Title
Perceptual mechanisms of pattern generalization in songbirds /

Author
Comins, Jordon Alexander

Publication Date
2014

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Perceptual mechanisms of pattern generalization in songbirds

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

in

Psychology

by

Jordan Alexander Comins

Committee in Charge:

Professor Timothy Gentner, Chair
Professor David Barner
Professor Sarah Creel
Professor Victor Ferreira
Professor Robert Kluender

2014

The Dissertation of Jordan Alexander Comins is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2014

DEDICATION

To all my peeps®.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply indebted to the many who have guided and supported me throughout the completion of this work. Without you, this would not have been possible.

I want to thank past and present members of the Gentner lab for their encouragement, feedback, and occasional commiseration: Dan, Emily, Jamie, Micah, Justin, Krista, Marvin, Margot, Patrick, Andrei and Becca.

Tim Gentner, my science mentor since 2007, thank you for all the time, training, patience and reinforcement you have shown. Thank you for providing me with access to so many wonderful research opportunities. I also want to thank the members of my committee, David Barner, Sarah Creel, Victor Ferreira and Robert Kluender.

To my family, Ric, Robin, Jamison, Scot and Natasha, for providing me the support I needed to succeed and for doing what families do best: inflate the ego of their kin!

Thank you to my friends in San Diego who made UCSD a second home: Jon, Kevin, Patrick, Scott, Evan, Sean, Nick, Erik, Tommy, Brian, Kendra, Ghassan and Youniss.

Stephanie, for all of your love and encouragement, I am eternally grateful.

Finally, in keeping with tradition, I want to thank 50p, my favorite rhesus macaque from when I worked in Cayo Santiago. Thank you for scratching me and making me think I was going to die from cercopithecine herpesvirus 1. You truly showed me that I do not want my brain to swell inside my skull.

This work was supported by NSF Graduate Research Fellowship 2011122846.

Chapter 1, in part, is material as it appears in Frontiers in Evolutionary Neuroscience, 2012. Kiggins, Justin K.; Comins, Jordan A.; Gentner, Timothy Q., Frontiers 2012. The material included only represents those portions of the article of which the dissertation author was the primary author.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Animal Cognition, 2014. Comins, Jordan A.; Gentner, Timothy Q., Springer 2014. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in Cognition, 2013. Comins, Jordan A.; Gentner, Timothy Q., Elsevier 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is currently material under review. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

Chapter 6, in part, is currently being prepared for submission for publication of the material. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

Chapter 7, in full, is currently material under review. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

VITA

EDUCATION

University of California, San Diego, La Jolla, CA 2011-2014
Psychology, Ph D

- National Science Foundation Graduate Research Fellow

Harvard University, Cambridge, MA 2005-2009
Human Evolutionary Biology, A.B.; Mind, Brain & Behavior, Cert.

- Mary Gordon Roberts Mind, Brain & Behavior Fellow
- Harvard College Research Program Award

MANUSCRIPTS

Comins JA, Gentner TQ (*Under Review*). Pattern-induced perceptual learning in songbirds.

Comins JA, Gentner TQ (*Under Review*). Temporal and syntactic processing by songbirds. *Current Opinion in Neurobiology.*

Comins JA, Gentner TQ (*In Prep*). An avian model of rule learning: The role of perceptual and decision-making processes..

Comins JA, Gentner TQ (*In Prep*). Male song syntax controls preference in female songbirds.

PUBLICATIONS

Comins JA, Gentner TQ (*2014*). Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. *Animal Cognition.*

Comins JA, Gentner TQ (2013). Perceptual categories enable pattern generalization in songbirds. *Cognition*, 128, 113-118.

Kiggins JT, Comins JA, Gentner TQ (2012). Targets for a comparative neurobiology of language. *Frontiers in Evolