 2010). The numerous parallels between humans and songbirds suggest similarities in underlying principles and mechanisms. Young birds and babies are faced with a similar challenge early in development – to learn how to produce mature sounds that follow the patterns of the ambient acoustic input.

From early immature precursors (plastic song in birds and babbling in human infants), vocal learning generates potent communication signals within a short period of time. In songbirds, specifically the zebra finch, much of the work on vocal learning has focused on neural mechanisms (Brainard & Doupe, 2002). However, in humans and many other songbird species, from song sparrows to brown-headed cowbirds, research emphasizes behavioral and social mechanisms of learning (e.g. Goldstein, King & West, 2003; Beecher & Brenowitz, 2005; Beecher, Burt, O’Loghlen, Templeton & Campbell, 2007; West & King, 1988). Here we attempt to construct a more integrative understanding of zebra finch song learning by examining the role of social feedback in vocal learning.

In the traditional view, songbirds learn to sing by memorizing the song of conspecifics and later by practicing to match their vocal production to the memorized "song template" (Marler & Tamura, 1964; Konishi, 1965; Marler, 1976). During practice, vocalizations progress through a series of developmental stages: the initial subsong is replaced by a more structured plastic song until crystallization occurs. Developmental progress from subsong toward plastic song is generally thought to be caused by maturation within the neural substrate and by increased training in the musculature of the vocal organ. Similar developmental milestones were described in the vocalization of human infants. The first quasi-resonant utterances are soon replaced by fully-resonant, and later with canonical (acoustically mature) babbling (Oller, 2000). Vocal development has a strong learned component, as babbling is modified by feedback from the baby's social environment. Contingent feedback by the caregiver that reinforces certain vocalizations over others is crucial in shaping the developmental trajectory (Goldstein et al., 2003; Goldstein & Schwade, 2008, 2010).

Could immature vocalizations have similar functional significance during vocal learning in songbirds? Might plastic song elicit reactions and be shaped by feedback from the juvenile bird's social environment? Such effects, termed *socially guided learning*, have been discovered in brown-headed cowbirds. After fledging, juveniles associate with adult conspecifics. The song of juvenile males elicits feedback from adult females in the form of rapid wing movements, called wing strokes, and the young males adjust their singing in response to the feedback. Individual differences in song are related to variation in these social interactions: when the responses of female cowbirds are decreased, the cowbird males learn much more slowly or not at all (West & King, 1988; King, West & Goldstein, 2005; Smith, King & West, 2000). Despite the importance of socially guided learning in human and cowbird vocal development, such effects have rarely been looked for in other songbird species.

Songbirds are an extremely diverse group in terms of their life history traits and an equally large diversity applies to their learning strategies (Beecher & Brenowitz, 2005). Song memorization and practice, sometimes called the sensory and sensorimotor learning periods, are temporally distinct phases in some species, but overlapping phases in others. Separate song memorization and song practice was first described in territorial, temperate zone North American sparrows, e.g. in swamp sparrows and white-crowned sparrows. In these species, memorization takes place during the first 2 months of life, and birds usually start rehearsing months later (Marler, 1970; Marler & Peters, 1981; Marler & Peters, 1987). Other species, like the zebra finch, exhibit a more compressed learning trajectory. In this species, sexual maturation and song learning occur during the first 4 months of life, and the sensory and sensorimotor periods have long been known to overlap (Bohner, 1983; Slater, Eales & Clayton, 1988; Zann, 1996). The temporal overlap between the two learning periods may create opportunities for socially guided song learning, as social feedback to plastic song can be used to modify the song representation. Thus vocal precursors may have functional significance in the learning process, if young males can use social feedback from conspecifics to modify their song.

A detailed look at finch development supports socially guided learning as a potential mechanism of song development. The zebra finch is a socially gregarious colonial species. Song memorization starts around 25 days and immature song production around 35-40 days of age, and young males are able to memorize new songs for another month (Immelmann, 1969; Roper & Zann, 2006). The extended sensitive period enables modification of the song representation at the same time as the juveniles are practicing singing. Juveniles reared with their fathers and then placed with a new song tutor between 35 to 65 days of age are able to learn song from their second tutor (Eales, 1989).

Furthermore, their typical social environment provides opportunities for socially guided learning. Juvenile males prefer to learn to sing from their fathers, both in captivity and in the wild, even if other potential song tutors are available (Bohner, 1983, Zann, 1990). This preference is based on behavioral interactions. In aviaries with multiple breeding pairs, males providing the greatest amount of interactions and parental care are preferred over others as song tutors (Williams, 1990). The level of aggression also influences song tutor selection. When two adult males are present, the one more aggressive toward juveniles is more likely to be chosen as song tutor (Clayton, 1987; Jones & Slater, 1996). Social interactions may even override interspecific boundaries, as juveniles readily learn from foster fathers from different species, such as the Bengalese finch (Clayton, 1988; Eales, 1987), even if other conspecific potential song tutors are within earshot (Immelmann, 1969).

Experimental evidence also indicates that song learning in zebra finch is influenced by interactions with social companions. Live tutors are generally superior to passive tape tutoring when training juvenile zebra finches (Eales, 1985, 1989; Derégnaucourt, Poirier, Kant, Linden & Gahr in press; but see Funabiki & Funabiki, 2009, London & Clayton, 2008), although juveniles do not fully imitate their tutor’s song. In family settings, domesticated zebra finch juveniles develop very similar but not identical songs to their fathers (73% similarity) in contrast when housed singly with their song tutors (91% similarity) (Tchernichovski & Nottebohm, 1998). Live interactions provide multimodal cues, including vocal, visual and perhaps tactile components that facilitate song learning. When one or several of these modalities are absent during passive tape tutoring, birds usually fail to learn. For example, when live adult song was broadcast through speakers to juveniles, they were unable to interact with the tutors, and did not learn the tutor song (Eales, 1989). Moreover, the timing of behavioral feedback may be crucial. Feedback made contingent on plastic song or other juvenile behavior may facilitate song learning compared to passive, unsynchronized exposure. When a juvenile action such as key peck is followed by tutor song playback, juveniles learn the model song better (Adret, 1993; Deregnaucourt et al., in press). However not all studies reported learning differences between treatment groups, as yoked subjects were equally similar to the tutor song as the contingent juveniles (e.g. Houx and ten Cate, 1999a,b), and there was high inter-individual variability in the final learning (Deregnaucourt, Mitra, Feher, Pytte, & Tchernichovski, 2005; Phan, Pytte & Vicario, 2006; for review see Deregnaucourt, 2011). One possibility is that juveniles are able to link their own actions to the tutor song presentation, and so are able to predict the song delivery, which then promotes learning. It is also possible that learning might occur because the procedure matches a typical pattern of social interaction.

In prior detailed behavioral observations, we have found that up to 9 % of immature song by zebra finch juveniles is followed by song from their fathers, and that the frequency of these co-occurrences predicts how similar juveniles’ mature song will be to fathers’ song (Menyhart et al., in prep.)

To move beyond descriptive studies and assess the causal power of social feedback to influence song learning, we provided learning opportunities to zebra finch in a strictly controlled environment by offering them contingent, species-typical feedback to their plastic song.

We tested the role of social interaction in song learning by designing an interactive training paradigm in which we monitored immature song of juveniles and manipulated the subsequent contingency of adult male song. In a contingent treatment group, immature juvenile song was always followed by adult song. In a yoked control group, subjects received song playbacks on an identical schedule as those that had been played to the contingent birds. Hence the feedback was not contingent on their own vocalizations, but was time locked to those of a matched contingent subject. We analyzed the songs of the juveniles to determine the quantity and quality of learning. Do identical playbacks of adult song that differ only in contingency predict how song is learned from the playback stimulus? We predicted that contingent but not yoked control subjects will learn more accurate and consistent song. We found that providing contingent feedback for juvenile vocalizations influenced singing behavior during training and the final song outcome as compared with passive tape exposure.

Materials and Methods

Breeding

Subject birds hatched in large (89 x 125 x 61 cm) aviaries containing multiple families. Each aviary was equipped with 8 nest boxes, nesting material, food and water ad libitum and was kept at a 14:10 light cycle to stimulate readiness to breed. Eight unmated adult males and 8 unmated adult females were placed in the aviary. Mating status of the birds and nest ownership was determined by daily behavioral observation of flight aviaries for 30 min. Each day we recorded the number of eggs and hatchlings. When juveniles reached 14 days of age, mating pairs and their offspring were placed into single cages (62 x 44 x 36 cm) with a nest box, and kept in the same room. Upon removal from the communal aviaries, juveniles were color banded and genetically sexed from pinfeathers. Juveniles do not learn their father’s song before the age of 25 days post hatching (Roper & Zann, 2006). Therefore families were kept together until the oldest juvenile reached 21 days of age. At that point the father was removed to prevent his sons from learning his song, and the family cage was moved to an isolated room. At 31 days, the target male and a single juvenile female (who was kept as a social companion) were placed in an isolated cage in a room with no other birds for the duration of the entire experiment. We used 14 juvenile males total, 7 in each treatment group, and one target male from each family.

Training procedure

Training started at 35 days. Each night before training, a bird was moved to a single cage measuring 46 x 44 x 36 cm with ad libitum food and water to allow habituation to the training room. Training started the following morning within the first hour after the lights came on. Immediately after training the birds were moved back to their housing room. One side of the training cage was covered with a semi-transparent sheet, disguising a speaker (Morel MTD29) and a mounted stuffed adult male zebra finch. The speaker was mounted immediately behind and above the stuffed male, and connected to a desktop computer. The experimenter sat in front of a computer at the other side of the room, behind an opaque white curtain. Training songs were played back manually using *Syrinx* software (John Burt, www.syrinxpc.com) at 70 dB at 30 cm distance from the speaker. A JVC GY-HM100U camera with an attached external Sennheiser K6/ME64 shotgun microphone was mounted in front of the cage to record the training session.

Playback stimuli were recorded from 7 adult zebra finch males in the presence of an adult female. All tutors were bred in captivity and developed a normal song. The same female was used to elicit song production for each of the selected males. From each tutor’s song introductory notes and 4 consequent motifs were selected as a playback stimulus. To simulate natural variance, a 3-motif version of the playback song was created by deleting one more motif. The 4- and 3-motif playback stimuli were played back alternately during training.

Each subject was exposed to six training sessions, and sessions were distributed on alternate days. A training session was successful when the experimental juvenile sang at least 10 songs during the Playback period, and hence was exposed to at least 10 contingent playbacks. If the juvenile did not sing this minimum number of songs, it was trained again the following day. Across all subjects the first successful training session started at the mean age of 38.5 ± 2.1 days, and the last training was administered at 50.14 ± 1.75 days of age. Each training session lasted for 2 hr total and followed an ABA design. The first 30 min of observation (Baseline I) were followed by a 60-min active playback period, and then the training session concluded with a second 30-min observation period (Baseline II). The purpose of Baseline I was to determine the typical singing behavior of the juvenile. During the active playback period the juvenile was exposed to the prerecorded playback stimuli on either a contingent or yoked schedule. In the contingent treatment, the playback was contingent upon juveniles' subsong. In the yoked-control treatment, juveniles were exposed to playback at the same times and frequencies as a contingent juvenile, but not dependent upon their own vocalizations. All the contingent and yoked subjects were matched for vocal maturity at the beginning of the first training session. Each of the seven contingent juveniles was exposed to song from a different tutor and each yoked control bird was exposed to the same tutor song as its matched contingent partner. After the active playback, juvenile singing behavior was again observed during the final Baseline II period. After the end of the sixth successful training session, subjects along with their female companions were housed in acoustic isolation until they reached 90 days of age, when their cages were placed in a communal housing room for the rest of the experiment.

Mature song recordings

Mature song was recorded for all experimental subjects starting on days 90, 120 and 150. For all of the recordings, males were placed in a sound-proof room overnight in a 46 x 44 x 36 cm large cage. The following morning, an unfamiliar adult female zebra finch was placed in an identical cage next to the male's cage and recording began. If the male did not sing in 60 min, further recordings were attempted after a day back with the colony until we obtained at least 10 song bouts in each recording. Due to technical difficulties or the bird's unwillingness to sing, song of only 13 males was recorded at all three recording ages. All recordings were made with a JVC GY-HM100U camera with an attached external Sennheiser K6/ME64 shotgun microphone. Sound files were created using Soundtrack Pro (version 3.0.1.) and uncompressed sound files were saved as separate wav files.

Behavioral analysis of training sessions

Behaviors on the video recordings of the six training sessions were annotated using *ELAN* EUDICO Linguistic Annotator Version 4.0.0. (Copyright © 2001 - 2012 Max- Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands). The software package is able to synchronize the video and sound files, and allows multiple categories of behavior ("tiers") from individual animals to be annotated onto different fields associated with the video record. The number, timing, and association of these behavioral categories can then be readily assessed. As an initial step, every occurrence of juvenile and playback song was annotated with frame accuracy in separate tiers. Juvenile song was defined as continuous singing that contained no pauses of one second or longer. Then, 15-second time windows were created in another tier before and after every song in order to assess the temporal relations between the song of juveniles and the playback stimuli. Based on the annotations, the song rate (number of songs per min) and mean duration of juvenile song were calculated for every observation period.

When animals (and humans) learn associations between a stimulus and its triggering behavior, eliminating the contingent feedback leads to a temporary increase in animals response rate, called extinction burst (for review see Lerman & Iwata, 1995) The magnitude of extinction burst has been associated with later learning (e.g., Goldstein, Schwade & Bornstein, 2009). Therefore we assessed the direction and magnitude of change in song duration and song rate between the active and final baseline periods.

For the active playback period, the following additional features were also calculated: mean time delay of playback stimuli after juvenile song, percent contingency of playback to juvenile song, percent of juvenile song interrupted by playback, percent of playback interrupted by juvenile song. Percent contingency was defined as the number of juvenile songs followed by playback divided by the total number of juvenile song in an active period. Instances where the tutor playback overlapped with the juvenile song were not calculated into percent contingency, but qualified as song interruptions. Juvenile song was interrupted when the playback started before the juvenile song was over. Percent of juvenile song interrupted was calculated as the number of juvenile songs interrupted divided by the total number of juvenile songs in the active period. Playback was interrupted when the juvenile started to sing before the tutor playback finished playing. Percent of playback interrupted was computed as the number of interrupted playbacks divided by the total number of playbacks during the active period.

Juveniles sometimes engaged in vocal interactions with the tutor playback. A vocal interaction was defined as countersinging, i.e., the juvenile subject began to sing within 2.5 sec after the tutor song ended. Interactions were not coded when the juvenile started to sing while the tutor song was still ongoing, or juvenile and tutor song began at the same time. We assessed both the absolute number of countersinging interactions and the percentage of juvenile songs that occurred in a countersinging interaction relative to the total number of songs produced during the active playback period.

Mature song analysis

Song structure

Adult zebra finch males sing only one song type that is between 0.5-1.5 sec in length and composed of a series of introductory notes at the beginning of the song. Males sing up to 14 syllable types that are organized into a roughly stereotyped pattern called a motif. Individual motifs are repeated between 1-8 times during a single song bout (Price, 1979). Although zebra finch song is assumed to be highly stereotyped, syllable deletions, additions, and repetitions occur in more than one third of the motifs produced (Sturdy, Phillmore & Weisman, 1999). To quantify variability in song production, we counted first how many different motif types were produced in each juvenile's mature song. We defined song bouts as containing silent intervals shorter than 500 msec. Then we assessed the following measures in the 90, 120 and 150 days recordings: the percentage of the dominant (primary) motif, the percentage of the second most common motif (secondary), and the percentage of the remaining types, which we defined as miscellaneous motifs. For motif categorization we employed a conservative approach: a single syllable deviation resulted in a new motif category. Motif structure was assessed blind to the treatment group.

To check the validity of our manual motif categorization, we also assessed the song structure quantitatively. First, a syllable catalog was created for every subject; then every song bout within each recording was transcribed into a string of these constituent syllables using Syrinx software (John Burt, www.syrinxpc.com). The resulting individual song corpora were analyzed with a biologically inspired model of language acquisition, called BAGEL (Kolodny, Lotem & Edelman, 2011). The model searches for syllable collocations, namely for sequence of basic units recurring more likely than chance. From the output of the model, we calculated the percentage of the most common sequence, the second most common sequence, and the percentage of miscellaneous sequences. (More detailed description of the model is available in Menyhart et al. in prep).

Similarity to playback stimuli

From every recording of mature song, we attempted to select the first 30 motifs not occluded by cage or female noise for comparison with the playback stimuli. In the recordings at 90 days, two males produced respectively 25 and 23 motifs that were used for the comparisons. In the recordings at 120 days, three males produced 15, 17, and 20 motifs that were used for the comparisons. The rest of the males produced the required number of motifs. Juvenile and playback vocalizations were compared using *Sound Analysis Pro 2011 (SAP2011)* "M x N" comparisons within the Similarity batch module (Tchernichovski, Nottebohm, Ho, Pesaran & Mitra, 2000). Similarity generated by *SAP2011* has three major components: the percent similarity score is computed over longer intervals of sound (typically 50 - 70 msec) and reflects the amount of song material included from the tutor's song in the juvenile's motif. Accuracy is computed across shorter time windows (5 - 10 msec) and indicates how well the juvenile sound matches the tutor's sound across the similar segments. Sequential match accounts for the temporal order (syntax) of sounds, and refers to the similarity of temporal order of *final sections* (as defined in the SAP 2011 manual) between the first reference sound and the second sound. A combined similarity score between juvenile and tutor song was calculated as percent similarity times accuracy.

Additionally, from the 120-day recordings an observer experienced with zebra finch song (OM) conducted visual comparison between the syllables of juvenile and tutor song, and identified the percentage of juvenile repertoire copied from the tutor and the repertoire size (total number of different syllable types the juvenile produced).

Self-similarity

In each final recording, the first 20 motifs that did not contain background noise were compared using 20 x 20 comparisons within the Similarity batch module of *Sound Analysis Pro 2011*. We assessed percent similarity, accuracy and sequential match between motifs of the juvenile song. The overall similarity score combined percent similarity, accuracy, and sequential match scores to assess singing consistency of the juveniles.

Statistical analysis

Juveniles sometimes did not produce any song during the observation periods. However, we first used logistic regression to test whether absence of song was confounded with treatment. Because absence of singing was not associated with treatment, training session, or observation period, observation periods with no activity were removed from later analyses. We analyzed treatment effect on song duration and frequency (song rate) with a linear mixed model. We included treatment, training session, observation period, and their interaction terms as fixed effects, and pair membership and its interactions with subject identity, training session, and period as random effects. We sequentially deleted nonsignificant interaction terms until we reached the model that accounted for a significant portion of the variance in the data. All significant or near significant interactions were decomposed by pairwise comparison of estimated marginal means, as examining marginal interactions does not increase the likelihood of Type I error (Howell, 2010). Due to multiple comparisons, significant results are reported along with the adjusted alpha values (see Tables 4.2 and 4.4 for the detailed models).

We tested treatment effects on the song structure parameters and similarity measures with a repeated measures linear mixed model. We included treatment and recording times (90, 120, or 150 days) as fixed effects, and subject identity and pair membership as random effects. The following covariates were also included in our model: mean time delay of playback stimuli after juvenile song, percent contingency of playback to juvenile song, percent of juvenile song interrupted by playback, and percentage of countersinging. We ran all model analyses on SPSS version 20. We checked residuals for normality and random distribution of variances, and when necessary, data were log transformed (Field, 2009), which is indicated at each model. Relationships between the properties of training sessions and song outcome measures were analyzed with nonparametric correlations due to the non-normal distribution of the majority of song outcome measures.

Results

Experimental treatment: timing of playback

We defined contingency as tutor song feedback immediately following juvenile song. Our design was 100% contingency, to provide experimental birds with song feedback to every instance of song practice. In practice, manual delivery of the feedback sometimes interrupted juvenile songs (when songs came in quick succession after stimulus presentation), or juveniles interrupted playback song. When those instances were removed, contingency was slightly less than 100%. While, by chance, some of the presentations to yoked controls followed immediately after their vocalizations, overall contingency in the contingent group ($M = 78.59$, $SD = 10.25$) was far greater than in the controls ($M = 2.84$, $SD = 4.17$), $t(6) = 20.75$, $p < 0.001$. In the contingent treatment group the manual playback presentation resulted in variable time delay between juvenile song and the contingent stimulus ($M = 1.76$, $SD = 0.31$, min: 1.29, max: 2.15 sec).

Juveniles were equally likely to interrupt the tutor song playback in both treatment groups (contingent: $M = 1.65$ $SD = 2.57$, yoked: $M = 0.27$ $SD = 0.71$), $t(6) = 1.30$, $p = 0.241$. However, more contingent subject’s songs were interrupted by the tutor song playback ($M = 8.06$, $SD = 6.56$) compared to controls ($M = 1.38$, $SD = 1.01$), $t(6) = 2.75$, $p = 0.033$.

Contingency affects delivery of juvenile song during training

Song durations

We tested the effect of contingent feedback on song duration during training (Table 4.1) with a linear mixed model using a 2 (treatment) x 6 (training session) x 3 (observation period) design. There was a significant interaction between treatment and observation period on song duration $F(2, 105.14) = 11.66$, $p < 0.001$. We also found a significant main effect of observation period $F(2, 104.67) = 23.6$, $p < 0.001$ (the full model is described in Table 4.2). Decomposing the interaction revealed that contingent subjects sang significantly shorter songs during the active playback period compared to their matched controls ($p = 0.006$ at alpha = 0.017), but not in the two baseline periods. In contingent subjects, song duration in the active playback period was significantly shorter than in both the initial and final baselines ($p < 0.001$ in both cases at alpha = 0.008). There was no difference in song duration between observation periods within the yoked control treatment (Figure 4.1).

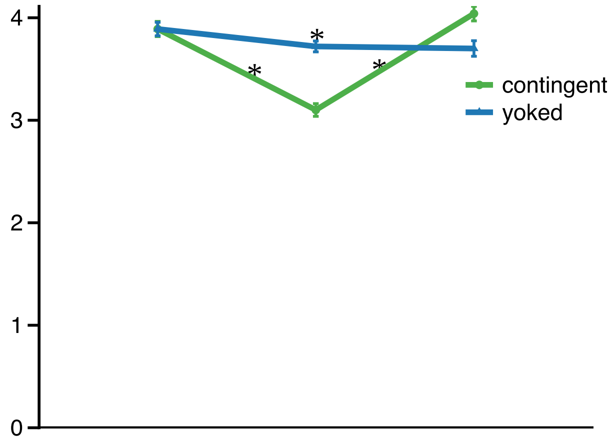

Figure 4.1 Song duration between contingent and yoked subjects across the three observation periods.

Table 4.1 Mean song duration between treatment groups during the three observation periods across the entire training.

Observation period	Treatment group	Mean song duration (sec) ± SD
Baseline I	contingent	3.89 ± 2.93
	yoked	3.89 ± 3.24
active playback	contingent	3.1 ± 2.72
	yoked	3.72 ± 3.12
Baseline II	contingent	4.04 ± 3.15
	yoked	3.7 ± 3.19

Table 4.2 Linear mixed model testing for the effects of treatment, training session, observation period, and their interaction terms on song duration during training. Song duration was log transformed to normalize distribution of data.

Final model	*d. f.*	*F*	*p*
Treatment	1, 11.86	1.02	0.333

Observation period	2, 104.7	23.6	**0**
Training session	5, 53.00	0.44	0.822
Treatment x period	2, 105.14	11.66	**0**
Rejected terms			
Treatment x Training session x Observation period			
Training session x Observation period			
Treatment x Training session			

Song rate

The effect of contingent feedback on song rate (Table 4.3) during training was assessed with a linear mixed model using a 2 (treatment) x 6 (training session) x 3 (observation period) design. We found a marginally significant interaction between treatment and observation period, $F(2, 167.88) = 2.8$, $p = 0.064$, significant main effects of training session and observation period, and marginally significant effect of treatment (the full model is described in Table 4.4). According to Howell (2010) examining marginal interactions does not increase the likelihood of Type I error. Decomposing the interaction revealed that contingent subjects sang with a significantly lower rate during the active playback period compared to yoked controls ($p = 0.008$ at alpha = 0.017), but not during any of the baseline periods. Contingent subjects significantly decreased their song rate during the active playback period compared to the initial baseline ($p < 0.001$ at alpha = 0.008). Then, after playback ceased, song rate increased significantly in the final baseline ($p < 0.001$ at alpha = 0.008). Control subjects also decreased their song rate in the active playback period compared to the initial baseline ($p < 0.004$ at alpha = 0.008), but unlike their contingent partners, their song rate did not change from the active to the final baseline period ($p = 0.183$ at alpha = 0.008) (Figure 4.2).

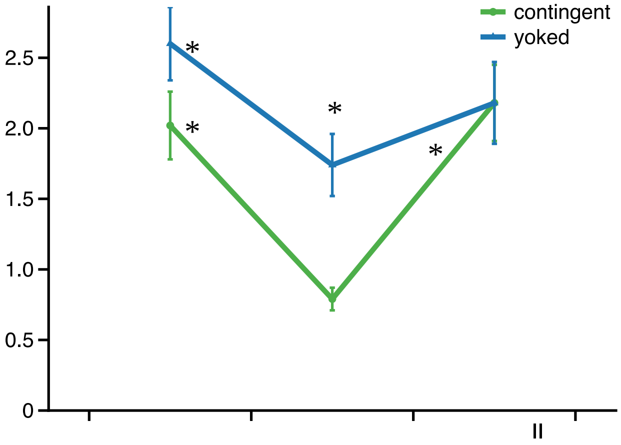

Figure 4.2 Song rate (number of songs per min) between contingent and yoked subjects across the three observation periods.

Table 4.3 Mean song rate per minute summarized for the three observation periods across contingent and yoked subjects.

Observation period	Treatment group	Mean song rate per min ± SD
baseline 1	contingent	2.02 ± 1.2
	yoked	2.6 ± 1.43
active playback	contingent	0.79 ± 0.53
	yoked	1.74 ± 1.25
baseline 2	contingent	2.18 ± 1.41
	yoked	2.18 ± 1.49

Table 4.4 Linear mixed model testing the effects of treatment, training session, observation period, and their interactions on song rate during training.

Final model	*d. f.*	*F*	*p*
Treatment	1, 11.874	4.186	0.064
Session	5, 166.652	5.472	**0**
Period	2, 168.067	14.273	**0**
Treatment x Period	2, 167.877	2.799	0.064
Rejected terms			
Treatment x Session x Period			
Session x Period			
Treatment x Session			

Changes in song activity after the playback period

The abrupt termination of the contingent feedback after the active playback period may have influenced juvenile singing behavior in the final observation period. We calculated changes in song duration and song rate between the active playback and final baseline periods by subtracting singing activity during playback from the activity during the final baseline. After the removal of contingent feedback in the final baseline, there was no significant effect of treatment on changes in song duration or song rate, when difference scores were averaged across the six training sessions (song duration $t(6) = 0.87$, $p = 0.42$, song rate $t(6) = -0.473$, $p = 0.65$) (Table 4.5).

To obtain a more detailed insight into the change in song duration and song rate after playback termination, we divided the entire training session into three equal time periods (training session 1+2, 3+4, 5+6) to assess how the difference score changes across the beginning, the middle, and the end of the training. The effect of playback termination on change of song activity was assessed with a separate 2 (treatment) by 3 (training sessions: beginning, middle, end of training) mixed ANOVA for the change in song duration and song rate. Contingent and yoked subjects song duration did not differ after playback termination, as there was no significant effect of treatment or training sessions on the difference score of song duration (*p*s > 0.1).

However, song rate was influenced differently across treatment groups and training sessions. We found a significant interaction between treatment and training session $F(2, 24) = 2.85$, $p = 0.012$, a marginally significant effect of training session $F(2, 24) = 5.35$, $p = 0.077$, but no main effect of treatment $F(1, 12) = 1.95$, $p = 0.188$. Decomposing the interaction revealed that the termination of playback resulted in divergence in song rate between contingent and yoked subjects in the middle of the training session (training sessions 3+4) (p = 0.014 at alpha = 0.017). While contingent subjects tended to increased their song rate after playback termination in the middle of training compared to the first two training sessions (p = 0.022 at alpha = 0.008), yoked subjects sang less frequently (p = 0.039 at alpha = 0.008). Yoked subjects tended to increase their song rate from the middle to the last two training sessions (p = 0.024 at alpha = 0.008) (Figure 4.3).

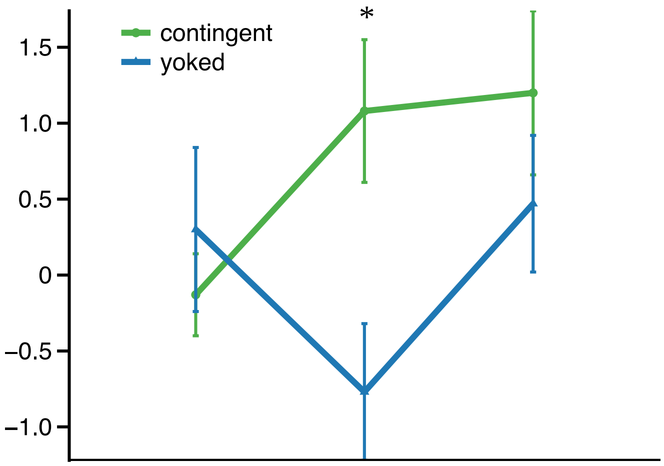

Figure 4.3 The extent of change in song rate after playback termination calculated as song rate during the active playback subtracted from the song rate during the final baseline period. Difference score was calculated for the first two, the middle and last two training sessions.

Changes in song activity after the playback period correlate with learning

We assessed if the overall extent of change in song duration and song rate between the active playback and the final observation period predicts song learning. When activity in all training sessions was included, birds that increased their song duration more after contingent playback ended learned a larger proportion of their repertoire from the tutor, r_s (13) = 0.666, p =

0.009) and developed a smaller total repertoire size measured at 120 days, $r_s(14) = -0.703$, $p = 0.005$. Juveniles that sang more frequently in the final baseline compared to the playback period developed more similar song to the tutor measured at 150 days (combined similarity score, $r_s(13) = 0.632$, $p = 0.021$) (Figure 4.4).

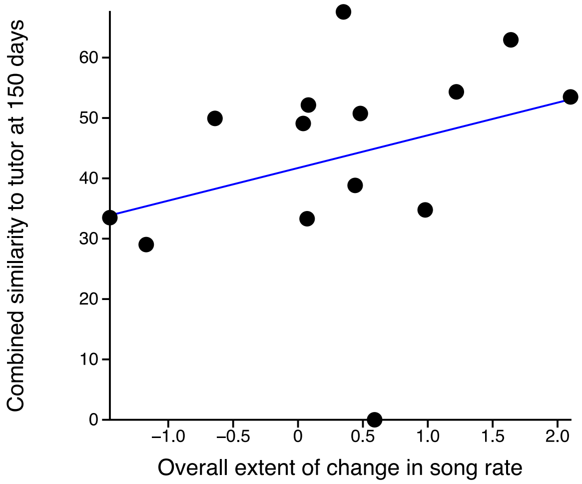

Figure 4.4 Subjects that increased their song rate more after contingent playback ended learned a more similar copy of their song model.

Table 4.5 The extent of change in song duration and song rate after playback terminated between contingent and yoked subjects across all six training sessions. Positive numbers indicate increase in song activity from the active to the final observation period.

	Contingent Mean ± SD	Yoked Mean ± SD
Song duration	0.26 ± 0.78	- 0.42 ± 1.85
Song rate	0.68 ± 0.7	-0.0003 ± 1.16

Countersinging

Treatment affected the timing of the singing behavior of our subjects. Out of the 14 subjects, 8 birds participated in countersinging interactions. A linear mixed model with a 2 (treatment) x 6 (training session) design revealed that contingent subjects were more likely to participate in countersinging than their matched controls, $F(1, 3.81) = 9.49$, $p = 0.039$. Moreover, according to a paired-sample t-test, contingent juveniles sang a larger proportion of their total songs within 2.5 sec after the tutor playback (*M* = 0.086, *SD* = 0.071) compared to yoked controls (*M* = 0.002, *SD* = 0.002), *t*(6) = 2.98, *p* = 0.025. Within the contingent treatment group, juveniles that sang at a higher rate sang more of their song immediately after the playback stimulus, $r_s(7) = 0.811$, *p* = 0.027.

Final song outcome measures between treatment groups

Mature song structure

To test the effects of training on song outcome, we conducted a 2 (treatment) x 3 (recording time) mixed ANOVA on singing activity. As adults, animals from the two treatment groups sang similar numbers of mature song motifs every time they were recorded (Table 4.6) (time *F*(1.67, 16.73) = 0.42 , *p* = 0.63, time x treatment *p* = 0.664).

Moreover, percentage of repertoire copied from tutor and total repertoire size measured at 120 days did not differ between treatment groups (paired t-tests, *p*s > 0.1). To compare changes in the song structure (proportion of primary, secondary and miscellaneous motifs) over time, we conducted separate linear mixed models for each song structure measure, each with a 2 (treatment) x 3 (recording time) design.

The percentage of song motifs comprised by the dominant motif did not differ significantly at any recording time between contingent and yoked subjects (treatment *F*(1, 12.44) = 1.82, *p* = 0.202; time *F*(2, 23.05) = 0.25, *p* = 0.783; treatment x time *F*(2, 23.05) = 2.71, *p* = 0.088).

However, use of a secondary motif increased substantially over time in the contingent group, but not in the yoked control subjects, as indicated by a significant treatment x time interaction *F*(2, 22.5) = 6.12, *p* = 0.008 (Figure 4.5). Decomposing the interaction revealed that only the contingent subjects increased the proportion of secondary motifs between 90 and 120 days ($p < 0.001$ at alpha = 0.008), and between 90 and 150 days ($p = 0.006$ at alpha = 0.008) (Figure 4.6).

Contingent birds sang a lower percentage of miscellaneous motifs across the three recording times than did yoked birds (treatment: $F(1, 12.49) = 5.75$, $p = 0.033$) (Figure 4.7). No other effects were significant for miscellaneous motifs.

We checked the validity of our manual song structure assessment with a linguistic tool BAGEL, by searching for sequences among syllables occurring more likely than chance, called syllable collocations. Similar to song structure, we obtained the percentage of the most common, second-most-common and miscellaneous syllable sequences. Analysis of syllable collocations produced findings similar to the manual motif annotation. Based on the 120-day recordings, contingent birds produced more of the second-most-common sequence than did yoked controls $t(5) = 3.58$, $p = 0.016$. Contingent and yoked subjects did not differ in the percentage of the most common ($t(5) = -0.518$, $p = 0.627$) and miscellaneous sequences ($t(5) = -0.588$, $p = 0.582$) extracted from syllable collocations.

Table 4.6 Number of mature motifs produced during the 90-, 120-, and 150-day song recordings.

	Mean ± SD	
Recording times	Contingent	Yoked
90 days	173.5 ± 139.47	246.14 ± 271.14
120 days	220.33 ± 168.24	170.14 ± 120.89
150 days	309.17 ± 284.67	217.71 ± 139.88

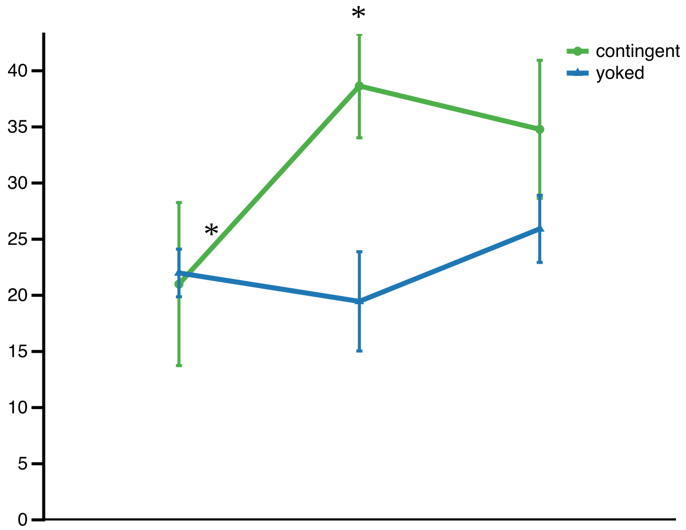

Figure 4.5 Only birds in the contingent treatment increased the percentage of secondary motifs from 90 to 120 days and from 90 to 150 days.

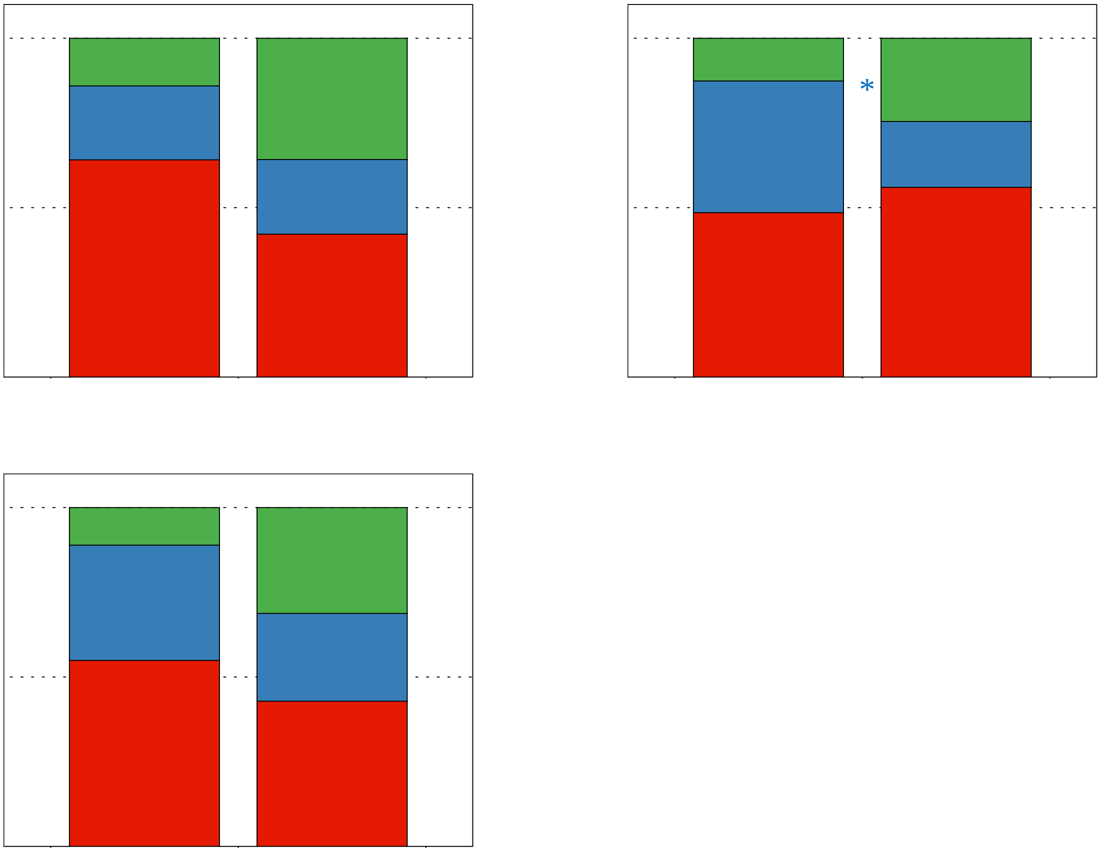

Figure 4.6 Mean percentage of primary (red), secondary (blue) and miscellaneous (green) motif types at 90, 120 and 150 days by treatment groups.

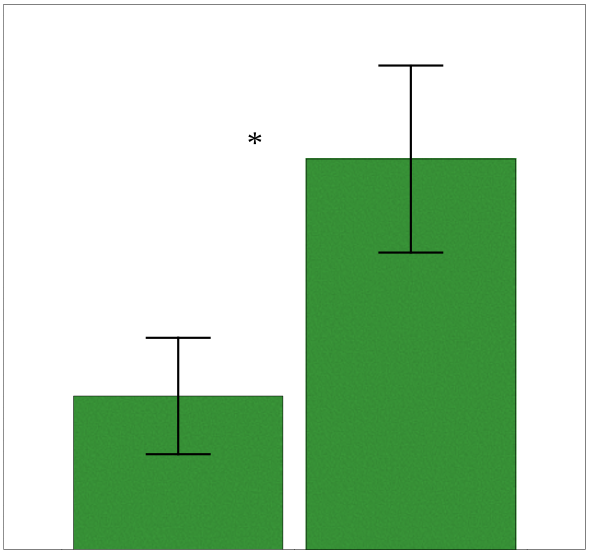

Figure 4.7 Percentage of miscellaneous motif types was significantly different between contingent and yoked subjects pooled across the three recording times.

Similarity to tutor

To assess similarity between juvenile and tutor songs, we conducted separate 2 (treatment) x 3 (recording time) linear mixed models for each similarity measure (percent similarity, accuracy, combined similarity and accuracy, sequential match). Based on the first 30 motifs from the mature song recordings, contingent and yoked subjects copied the playback stimulus similarly well (Table 4.7). Percent similarity to tutor song was not dependent on treatment groups *F*(1, 6.01) = 3.25, *p* = 0.12, or on the time of the mature song recordings, *F*(2, 24.16) = 2.89, *p* = 0.075. Similarly, accuracy was not different between treatments, *F*(1, 6.01) = 1.1, *p* = 0.34, or across recording times, *F*(2, 24.28) = 0.16, *p* = 0.85. The combined score of percent similarity and accuracy also did not differ across treatments or recording times.

Sequential match score increased across recording times, $F(2, 24.08) = 4.59$, $p = 0.02$, with significant increases from 120 to 150 days of age. However, there was no significant difference between treatment groups $F(1, 12.03) = 0.481$, $p = 0.5$.

Table 4.7 Subject's similarity to tutor (percent similarity, accuracy, sequential match and combined scores of percent similarity and accuracy) measured at 90, 120 and 150 days of age.

Percent similarity to tutor

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	57.88 ± 12.80	70.76 ± 14.00
120 days	61.52 ± 9.94	72.58 ± 18.83
150 days	59.32 ± 11.75	66.97 ± 18.68

Accuracy of learning

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	72.71 ± 1.43	74.24 ± 3.71
120 days	72.7 ± 2.29	74.06 ± 3.05
150 days	73.83 ± 2.34	73.89 ± 3.05

Sequential match

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	71.7 ± 12.65	67.51 ± 18.54
120 days	73.27 ± 12.1	60.87 ± 24.05
150 days	75.68 ± 7.67	67.22 ± 18.4

Combined similarity score (percent similarity * accuracy)

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	42.2 ± 10.03	52.53 ± 10.73
120 days	44.86 ± 8.43	53.53 ± 13.15
150 days	43.88 ± 9.25	49.48 ± 14.18

Self-similarity

Self-similarity scores were assessed by comparing the first 20 motifs the juveniles produced to themselves at each recording time (Table 4.8). Separate 2 (treatment) x 3 (recording time) linear mixed models were computed for each of the four similarity measures (percent self-similarity, accuracy, sequential match, and overall combined similarity score).

For percent of self-similarity there was a significant main effect of time ($F(2, 29.6) = 7.98$, $p = 0.002$). Song of both contingent and yoked subjects showed increased self-similarity between 120 and 150 days ($ps < 0.001$ at alpha = 0.008). Similarly, for accuracy we found a main effect of time ($F(2, 29.57) = 5.03$, $p = 0.013$), as both contingent and yoked subjects increased self-accuracy from 120 to 150 days, although the increase was not significant after Bonferroni correction for multiple comparisons (contingent: $p = 0.043$; yoked $p = 0.028$ at alpha = 0.008) (Figure 4.8A, B). There was no significant effect of treatment group for either similarity measure.

For sequential scores, providing contingent feedback changed similarity. Self-similarity in song sequence was influenced by a significant interaction between time and treatment $F(2, 22.64) = 5.49$, $p = 0.011$, and a significant effect of time $F(2, 22.64) = 18.59$, $p < 0.001$. Self-similarity in song sequence appeared to be more plastic in the contingent treatment, and tended to decrease between 90 and 120 days ($p = 0.01$ at alpha = 0.008) but increased significantly by day 150 ($p < 0.001$ at alpha = 0.008) recordings. In yoked controls, change over time was not significant (Figure 4.8C). At 150 days contingent subjects reached higher sequence consistency scores than controls, although the difference was not significant after Bonferroni correction ($p = 0.051$ at alpha = 0.017).

For the combined score of self-similarity there was a main effect of time $F(2, 35) = 11.34$, $p < 0.001$, and marginally significant effect of treatment $F(1, 35) = 3.99$, $p = 0.054$, but the interaction was not significant $F(2, 33) = 1.023$, $p = 0.371$. Combined score of self-similarity increased between 90 and 120 ($p < 0.001$ at alpha = 0.008) and between 90 and 150 days ($p = 0.006$ at alpha = 0.008) in both treatment groups (Figure 4.8D).

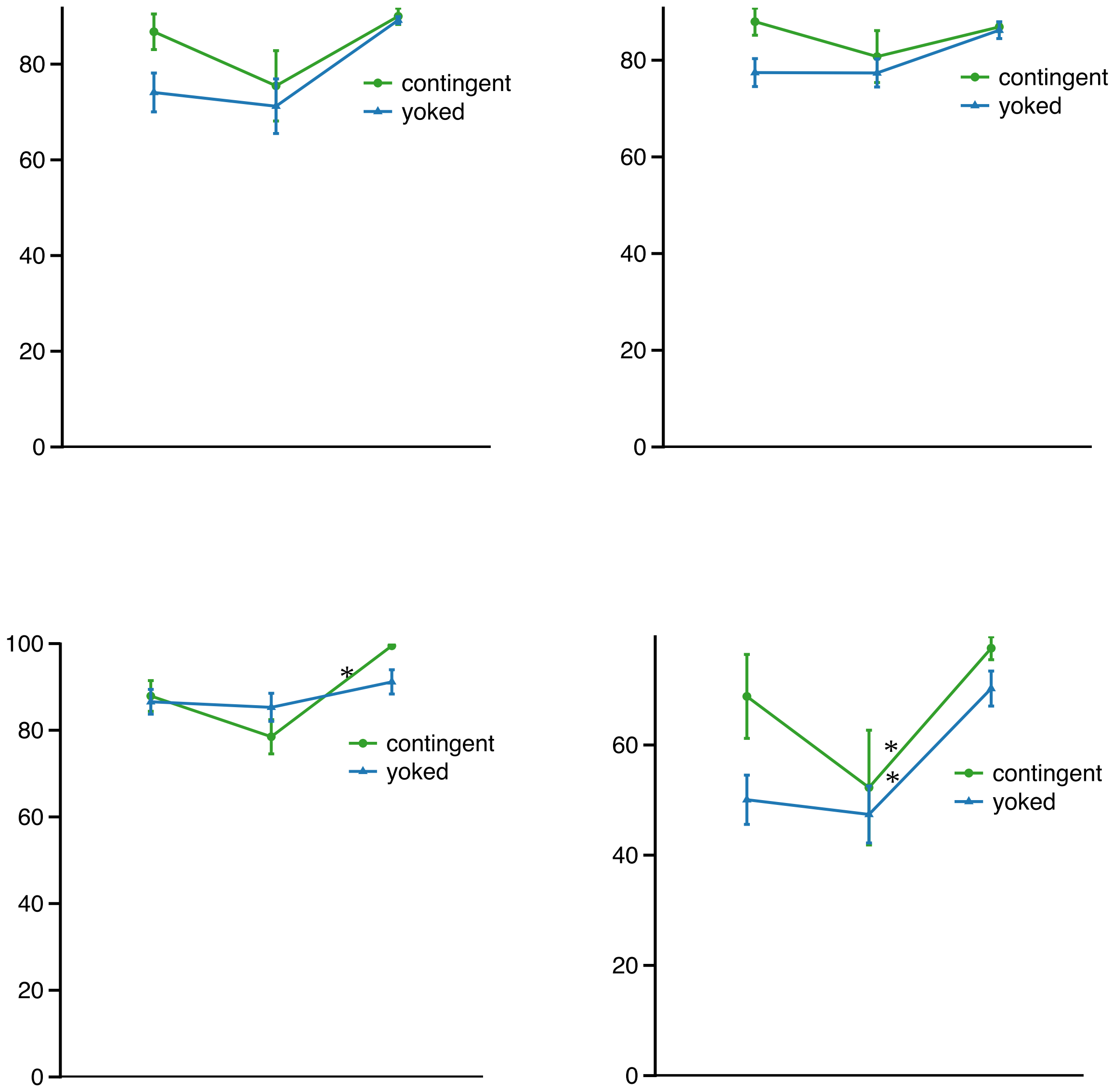

Figure 4.8 A, B, C, D. Self-similarity between treatment groups at the three recording times (90, 120 and 150 days). A: percent similarity, B: accuracy, C: sequential match, D: combined score (percent similarity * accuracy * sequential match).

Table 4.8 Self-similarity scores (percent similarity, accuracy, sequential match, and their combined value) measured at 90, 120 and 150 days of age.

Percent similarity

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	86.77 ± 9.13	74.08 ± 10.75
120 days	75.48 ± 18.01	71.22 ± 15.07
150 days	90.00 ± 4.19	89.20 ± 1.87

Accuracy

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	86.49 ± 6.67	77.44 ± 7.64
120 days	78.71 ± 11.83	77.36 ± 7.71
150 days	87.46 ± 2.41	86.21 ± 4.61

Sequential match

Recording times	Contingent mean±SD	Yoked mean±SD
90 days	88.21±7.12	86.57±7.58
120 days	79.03±8.01	85.29±8.58
150 days	98.99±1.48	91.16±7.42

Combined score (percent similarity * accuracy * sequential match)

Recording times	Contingent mean ± SD	Yoked mean ± SD
90 days	67.23 ± 16.5	50.06 ± 11.83
120 days	49.15 ± 22.17	47.39 ± 13.67
150 days	77.93 ± 4.27	70.24 ± 8.42

Training properties affect final song outcomes

Percent contingency

To assess the effects of training on song outcome, we correlated the mean number of tutor playbacks per training session (M = 39.97, SD = 13.55) with the measures of song structure, similarity to tutor, and self-similarity at the three recording times by conducting separate correlations for each treatment group. In the contingent group, birds that received more tutor song playbacks developed a larger total repertoire size obtained from the 120-day recordings (r_s (7) = .0811, p = 0.027). Yoked subjects with increased exposure to passive playback developed a less stable song structure by 120 days (percentage of the secondary motif type, r_s (7) = -0.786, p = 0.036) and tended to learn the tutor song less accurately measured at 150 days (accuracy between juvenile and tutor, r_s (7) = -0.75, p = 0.052).

Because there was a substantial range of contingency in the contingent group (M = 78.59, SD = 10.25, range: 59.94 - 88.25%), we measured correlations between the percentage of contingent playbacks and final song outcomes only for the contingent treatment group. Juveniles who received more contingent playback tended to learn less from the tutor (percent similarity: r_s

(6) = -0.771, $p = 0.072$, combined similarity score: r_s (6) = -0.771, $p = 0.072$ measured at 90 days). The relationship between percent contingency and similarity to tutor is illustrated in Figure 4.9.

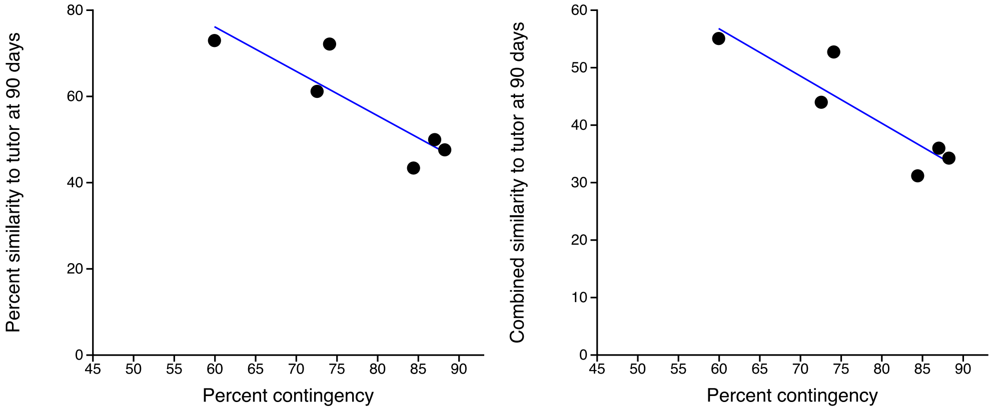

Figure 4.9 Relationship between the percentage of contingent feedback on juvenile song and the following song outcome measures: A: percent similarity to tutor obtained from 90 days recordings; B: combined similarity score to tutor at 90 days.

Countersinging interactions

Contingent subjects were more likely to participate in countersinging interactions compared to the yoked subjects $F(1, 3.81) = 9.49$, $p = 0.039$. To assess the possible effects of countersinging on song learning in the contingent treatment group, we correlated the proportion of juvenile song immediately after playback with the song outcome measures obtained at the three recording times. Juveniles who sang a larger proportion of their song immediately after playback developed a more consistent song, as obtained by the percent self-similarity (r_s (6) = 0.812, $p = 0.05$), self-accuracy (r_s (6) = 0.812, $p = 0.05$), and combined self - similarity scores (r_s (6) = 0.812, $p = 0.05$) from the 90-day recordings (Figure 4.10).

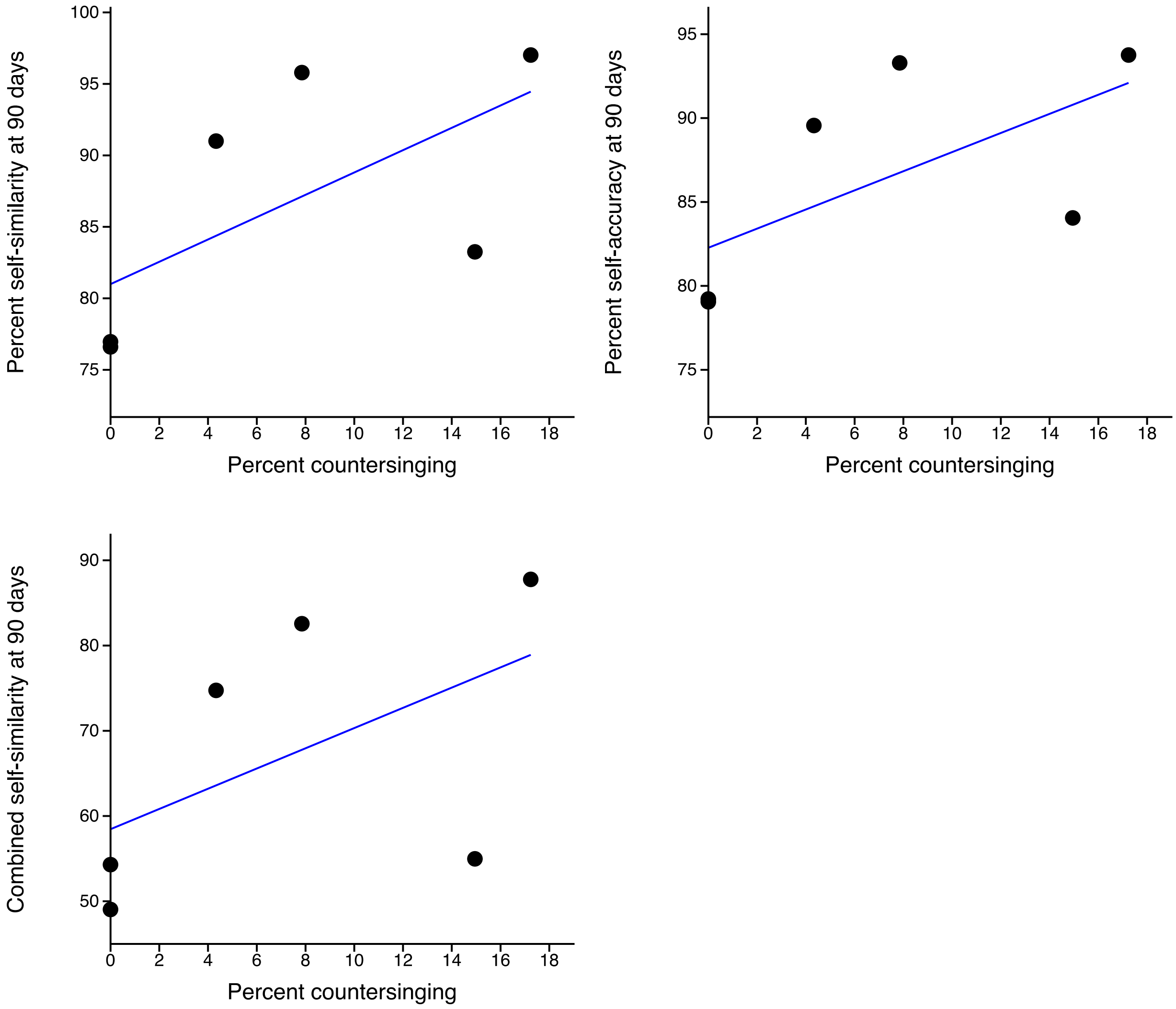

Figure 4.10 Relationship between percent countersinging (defined as the proportion of juvenile songs following the playback within 2.5 sec), and self-similarity measures obtained from the 90-day recordings: A: percent self-similarity; B: percent self-accuracy; C: combined self-similarity score.

Song interruption

The song of contingent birds was more likely to be interrupted than was the song of the yoked controls by the tutor song playback (contingent: $M = 8.06$, $SD = 6.56$, yoked: $M = 1.38$, $SD = 1.01$), $t(6) = 2.75$, $p = 0.033$). For contingent birds, longer songs were more likely to be interrupted $r_s\,(7) = 0.75$ $p = 0.052$).

To test the effect of song interruption on final song outcome in the contingent group, we correlated the percentage of juvenile song interrupted by playback with the song outcome measures obtained from the three recording times. Juveniles with a higher proportion of interrupted song developed a less overall consistent song measured at 150 days (correlation with combined self-similarity (r_s (6) = -0.829, p = 0.042). However, a higher proportion of song interruption correlated negatively with the percentage of miscellaneous motifs at 150 days (r_s (6) = -0.943, p = 0.005, Figure 4.11).

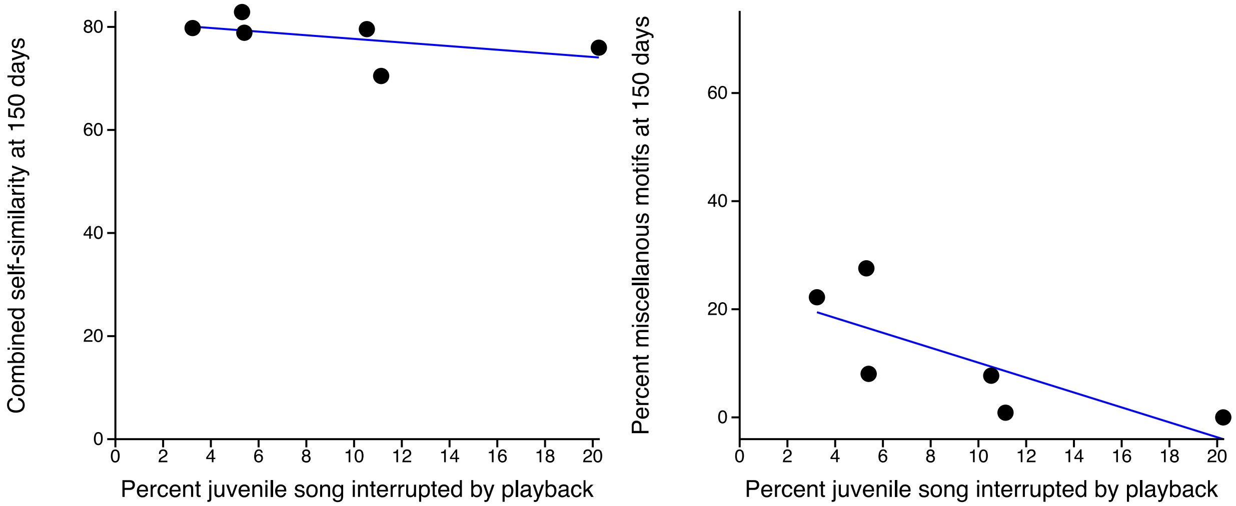

Figure 4.11 Percentage of juvenile song interrupted by tutor playback predicted the combined self-similarity score (self-similarity*accuracy*sequential match) and percentage of miscellaneous motifs at 150 days.

Time delay of playback delivery

To analyze the effect of time delay of the playback delivery on song learning we correlated the mean time delay across training sessions with the final song outcome measures. Birds who received more delayed playback on average learned less of their repertoire from the tutor based on the 120-day song recordings (r_s (13) = -0.645, p = 0.017). Moreover, juveniles experiencing longer average delay developed a more variable song structure, as they sang fewer dominant (r_s (12) = -0.587, p = 0.045) and more miscellaneous motifs in their song (r_s (12) = 0.637 p = 0.026) at 150 days. Consistently, subjects who received a more delayed playback acquired a lower combined self-similarity score measured at 150 days (r_s (12) = -0.643 p = 0.024, Figure 4.12).

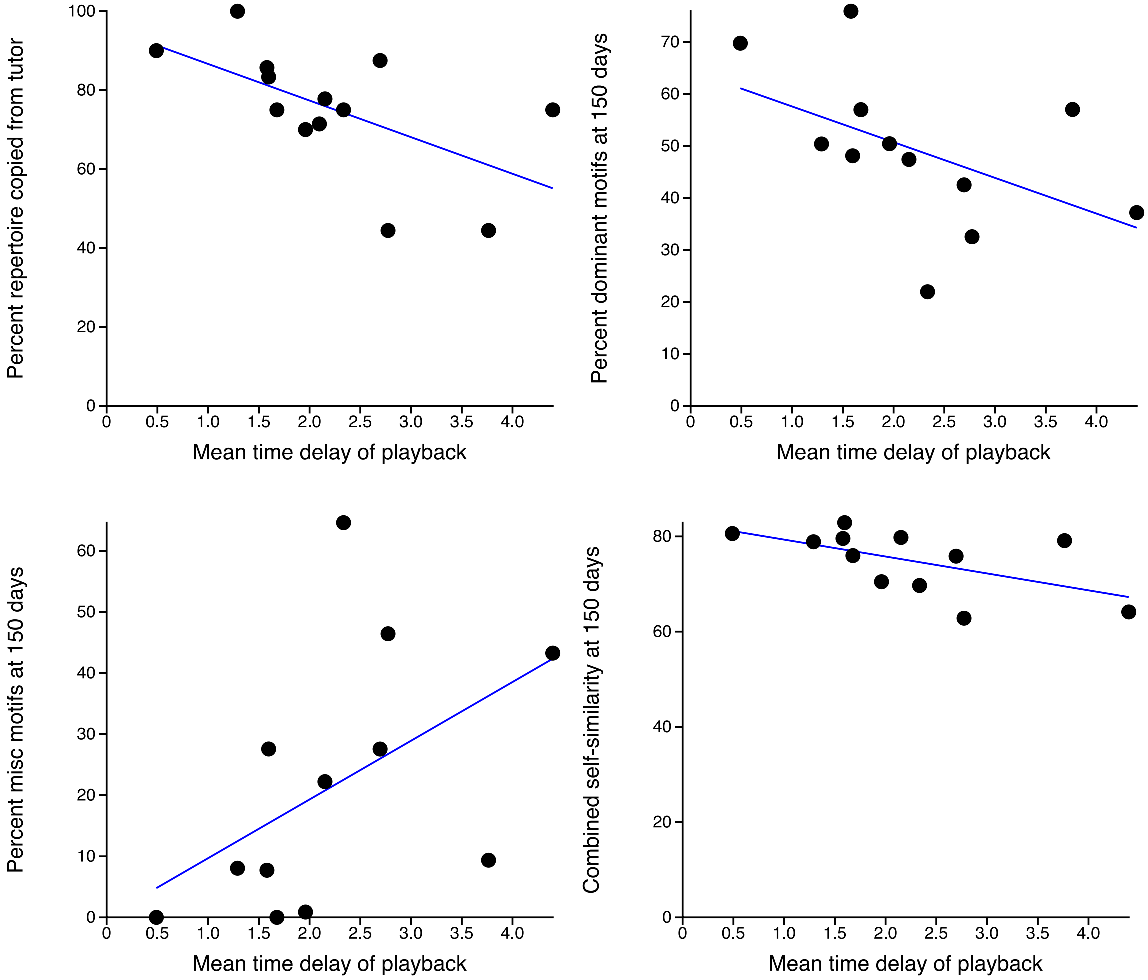

Figure 4.12 Longer time delay of playback delivery predicted the following song outcome measures: A: percentage of repertoire copied from the tutor measured at 120 days; B: percentage of dominant (most common) motif calculated at 150 days; C: percentage of miscellaneous motifs measured at 150 days; D: combined self-similarity score (self-similarity*accuracy*sequential match) measured at 150 days.

Relationship between song outcome measures

Correlation between song features at 90 days

To assess which song measures describe similar features of the final song outcome, we correlated song structure, tutor similarity, and self-similarity measures. The analysis revealed that the song structure is either stable or variable: a higher percentage of dominant motifs results in a lower proportion of miscellaneous motif types (r_s (13) = -0.879 $p < 0.001$, Figure 4.13A).

Birds singing less variable song structure also produced songs with increased self-similarity, as higher proportion of dominant motifs correlated positively with the combined self-similarity score, r_s (13) = 0.588, $p = 0.035$. Subjects with more variable song structure (higher proportion of miscellaneous motifs) learned the tutor song sequence less well (sequential match with tutor, r_s (13) = -0.643, $p = 0.018$). Moreover, birds singing a more variable song structure (proportion of miscellaneous motifs) also sang a less consistent song sequence (self-sequential match, r_s (13) = -0.709, $p = 0.007$) and less consistent overall song (combined self-similarity score, r_s (13) = -0.676, $p = 0.011$) (Figure 4.13 B, C, D, E).

To test the relationship between the different similarity measures to the tutor, we cross-correlated the percent similarity, accuracy and sequential scores. Juveniles who learned a larger portion of the father's song tended to produce it more accurately, but percent similarity to tutor and accuracy were not significantly correlated (r_s (13) = 0.511, $p = 0.074$). Moreover, neither percent similarity nor accuracy significantly correlated with sequential match measured between juvenile and tutor (*p*s > 0.3).

The percent self-similarity and accuracy scores were strongly correlated (r_s (13) = 0.978, $p < 0.001$) (Figure 4.13F). Thus, birds that sang larger portions of their songs more consistently produced those portions more accurately as well. Percent self-similarity and accuracy scores did not correlate with self-sequential match (*ps* > 0.3).

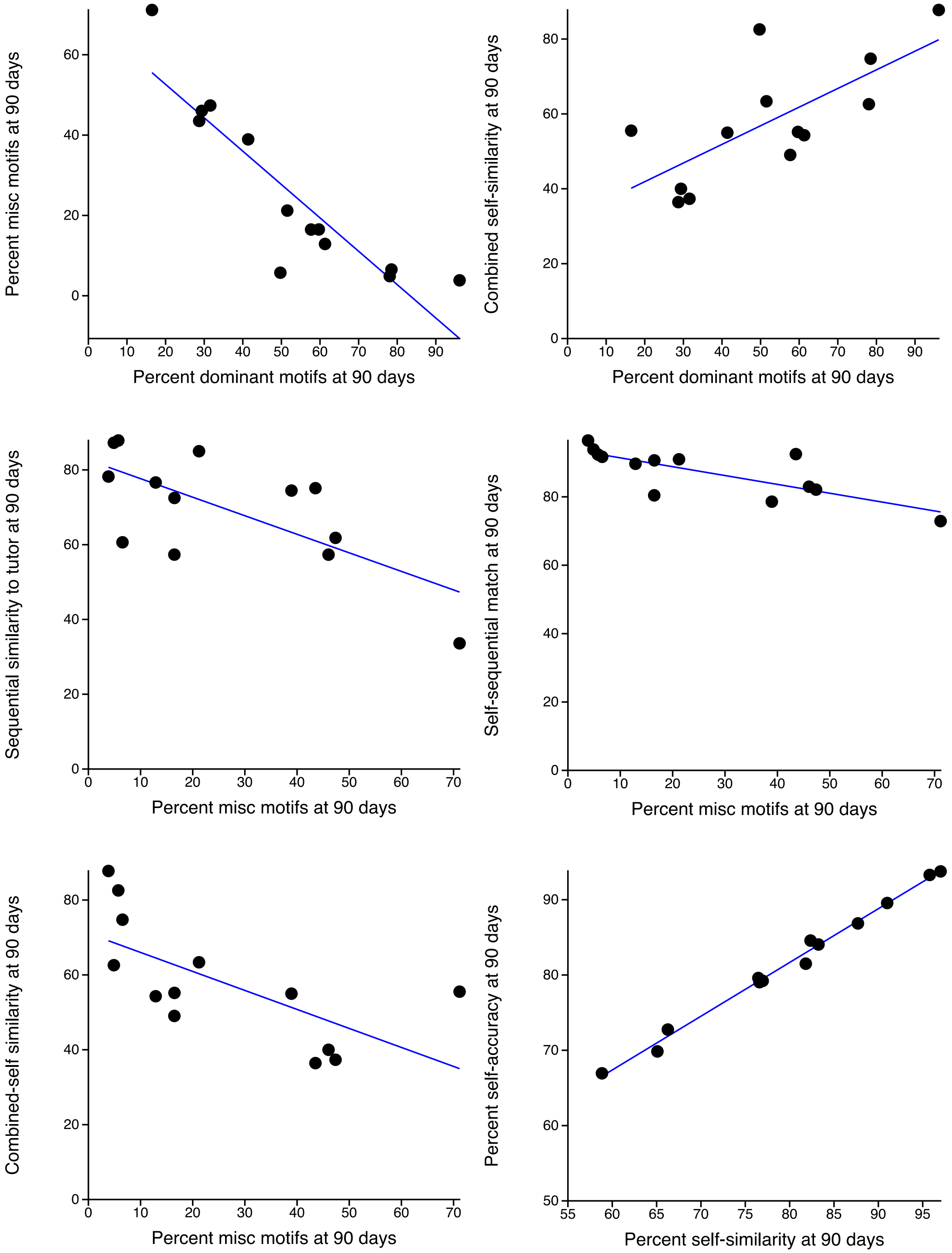

Figure 4.13 Relations between song outcome measures recorded at 90 days of age. Correlation between A: the percentage of dominant (most common) and miscellaneous motif types; B: the percentage of dominant motif and combined self-similarity measure (self-similarity*accuracy*sequential match); C: the percentage of miscellaneous motif types and sequential similarity to tutor D: the percentage of miscellaneous motif types and self-sequential match; E: the percentage of miscellaneous motif types and the combined self-similarity measure; F: percent self-similarity and self-accuracy measures.

The correlations revealed a trade-off between reproducing the tutor song and singing consistency. Juveniles who learned a larger portion of their tutor song produced songs with less overall consistency, as increased percent similarity to the tutor tended to predict decreased combined self-similarity score, although the correlation was not significant (r_s (13) = -0.051, p = 0.074). Juveniles singing more accurate copies of the tutor song sang less consistently, as increased accuracy between tutor and juvenile was negatively correlated with self-similarity (r_s (13)= -0.637, p = 0.019) and self-accuracy (r_s (13) = -0.566, p = 0.044). Consequently, birds singing an overall more similar song to the tutor tended to sing an overall less consistent song (r_s (13) = -0.544, p = 0.055). Juveniles that learned the tutor song sequence more faithfully also produced a more consistent song sequence (r_s (13) = 0.709, p = 0.007) (Figure 4.14 A, B, C, D, E).

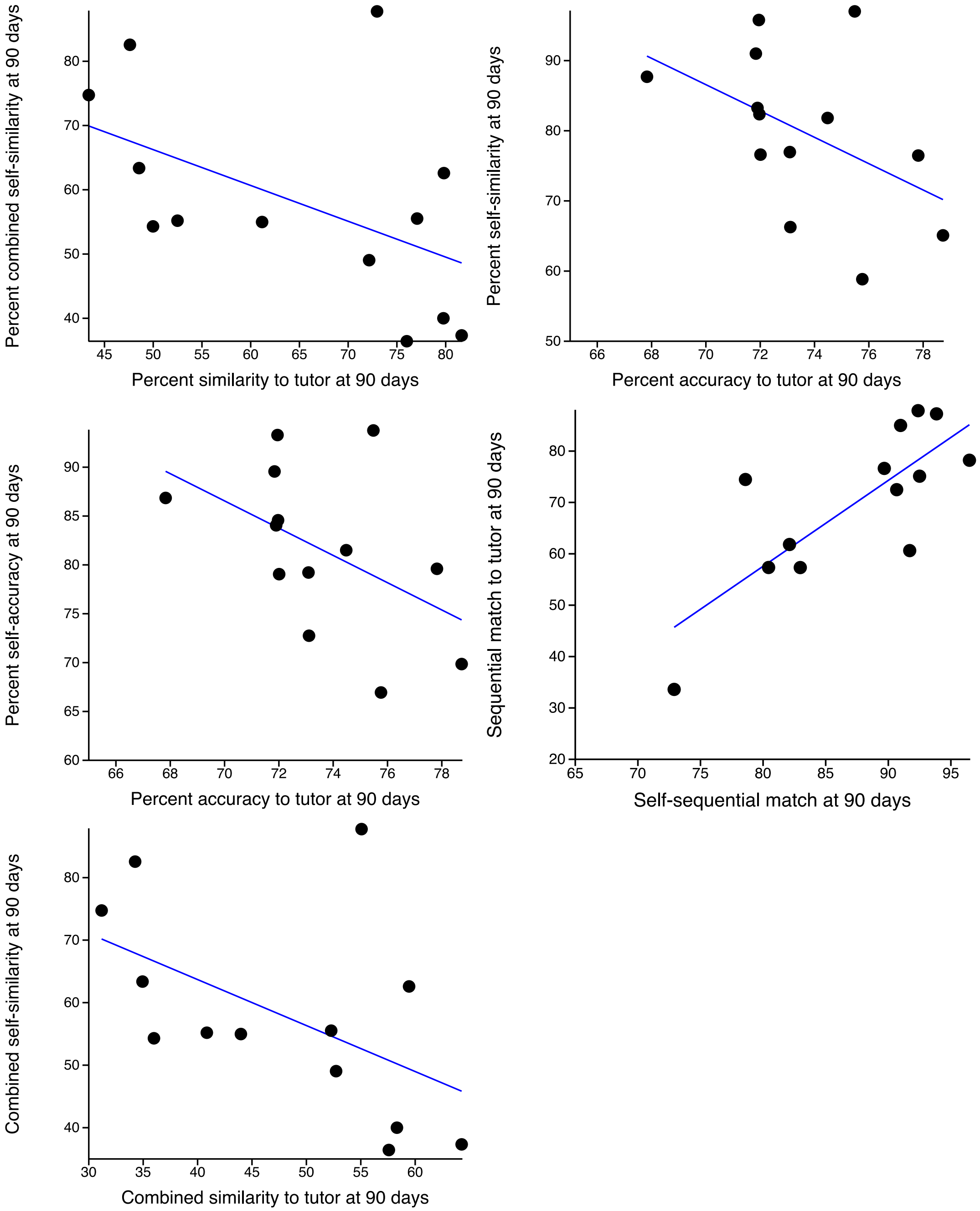

Figure 4.14 Relationship between tutor similarity and self-similarity measures. Correlation between A: percent similarity to tutor and combined self-similarity (self-similarity*accuracy*sequential match); B: accuracy to tutor and percent self-similarity; C: accuracy to tutor and self-accuracy; D: sequential similarity to tutor and self- sequential match, E: combined similarity score to tutor and combined self-similarity score.

Correlation between song outcome measures across recording times

To test stability in mature song we correlated song structure, similarity to tutor and self-similarity measures across the three recording times (90, 120, 150 days). None of the song structure measures (percent of the dominant, secondary and miscellaneous motifs) or self-similarity scores (percent similarity, accuracy, sequential match and their combined value) correlated significantly across the 90-, 120-, and 150-day recordings.

Percent similarity to tutor was positively correlated over time (between 90 and 120 days, r_s (13) = 0.797, $p = 0.001$; between 120 and 150 days, r_s (12) = 0.868, $p < 0.001$). Combined similarity score to tutor was also positively correlated over time (90 to 120 days, r_s (13) = 0.725, $p = 0.005$; 120 to 150 days, r_s (12) = 0.808, $p = 0.001$). Accuracy scores to tutor positively correlated between 120 and 150 days (r_s (12) = 0.571, $p = 0.041$). Sequential match between juvenile and tutor also correlated positively across all three recording times (between 90 and 120 days, r_s (13) = 0.874, $p < 0.001$; between 120 and 150 days, r_s (12) = 0.912, $p < 0.001$).

Discussion

We found that using a strictly controlled vocal environment to mimic naturalistic interactions influenced singing behavior during training and the ultimate song outcome at maturation. During the active playback period of the training sessions, contingent subjects sang shorter and less frequent songs than the yoked controls. Once the playback period ended, overall singing activity of contingent juveniles increased. This extinction burst in singing suggests that juveniles recognized a link between their own action and the response from their vocal environment. Moreover, the magnitude of the extinction burst predicted learning, as juveniles with higher overall increase in their song activity learned the playback stimulus better. Timing of playback exposure also influenced the structure of the mature song: contingent subjects developed a more stable song pattern by singing a more consistent syllable sequence than their matched yoked partners. The time delay of contingent feedback appeared to have important effects on learning, as longer delay of playback exposure was related to decreased learning from the stimuli. These observations support our initial prediction that contingent interactions organized by vocal precursors influence song development.

Our findings have multiple implications for understanding principles of song learning. The dominant model claims that songbirds imitate tutor song via auditory feedback in a two-phase process: memorization and practice, which are temporally distinct periods being weeks or months apart during development (Konishi, 1965; Marler, 1970; Marler, 1976). Learning is primarily instructive in nature: a passive sensory, memorization process is followed by song practice during which juveniles sing from memory to match the tutor song “template”. A more recently described alternative of the learning model incorporates a selective process. After the instructive memorization phase the social experience selectively shapes the song outcome during practice. This model of song acquisition has been frequently described in territorial species that use their song in distant communication. Immature songs gain functional significance during territory establishment, when vocal interactions determine which elements to keep in the final song from the pool of the already memorized, but not crystallized repertoire, a process called “action based learning” (Nelson & Marler, 1994; Nelson, 2000). Here, even passive overhearing or eavesdropping on vocal interactions facilitates learning (Beecher et al., 2007).

In contrast to the most frequently studied North American sparrows, the zebra finch is a gregarious species. Song is used within short distance and juveniles become sexually mature within 4 months (Zann, 1996). There is a substantial overlap between memorization and song practice (Slater et al., 1988). As a consequence, immature vocalizations play a different role during vocal development compared to territorial species. As supported by our findings, passive overhearing of song may not be an adequate stimulus to develop an appropriate singing pattern in the zebra finch. Our results suggest that juveniles use their vocalizations to control feedback from their environment, and use the feedback to modify their song outcome. Thus, learning is regulated by a dynamic bidirectional interaction between an organisms and its environment, and juveniles play an active role in their vocal development as both senders and receivers. When juveniles are denied from the active contribution in learning, the consequences are reflected in their song outcome.

What mechanisms are responsible for the differences detected in the song of contingent and yoked subjects? Three possibilities are considered below: associative learning of singing and receiving social responses, changes in the motivational state of the singer while singing, and reward based learning from social feedback. These mechanisms are not mutually exclusive and may be working in combination to mediate socially guided song learning.

Associative learning may play an important role in social learning of song structure. Actions of conspecifics occurring at the same time as singing might be linked together (Schultz & Dickinson, 2000), such that juveniles may learn that via singing they can elicit changes in their social environment. This is supported by the observation that during the playback phase, contingent subjects sang shorter songs and less frequently. Once playback ceased in the final baseline period, singing activity of contingent subjects increased compared to the yoked birds. The change in song rate between the active and final baseline periods was most apparent during the middle of the training period (joint training session 3 and 4) (Figure 4.3), suggesting that juveniles had built expectations about vocalizations operating on feedback. This pattern is consistent with animals learning associations between a stimulus and an outcome (e.g., pressing a lever leading to a food reward in rats). Such associative learning results in increased response activity when reward ceases (called an extinction burst), due to increased frustration or arousal (Tarabulsy, Tessier, & Kappas, 1996).

Links between social and motivational circuitry indicate that directed singing is a socially motivated act. In the zebra finch, activation of the social behavior network after social interaction indicates that conspecifics are assigned positive valence (Goodson, Evans, Lindberg & Allen, 2005; Goodson & Thompson, 2010). Therefore it is likely that young birds were motivated to work for social response (tutor song) because that song was associated with the presence of a conspecific and thus had positive valence. Not only social feedback, but also the act of vocalizing may alter the internal state of the actor. In ring doves, female “nest coo” vocalizations are self-stimulating and start a neuroendocrine cascade which changes physiological status (Cheng, 1992). Brain structures of the limbic system and the song system are interconnected. Consequently, singing may alter physiological state, and juveniles may become more aroused or be in a heightened state of attention. There is indirect evidence supporting the idea that the act of undirected singing, which is not used in courtship, is regulated by opioids, and as such may be linked to intrinsic reward. In the zebra finch, administration of high doses of the opioid antagonist naloxone dramatically reduces the production of song not intended toward other conspecifics (Khurshid, Jayaprakash, Hameed, Mohanasundaram & Iyengar, 2010). This is consistent with the assumption that opioids regulate undirected singing in the zebra finch. Presence of enkephalin opioids is also related to undirected singing in starlings in brain areas known to be related to motivation and reward in starlings (medial preoptic nucleus and ventral tegmental area) (Riters et al., 2005). Moreover, after undirected singing, males have decreased pain reactivity to a hot water bath, indirect evidence of increased opioid production, as opioids have analgesic effects (Kelm-Nelson, Stevenson & Riters, 2012). All of the above suggests that subjects immediately after singing may receive feedback at a time of an altered inner physiological state.

Contingent feedback may be perceived as reinforcement on immature vocal precursors, thus shaping vocal development. According to our findings, contingent playback changed not only the quantity of vocalizations during training, but also the quality of the ultimate song outcome. Our manipulation spanned 35 – 50 days of age, a period during which song memorization and practice overlap, and juvenile exposed to tutors during this period learned the song model more accurately compared to later in development (Eales, 1987b). Contingent playback therefore may influence both vocal production and the song representation at the same time. We also found that the extent of change in singing during the final baseline compared to the playback period correlated with increased similarity to tutor song in the mature song recordings, suggesting that heightened sensitivity to social contingency might influence memorization. Contingency therefore may influence learning as soon as a juvenile recognizes the temporal relation between his own actions and the contingent social feedback, and changes not only the quantity but also the quality of his vocalizations. The role of contingent feedback on juvenile vocalization in the zebra finch shows remarkable similarities to the vocal development of the brown-headed cowbirds, in which subsong and plastic song are modifiable by the listener’s social response. Female cowbirds react to immature vocal precursors with silent visual signals termed wing strokes, and social responsiveness of females shapes both the content and the potency of the final song outcome and affects the rate of song development (West & King, 1988; Smith et al., 2000). However, socially guided learning as described in the cowbird had not been considered to be a significant learning mechanism in the songbird realm. Brown-headed cowbirds are brood parasitic, and due to the species natural ecology vocal learning is expected to deviate from the traditional pathway of learning, as juveniles may be less dependent on conspecific song tutors.

However, the learning process we described in the zebra finch and the brown-headed cowbirds bears striking similarities to vocal development in human infants. Parents are responsive to infant babbling (Goldstein & West, 1999; Gros-Louis, West, Goldstein & King, 2006). Immature vocalizations have functional significance, as babies learn associations between their own vocalizations and responses from their social environment (Goldstein, et al., 2009), and modify their babbling accordingly (Goldstein et al., 2003; Goldstein & Schwade, 2008). Parallel learning mechanisms on the behavioral level between brown-headed cowbirds, human infants and the zebra finch suggest, that if the developmental niche of a young learner offers similar opportunities for learning, similar mechanisms may emerge, even in evolutionary distant taxa.

Our experiment is the first paradigm in which zebra finch juveniles are singing to produce tutor song, in contrast to earlier studies that used key pecking or perch landing to trigger song (ten Cate, 1991; Adret, 1993). Although earlier studies focused on similarity between juvenile and tutor song, and not on song pattern consistency, the findings are comparable. We have not found greater similarity to the playback stimulus between contingent and yoked subjects, which is consistent with several findings (Bolhuis, Van Mil & Houx, 1999; Houx & ten Cate, 1999a, b; Houx, ten Cate & Feuth, 2000). Others reported improved similarity and sequential match between models and pupils compared to passive tape exposure (Adret, 1993; Deregnaucourt et al., in press). Passive tutoring resulted in more consistent syllable transitions, as untutored birds produced a less stable sequential arrangement of syllables (Volman & Khanna, 1995). Because in our study contingent subjects sang less variable song pattern than the yoked controls, our results suggest that contingent feedback improved song pattern consistency further compared to passive song exposure.

As our manipulation was dependent on contingent subject's vocal activity, the amount of contingency, timing of feedback, and interruption of juvenile songs showed variability across training sessions. Interactions among these training parameters may have influenced the final song outcome, and may explain discrepancies from previous studies.

Learning was influenced by the proportion of songs that received contingent feedback. We aimed to provide contingent feedback on every instance of juvenile song. The manual delivery of playbacks resulted in about 78% contingency within the treatment group. Based on our observations of unmanipulated zebra finch families between 30 to 75 days of age, average levels of contingent feedback provided by the father was around 4 %, which correlated with higher fidelity of learning the tutor's song sequence (Menyhart et al., in prep). In numerous species, learning associations based on imperfect contingencies help to retain the learned pattern for a longer time when the contingency is removed, a phenomenon known as the partial reinforcement effect (Mackintosh, 1974). Our results suggest that too high contingency had adverse effect on learning the tutor song. Thus, to determine the ideal feedback rate for learning, it is would be informative to test a response rates approaching the naturally occurring levels of social contingency.

Length of delay of contingent feedback appeared to have important effects on learning. Longer delay of playback exposure was related to decreased learning from the stimuli. The time period in which the contingency of two events is most effectively associated is called the "contingency memory span" (Miller & Watson, 1979). The delay in the contingent treatment group ranged between 1.29 and 2.15 seconds. Perhaps feedback falling closer to the juvenile's own vocalization approaches the contingency memory span better. Juvenile zebra finch males generally learn better from operant conditioning paradigm in which tutor song is played back between 0.1-1.5 second after key-peck compared to passive tape exposure (Tchernichovski, Lints, Mitra & Nottebohm, 1999). The "best delay" for learning could therefore be under or around 1 second after the juvenile's own action.

Songs of contingent subjects were also more likely to be interrupted by the playback than the songs of yoked controls, and birds whose song was more often interrupted developed smaller proportion of miscellaneous motif types by 150 days of age. Temporal overlap in auditory signals is often interpreted to be an agonistic interaction in territorial songbirds (e.g. Todt and Naguib, 2000; Hall, Illes & Vehrencamp, 2006; Maynard, Ward, Doucet & Mennill, 2012). Although zebra finch is gregarious, the males defend nest sites, and sing frequently in their close proximity (Williams, 1990). The possibility that signal overlapping has functional significance in the zebra deserves further investigation.

The acoustic variability of the playback stimuli may have also played a role in song learning. All previous playback studies used a single song stimulus (Bolhuis et al., 1999; Houx and ten Cate, 1999a, b; Houx et al., 2000; Tchernichovski et al., 2001; Deregnaucourt et al., 2005; Phan et al., 2006; Deregnaucourt et al., in press). In contrast, we alternated the number of motif repetitions (4 or 3, respectively) to prevent early habituation, but playback stimuli were not otherwise varied. Repeated exposure to structured variability in a stimulus may facilitate learning via contingent feedback by enabling a powerful domain-general mechanism, called statistical learning. Statistical learning is based on recognizing high-probability patterns in structural information available in the continuous flow of a stimulus. Learning the co-occurrence and transitional probabilities among elements helps to segment continuous information into shorter chunks. Recombining those chunks into longer phrases is a helpful way to learn longer sequences and to learn diverse songs from multiple tutors. For example, the Bengalese finch copies chunks of songs from several tutors, and use statistical information to learn boundaries for song segmentation (Takahasi, Yamada & Okanoya, 2010). Moreover, white-crowned sparrows tutored with multiple phrase-pairs were able to assemble complete songs in a correct sequence, but those who were tutored with multiple but singly presented phrases failed to produce full songs (Rose et al., 2004). There are numerous indications that zebra finches are also sensitive to the structural variability in the songs they hear. Zebra finch song contains substantial variability (Sturdy et al., 1999; Helekar, Marsh, Viswanath & Rosenfield, 2000), and statistical regularities in the tutor song are mirrored in the song of their offspring (Menyhart et al., in prep.). Juveniles learn their song in chunks, and if housed socially, develop a hybrid song composed of chunks copied from multiple tutors (Williams, 1990; ten Cate and Slater, 1991; Williams & Staples, 1992).

The absence of increased imitation in the contingent treatment group was also consistent with findings in human infants, as babies can learn to produce phonological patterns from underlying regularities in caregivers’ speech, but they do not imitate the same phonemes as their caregivers produce (Goldstein & Schwade, 2008). Importantly, this learning does not occur when responses to babbling are unvaried and repetitive (Goldstein, Syal & Schwade, submitted). A subsequent study with more variable playback therefore could test the significance of socially facilitated statistical learning in the zebra finch.

Model abundance may also influence imitation of the tutor song. Unrestricted access to the song model decreases learning fidelity and ultimate motif duration. Even about 2 minutes of total song exposure (i.e. ~160 repetitions of the tutor song motif) hinders imitation, though the effects of model abundance on yoked subjects has not been tested (Tchernichovski et al., 1999). We exposed our subjects on average to about 40 playback songs, or about 140 repetitions of the tutor motif in each training session. We found a significant negative correlation between model abundance and learning accuracy in the yoked subjects, a relation that was absent in the contingent treatment group. Our results replicated the findings of a previous study in which yoked subjects showed less accurate imitation with increased exposure to the tutor song (Houx & ten Cate, 1999a). This phenomenon might be explained by decreased motivation to practice the model due to its reduced reward value as compared to contingently presented song. In addition, high exposure to noncontingent song may cause learned irrelevance and/or increased habituation.

Taking into account the various effects of variation in vocal activity of individual subjects, our results suggest that contingent feedback along with countersinging stabilized singing consistency, but time delay of playback may have reduced learning from the model. However, contingent versus yoked timing of song playback resulted in differences in ultimate song structure. The findings support our hypothesis that the temporal structure of naturalistic social interactions organized by juvenile subsong is a potent force in song learning.

The strongest effect of contingent feedback was on song consistency. What is the function of song stability in female preference formation and maintenance? When courting females, zebra finch males produce a so called directed song with more consistent syllable structure compared to undirected singing not aimed toward any conspecifics. Females strongly prefer directed songs over undirected ones, and variability in syllable pitch predicts female preferences (Woolley & Doupe, 2008). We do not know, however, whether sequence consistency is a feature into which females zebra finches pay attention during mate selection as preference for sequence consistency has not been tested systematically in the zebra finch. If the consistency of song syntax is an important component of female preference, then consistency may be a more central and demanding component of song production then its fidelity to a tutor's song. This is indirectly supported by a brood size manipulation treatment inducing developmental stress, which as a result, altered syntax learning and song duration consistency among treatment groups (Holveck, de Castro, Lachlan, ten Cate & Riebel, 2008).

Learning under natural circumstances offers opportunities for multimodal (e.g., visual or tactile) feedback also from non-singing family members (e.g. females) in the zebra finch (Jones & Slater, 1993; Adret, 2004; Williams, 2004). In observations on interactions within zebra finch families, we have found evidence that interactions with mothers also affect song learning (Menyhart et al., in prep). Thus, vocal learning and development are based on a network of social interactions far more complex –and interesting—than originally believed.

REFERENCES

Adret, P. (1993). Operant-conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour, 46*(1), 149-59. doi: 10.1006/anbe.1993.1170

Adret, P. (2004). Vocal imitation in blindfolded zebra finches (Taeniopygia guttata) is facilitated in the presence of a non-singing conspecific female. *Journal of Ethology, 22*(1), 29-35. doi: 10.1007/s10164-003-0094-y

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution, 20*(3), 143-149. doi: 10.1016/j.tree.2005.01.004

Beecher, M. D., Burt, J. M., O'Loghlen, A. L., Templeton, C. N., & Campbell, S. E. (2007). Bird song learning in an eavesdropping context. *Animal Behaviour, 73,* 929-935. doi: 10.1016/j.anbehav.2006.10.013

Bohner, J. (1983). Song learning in the zebra finch (taeniopygia-guttata) - selectivity in the choice of a tutor and accuracy of song copies. *Animal Behaviour, 31*(1), 231-237. doi: 10.1016/s0003-3472(83)80193-6

Bolhuis, J. J., Van Mil, D. P., & Houx, B. B. (1999). Song learning with audiovisual compound stimuli in zebra finches. *Animal Behaviour, 58*(6)1285-1292. doi: 10.1006/anbe.1999.1266

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature, 417*(6886), 351-358. doi: 10.1038/417351a

Cheng, M. F. (1992). For whom does the female dove coo - a case for the role of vocal self-stimulation. *Animal Behaviour, 43*(6), 1035-1044. doi: 10.1016/s0003-3472(06)80016-3

Clayton, N. S. (1987b). Song tutor choice in zebra finches. *Animal Behaviour, 35*(3), 714-721. doi: 10.1016/s0003-3472(87)80107-0

Clayton, N. S. (1988). Song learning and mate choice in estrildid finches raised by 2 species. *Animal Behaviour, 36*(6), 1589-1600. doi: 10.1016/s0003-3472(88)80101-5

Deregnaucourt, S. (2011). Birdsong learning in the laboratory, with especial reference to the song of the Zebra Finch (Taeniopygia guttata). *Interaction Studies, 12*(2), 324-350. doi: 10.1075/is.12.2.07der

Deregnaucourt, S., Mitra, P. P., Feher, O., Pytte, C., & Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. *Nature, 433*(7027), 710-716. doi: 10.1038/nature03275

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*(1), 567-631. doi: 10.1146/annurev.neuro.22.1.567

Eales, L. A. (1985). Song learning in zebra finches - some effects of song model availability on what is learnt and when. *Animal Behaviour, 33*(4), 1293-1300. doi: 10.1016/s0003-3472(85)80189-5

Eales, L. A. (1987a). Do zebra finch males that have been raised by another species still tend to select a conspecific song tutor. *Animal Behaviour, 35*(5), 1347-1355. doi: 10.1016/s0003-3472(87)80007-6

Eales, L. A. (1987b). Song learning in female-raised zebra finches - another look at the sensitive phase. *Animal Behaviour, 35*(5), 1356-1365. doi: 10.1016/s0003-3472(87)80008-8

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour, 37*(3), 507-508. doi: 10.1016/0003-3472(89)90097-3

Field, A. P. (2009) *Discovering statistics using SPSS*. Los Angeles: SAGE.

Funabiki, Y., & Funabiki, K. (2009). Factors Limiting Song Acquisition in Adult Zebra Finches. *Developmental Neurobiology, 69*(11), 752-759. doi: 10.1002/dneu.20738

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 8030-8035. 8030-8035. doi: 10.1073/pnas.1332441100

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515-523. doi: 10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The Value of Vocalizing: Five-Month-Old Infants Associate Their Own Noncry Vocalizations With Responses From Caregivers. *Child Development, 80*(3), 636-644. doi: 10.1111/j.1467-8624.2009.01287.x

Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience.* Oxford University Press.

Goldstein, M. H., Syal, S., & Schwade, J. A. (submitted). Socially guided statistical learning drives the emergence of speech in infants.

Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*(1), 52-58. doi: 10.1037/0735-7036.113.1.52

Goodson, J. L., Evans, A. K., Lindberg, L., & Allen, C. D. (2005). Neuro-evolutionary patterning of sociality. *Proceedings of the Royal Society B-Biological Sciences, 272*(1560). doi: 10.1098/rspb.2004.2892

Goodson, J. L., & Thompson, R. R. (2010). Nonapeptide mechanisms of social cognition, behavior and species-specific social systems. *Current Opinion in Neurobiology, 20*(6), 784-794. doi: 10.1016/j.conb.2010.08.020

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*(6), 509-516. doi: 10.1177/0165025406071914

Hall, M. L., Illes, A., & Vehrencamp, S. L. (2006). Overlapping signals in banded wrens: long-term effects of prior experience on males and females. *Behavioral Ecology, 17*(2), 260-269. doi: 10.1093/beheco.arj022

Helekar, S. A., Marsh, S., Viswanath, N. S., & Rosenfield, D. B. (2000). Acoustic pattern variations in the female-directed birdsongs of a colony of laboratory-bred zebra finches. *Behavioural Processes*, *49*(2), 99-110. doi: 10.1016/s0376-6357(00)00081-4

Holveck, M. J., de Castro, A. C. V., Lachlan, R. F., ten Cate, C., & Riebel, K. (2008). Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behavioral Ecology, 19*(6), 1267-1281. doi: 10.1093/beheco/arn078

Houx, B. B., & ten Cate, C. (1999a). Do stimulus-stimulus contingencies affect song learning in zebra finches (Taeniopygia guttata)? *Journal of Comparative Psychology, 113*(3), 235. doi: 10.1037/0735-7036.113.3.235

Houx, B. B., & ten Cate, C. (1999b). Song learning from playback in zebra finches: is there an effect of operant contingency? *Animal Behaviour, 57*(4), 837-845.

Houx, B. B., ten Cate, C., & Feuth, E. (2000). Variations in zebra finch song copying: An examination of the relationship with tutor song quality and pupil behaviour. *Behaviour, 137*(10), 1377-1389. doi: 10.1163/156853900501980

Howell, D. C. (2010). *Statistical methods for psychology* (8th ed.). Belmont, CA: Cengage Wadsworth. ISBN-13: 978-1-111-83548-4

Immelmann, K. (1969). Song development in the zebra finch and other Estrildid finches. In R. A. Hinde (Ed.), *Bird Vocalizations* (pp. 61-74). Cambridge: Cambridge University Press.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. *Journal of Ornithology, 148,* S35-S44. doi: 10.1007/s10336-007-0243-0

Jones, A. E., & Slater, P. J. B. (1993). Do young male zebra finches prefer to learn songs that are familiar to females with which they are housed. *Animal Behaviour, 46*(3), 616-617. doi: 10.1006/anbe.1993.1233

Jones, A. E., & Slater, P. J. B. (1996). The role of aggression in song tutor choice in the zebra finch: Cause or effect? *Behaviour,133,* 103-115. doi: 10.1163/156853996x00053

Kelm-Nelson, C. A., Stevenson, S. A., & Riters, L. V. (2012). Context-Dependent Links between Song Production and Opioid-Mediated Analgesia in Male European Starlings (Sturnus vulgaris). *Plos One, 7*(10), e46721. doi: 10.1371/journal.pone.0046721

Khurshid, N., Jayaprakash, N., Hameed, L. S., Mohanasundaram, S., & Iyengar, S. (2010). Opioid modulation of song in male zebra finches (Taenopygia guttata). *Behavioural Brain Research, 208*(2), 359-370. doi: 10.1016/j.bbr.2009.12.003

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: Parallels to human speech development. *Ethology, 111*(1), 101-117. doi: 10.1111/j.1439-0310.2004.01039.x

Kolodny O., Lotem A., Edelman S., (2011). Learning a graph-structured generative probabilistic grammar of linguistic experience (Proc. Cog. Sci. Society, Boston).

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift fur Tierpsychologie, 22*(7), 770-783.

Lerman, D. C., & Iwata, B. A. (1995). Prevalence of the extinction burst and its attenuation during treatment. *Journal of Applied Behavior Analysis, 28*(1), 93-94. doi: 10.1901/jaba.1995.28-93

London, S. E., & Clayton, D. F. (2008). Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience, 11*(5), 579-586. doi: 10.1038/nn.2103

Mackintosh, N. J. (1974). The psychology of animal learning. London: Academic Press.

Marler, P. (1970). A comparative approach to vocal learning - song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology, 71*(2p2),1. doi: 10.1037/h0029144

Marler, P. (1976). Sensory templates in species specific behavior. *Fentress, John C. (Ed.). Simpler Networks and Behavior. X+403p. Illus. Sinauer Associates, Inc.: Sunderland, Mass., USA. Isbn 0-87893-182-1*, 314-329.

Marler, P., & Peters, S. (1981). Sparrows learn adult song and more from memory. *Science, 213*(4509), 780-782. doi: 10.1126/science.213.4509.780

Marler, P., & Peters, S. (1987). A sensitive period for song acquisition in the song sparrow, Melospiza-melodia - a case of age-limited learning. *Ethology, 76*(2), 89-100.

Marler, P., & Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science, 146*(365). doi: 10.1126/science.146.3650.1483

Maynard, D. F., Ward, K. A. A., Doucet, S. M., & Mennill, D. J. (2012). Calling in an acoustically competitive environment: duetting male long-tailed manakins avoid overlapping neighbours but not playback-simulated rivals. *Animal Behaviour, 84*(3), 563-573. doi: 10.1016/j.anbehav.2012.06.008

Miller, W. S. & Watson, J. S. (1979). The effect of delayed feedback on infant learning reexamined. *Child Development*, 50, 747-751

Nelson, D. A. (2000). Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour, 60*, 887-898. doi: 10.1006/anbe.2000.1560

Nelson, D. A., & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences of the United States of America, 91*(22), 10498-10501. doi: 10.1073/pnas.91.22.10498

Oller, D. K. (2000). *The Emergence of the Speech Capacity*. Mahwah, NJ: Erlbaum.

Phan, M. L., Pytte, C. L., & Vicario, D. S. (2006). Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proceedings of the National Academy of Sciences of the United States of America, 103*(4), 1088-1093. doi: 10.1073/pna5.0510136103

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology, 93*(2), 260-277. doi: 10.1037/h0077553

Riters, L. V., Schroeder, M. B., Auger, C. J., Eens, M., Pinxten, R., & Ball, G. F. (2005). Evidence for opioid involvement in the regulation of song production in male European starlings (Sturnus vulgaris). *Behavioral Neuroscience, 119*(1), 245. doi: 10.1037/0735-7044.119.1.245

Roper, A., & Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology, 112*(5), 458-470. doi: 10.1111/j.1439-0310.2005.01169.x

Rose, G. J., Goller, F., Gritton, H. J., Plamondon, S. L., Baugh, A. T., & Cooper, B. G. (2004). Species-typical songs in white-crowned sparrows tutored with only phrase pairs. *Nature, 432*(7018), 753-758. doi: 10.1038/nature02992

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23*(1), 473-500. doi: 10.1146/annurev.neuro.23.1.473

Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches (Taeniopygia guttata): progress and prospects. *Advances in the Study of Behavior, 18,*1.

Smith, V. A., King, A. P., & West, M. J. (2000). A role of her own: female cowbirds, Molothrus ater, influence the development and outcome of song learning. *Animal Behaviour, 60*, 599-609. doi: 10.1006/anbe.2000.1531

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (1999). Note types, harmonic structure, and note order in the songs of zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology, 113*(2), 194-203. doi: 10.1037/0735-7036.113.2.194

Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and Prosodic Cues for Song Segmentation Learning by Bengalese Finches (Lonchura striata var. domestica). *Ethology, 116*(6), 481-489. doi: 10.1111/j.1439-0310.2010.01772.x

Tarabulsy, G. M., Tessier, R., & Kappas, A. (1996). Contingency detection and the contingent organization of behavior in interactions: Implications for socioemotional development in infancy. *Psychological Bulletin, 120*(1), 25-41. doi: 10.1037/0033-2909.120.1.25

Tchernichovski, O., Lints, T., Mitra, P. P., & Nottebohm, F. (1999). Vocal imitation in zebra finches is inversely related to model abundance. *Proceedings of the National Academy of Sciences of the United States of America, 96*(22), 12901-12904. doi: 10.1073/pnas.96.22.12901

Tchernichovski, O., & Nottebohm, F. (1998). Social inhibition of song imitation among sibling male zebra finches. *Proceedings of the National Academy of Sciences of the United States of America, 95*(15), 8951-8956. doi: 10.1073/pnas.95.15.8951

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour, 59*, 1167-1176. doi: 10.1006/anbe.1999.1416

ten Cate, C. (1991). Behavior-contingent exposure to taped song and zebra finch song learning. *Animal Behaviour, 42,*857-859.

ten Cate, C., & Slater, P. J. B. (1991). Song learning in zebra finches - how are elements from 2 tutors integrated. *Animal Behaviour, 42*(1), 150-152.

Volman, S. F., & Khanna, H. (1995). Convergence of untutored song in group-reared zebra finches (taeniopygia-guttata). *Journal of Comparative Psychology, 109*(3), 211. doi: 10.1037//0735-7036.109.3.211

West, M. J., & King, A. P. (1988). Female visual-displays affect the development of male song in the cowbird. *Nature, 334*(6179), 244-246. doi: 10.1038/334244a0

Williams, H. (1990). Models for song learning in the zebra finch - fathers or others. *Animal Behaviour, 39*, 745-757. doi: 10.1016/s0003-3472(05)80386-0

Williams, H. (2004). Birdsong and singing behavior. In H. P. M. P. Zeigler (Ed.), *Behavioral Neurobiology of Birdsong* (Vol. 1016, pp. 1-30).

Williams, H., & Staples, K. (1992). Syllable chunking in zebra finch (taeniopygia-guttata) song. *Journal of Comparative Psychology, 106*(3), 278-286. doi: 10.1037/0735-7036.106.3.278

Woolley, S. C., & Doupe, A. J. (2008). Social context - induced song variation affects female behavior and gene expression. *Plos Biology, 6*(3), e62. doi: 0.1371/journal.pbio.0060062

Zann, R. (1990). Song and call learning in wild zebra finches in south-east australia. *Animal Behaviour, 40(*5), 811-828. doi: 10.1016/s0003-3472(05)80982-0

Zann, R. A. (1996). *The Zebra Finch: Synthesis of Field and Laboratory Studies*. New York: Oxford University Press.

CHAPTER 5

CONCLUSIONS

In this thesis, I examined mechanisms of song learning in the zebra finch from an ecological perspective. The results suggest that when studied in more naturalistic settings, song learning in the zebra finch is a more multi-faceted process than previously thought. Contrary to prevailing assumptions, song learning does not appear to progress via simple imitation. Responsiveness of the social environment and variability in the acoustic input predict mature song outcome. The studies included in this dissertation only begin to unpack the complex nature of vocal ontogeny, but the results generate numerous hypotheses for future investigations.

In *Chapter 2*, I observed the microstructure of social interactions organized around juvenile vocalizations. I found that multisensory stimulation from both parents organized around juvenile immature song predicts song similarity between juveniles and their fathers.

The results of *Chapter 3* further support that learning is more than imitation, as juveniles learned the statistical structure of the tutor song. Song structures analyzed as finite-state grammars were more similar between fathers and sons compared to pairs of unrelated males, even without syllable-level correspondence among individuals.

Chapter 4 used an operant paradigm to test the efficacy of tutor song playback contingent on immature juvenile song and found that contingency changed the way juvenile song is learned compared to yoked controls.

Chapter 2: Socially guided song learning in the zebra finch: effects at multiple time scales

Socially guided learning influences vocal ontogeny in both brown-headed cowbirds (West & King, 1988) and human infants (Goldstein, King & West, 2003), where feedback from social companions facilitates developmentally more advanced vocalizations. My observations indicate similar mechanisms at work in the zebra finch. Increased amounts of contingent parental feedback predicted learning of the paternal song sequence, but when juveniles sang more before the male sang, they developed a less accurate copy of the father's song. Adult female fluff-ups around juvenile song were positively correlated with sequence learning, and wing strokes were linked to syllable pitch differences, as syllables after a wing stroke were closer to the father's pitch compared to control song regions not eliciting female responses.

Although both male and female feedback occurred infrequently, it facilitated learning differences among juveniles. The impact of covert and rare behaviors suggests that although male siblings from the same family were confined to a restricted environment living with the same parents, they were situated in distinct developmental niches. Thus the singing activity of the juveniles may differentiate their learning opportunities, resulting in an altered pace of development. Moreover, parental responsiveness had a substantial range, as for example vocal contingency under natural circumstances ranged between 1 - 9%. Does altered parental sensitivity canalize learning in juveniles early in development? In human infants, maternal responsiveness may exert lasting effects on cognitive and social development (e.g. Bornstein & Tamis-LeMonda, 2006). Variability in both juvenile activity and parental responsiveness deserves further attention in more controlled environments to explore their full potential in vocal ontogeny.

There are numerous questions yet unanswered about the time course of learning as song development is influenced by social interactions between juveniles and parents. At this point, I have obtained an overall assessment of parental impact for the entire observation period. I do not know yet whether parents scaffold the developmental process by adjusting their feedback based on the juvenile's current performance. A more complete understanding of the process of development, including qualitative and quantitative changes in juvenile song output over time and their relation to parental feedback, awaits further analyses. Moreover, as both contingent male songs and female fluff-ups predicted the same similarity measure, it raises questions as to whether these behaviors are related. Are female fluff-ups prompting juvenile or adult males to sing? Microanalysis of real-time interactions among triads of individuals (father-mother-juvenile male) may reveal who initiates interactions.

The results indicate that studies of song learning should consider the role of females. I only quantified responses from adult females. However, all but one family contained female siblings, whose contribution to the learning process is yet unknown. In brown-headed cowbirds, juvenile females show less selectivity than adult females in both in their responsiveness to immature or adult song (White, Gros-Louis, King & West, 2006) and their social interactions (Miller, King & West, 2008); their preference changes with experience (West, King, White, Gros-Louis & Freed-Brown, 2006). In the zebra finch, females learn to discriminate high from low-quality songs (Lauay, Gerlach, Adkins-Regan & DeVoogd, 2004). Juvenile females develop preference for the song they are exposed to during development (Miller, 1979; Riebel, 2000, 2003) and also for their mate's song (Woolley & Doupe, 2008). Rearing conditions, such as brood-size, alter later song preference: females from large broods ("low-quality") prefer the song of “low-quality” males (Holveck & Riebel, 2010). The development of female responsiveness as a function of the social environment deserves further attention.

In sum, the study of social influences on vocal ontogeny in in the zebra finch holds great promise. In this species, it should be possible to determine behavioral and neural mechanisms by which social reinforcement has the potential to shape ultimate song outcome, amplifying mechanisms that have been described in brown-headed cowbirds and human infants.

Chapter 3: Juvenile zebra finches learn the underlying statistical regularities in their fathers’ song.

Results of *Chapter 3* revealed that song syntax has functional significance, as finite-state grammars extracted from the mature song were more similar between fathers and sons compared to unrelated individuals. Our findings are important for multiple reasons: first, our results further indicate that imitation is insufficient as the sole learning mechanism in the zebra finch, as it is not possible to imitate an underlying abstract structure. Second, the findings suggests a further parallel between humans and songbirds, as extracting abstract rules from the auditory input seems to play an important role during vocal development in both taxa. Third, our method based on graph-distances does not require a common labeling scheme across multiple individuals. Thus, extracting grammatical structure is more accessible for species with large repertoires, and for other types of hierarchically organized behaviors.

Taking together the results of *Chapter 2* and *3*, an important question emerges: how is statistical learning of sound sequences integrated into the larger pattern of social statistics that describe interaction patterns? A current synthesis across multiple disciplines provides a theoretical and computational framework to integrate learning mechanisms across space and time (Goldstein et al., 2010). Although this synthesis is elucidated for findings in human communication, similar mechanisms may be applicable to other vocal learners. The theoretical model predicts that statistical structure in the information stream is differentially appreciated if it is embedded in social context. Statistical learning and rule abstraction from the information stream will be most efficiently enhanced when information presented by the caregiver (such as in her speech) occurs contingently on the infant's action (such as on her babbling). In the zebra finch, clustering of family members based on song structure is somewhat noisy (e.g., Figure 3. 2), indicating that grammars are similar but not identical among fathers and juveniles. As in human speech, the information content of parental signal (adult male song) may affect juvenile memorization most when it is provided contingently on the juvenile male's song. Looking at real time changes in the syntax of juvenile male song immediately before and after adult male feedback could reveal another important parallel in learning mechanisms among humans and songbirds.

Chapter 4: Contingent playback upon juvenile vocalizations shapes final song production in the zebra finch

In *Chapter 4* I found that contingent interactions organized by juvenile vocalizations influenced song development. Mimicking naturalistic interactions by providing contingent song playback affected structural consistency of the ultimate song outcome: contingent subjects developed a more stable song pattern than their matched yoked partners, although subjects in both treatment groups learned the song model similarly well. Contingent subjects decreased both their song rate and song duration during the active period of training. Once training ended, an "extinction burst" in song activity correlated with the amount of information learned from the father. Moreover, longer playback delays correlated with decreased learning.

Comparing the observations in Chapter 2 to the outcome of the operant training in Chapter 4 suggests that contingency affects song organization in different ways. In the longitudinal study, contingent paternal song predicted increased sequence learning from the father. Although treatment groups learned the tutor song comparably well, contingent subjects developed a more consistent song structure during operant training. The major differences in contingency levels (4% mean contingency in the natural observation compared to 78 % in the operant study), structural variability in the auditory information (or lack of variability in the operant study), and unisensory (auditory) versus multimodal (auditory and visual) presentation of the song stimuli may account for the observed differences. In the operant study, juveniles who received a higher percentage of contingency tended to learn less from the father. Thus, in subsequent studies it would be informative to explore the effect of contingent feedback when given with frequencies that more closely resemble those in natural interactions. The operant paradigm also offers an effective tool for exploring the amount of contingency a well as the significance of statistical learning in social contexts.

There were unexpected findings related to measures of song similarity across both the longitudinal (*Chapter 2*) and the operant study (*Chapter 4*). In the longitudinal study, the sequential match measure seemed to be positively influenced by contingent paternal feedback; however accuracy of learning tended to decrease at the same time. In the operant study, juveniles who learned the tutor song more accurately sang with less overall consistency. The analogous findings across the two studies suggest compromises in learning, such that all aspects of the song cannot be learned equally well. The interplay between the various song outcome measures and their relationship to memory processes deserves further investigation.

Concluding thoughts

Learning an accurate tutor song sequence is affected by contingencies in the social environment. Song syntax is shaped by the structure of the auditory stimuli. Structural consistency is influenced by contingent operant playback. The findings of all three studies point toward the conclusion that song structure is modified by the social rearing environment. Other studies confirm that song structure is influenced by early rearing conditions: brood size manipulation influenced syntax and singing consistency, as males from large broods sang less consistent song syntax (Holveck, de Castro, Lachlan, ten Cate & Riebel, 2008). Similar results were obtained by manipulating food availability (Brumm, Zollinger & Slater, 2009), where juveniles exposed to nutritional stress copied the paternal song sequence less accurately.

The findings of this thesis raise the question whether learning performance of the song structure translates to selective advantage during mating. However, our knowledge about the role of song structure in mate choice in the zebra finch is limited. Untutored female zebra finches inherently respond more to complex auditory stimuli than to prototypical male song bouts (Vyas, Harding, Borg & Bogdan, 2009). When exposed to sequences of four song motifs, females prefer to listen to motifs in which syllable sequences were rearranged into four different strings rather than to four repetitions of the same motif (Collins, 1999). As the complexity of song stimuli in the above studies were artificially enhanced, it is hard to reconcile the above results with the variability found in natural songs. Additionally, zebra finch females exhibit assortative mating dependent on their rearing environment. Low quality females from large broods prefer the song of low quality males from large broods in operant preference testing paradigms (Holveck & Riebel, 2010). Low quality males sang with decreased consistency (Holveck et al., 2008) which indirectly indicates that consistency of song structure may play a role during mate selection, although the song characteristics underlying female choice during testing are not yet known. Moreover, females tutored with complex songs show later preference for complex songs compared to prototypical or long song bouts lacking complexity. However, when tutored with prototypical or long song bouts, they do not show increased preference for complex song sequences (Vyas et al., 2009).

If song structure is an important trait and can be shaped by interactions within the early rearing environment, socially guided learning in the zebra finch may have far reaching consequences from an evolutionary perspective. Brown-headed cowbird females from two distant populations have different preferences and influence not only the song content of the juvenile males but the rate of song ontogeny (King, West & Goldstein, 2005). Traits that are important in sexual selection, such as properties of the adult song, therefore may be shaped by behavioral interactions during development. Studies on sexual selection focus mainly on adult features that indicate the male’s quality (Andersson, 1994). This approach may underestimate the contribution of developmental processes to the final form of song. When the developmental origins of features under selection pressure are revealed, the entire system could be understood better. I propose an integration of developmental mechanisms with the current framework of behavioral ecology to get a fuller understanding of evolutionary processes. Thus, a comparative approach focusing on adult life history strategies as well as developmental mechanisms in species under different environmental pressure can be very informative.

Finally, as similar developmental mechanisms were observed in brown-headed cowbirds, human infants, and zebra finches, developing individuals from unrelated species exposed to similar forces may develop similar mechanisms. This idea is borrowed from S. J. Gould, who describes in “The Panda’s Thumb” that odd regularities observed in nature are subject to physics, and animals reach the same solution because it is the easiest way to reach the adaptation (Gould, 1992). Human infants, juvenile cowbirds, and zebra finches are all altricial and require caregivers to survive. As such, juveniles are immediately integrated into a social communication network that functions by certain rules. The job of the developing organisms is to extract information from the structure the natural environment provides. Thus, to understand mechanisms of development, juveniles need to be re-integrated with their natural social environment, and the entire system needs to be studied as a functional unit.

REFERENCES

Andersson, M. (1994). *Sexual selection*: Princeton University Press.

Bornstein, M. H., & Tamis-LeMonda, C. S. (2006). Maternal responsiveness and cognitive development in children. *New Directions for Child and Adolescent Development, 1989*(43), 49-61.

Brumm, H., Zollinger, S. A., & Slater, P. J. B. (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. *Behavioral Ecology and Sociobiology, 63*(9), 1387-1395. doi: 10.1007/s00265-009-0749-y

Collins, S. A. (1999). Is female preference for male repertoires due to sensory bias? *Proceedings of the Royal Society of London. Series B: Biological Sciences, 266*(1435), 2309-2314.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences, 14*(6), 249-258. doi: 10.1016/j.tics.2010.02.004

Gould, S. J. (1992). *The panda's thumb: More reflections in natural history*: WW Norton & company.

Holveck, M.-J., & Riebel, K. (2010). Low-quality females prefer low-quality males when choosing a mate. *Proceedings of the Royal Society B: Biological Sciences, 277*(1678), 153-160.

Holveck, M. J., de Castro, A. C. V., Lachlan, R. F., ten Cate, C., & Riebel, K. (2008). Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behavioral Ecology, 19*(6), 1267-1281. doi: 10.1093/beheco/arn078

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: Parallels to human speech development. *Ethology, 111*(1), 101-117. doi: 10.1111/j.1439-0310.2004.01039.x

Lauay, C., Gerlach, N. M., Adkins-Regan, E., & DeVoogd, T. J. (2004). Female zebra finches require early song exposure to prefer high-quality song as adults. *Animal Behaviour, 68*(6), 1249-1255.

Miller, D. B. (1979). Long-term recognition of fathers song by female zebra finches. *Nature, 280*(5721), 389-391. doi: 10.1038/280389a0

Miller, J. L., King, A. P., & West, M. J. (2008). Female social networks influence male vocal development in brown-headed cowbirds, Molothrus ater. *Animal Behaviour, 76*(3), 931-941.

Riebel, K. (2000). Early exposure leads to repeatable preferences for male song in female zebra finches. *Proceedings of the Royal Society B-Biological Sciences, 267*(1461), 2553-2558. doi: 10.1098/rspb.2000.1320

Riebel, K. (2003). The "mute" sex revisited: Vocal production and perception learning in female songbirds. *Advances in the Study of Behavior, Vol 33, 33*, 49-86. doi: 10.1016/s0065-3454(03)33002-5

Vyas, A., Harding, C., Borg, L., & Bogdan, D. (2009). Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Hormones and Behavior, 55*(1), 50-59. doi: 10.1016/j.yhbeh.2008.08.005

West, M. J., & King, A. P. (1988). Female visual-displays affect the development of male song in the cowbird. *Nature, 334*(6179), 244-246. doi: 10.1038/334244a0

West, M. J., King, A. P., White, D. J., Gros-Louis, J., & Freed-Brown, G. (2006). The development of local song preferences in female cowbirds (Molothrus ater): Flock living stimulates learning. *Ethology, 112*(11), 1095-1107. doi: 10.1111/j.1439-0310.2006.01264.x

White, D. J., Gros-Louis, J., King, A. P., & West, M. J. (2006). A method to measure the development of song preferences in female cowbirds, Molothrus ater. *Animal Behaviour, 72,* 181-188. doi: 10.1016/j.anbehav.2006.01.008

Woolley, S. C., & Doupe, A. J. (2008). Social context–induced song variation affects female behavior and gene expression. *Plos Biology, 6*(3), e62.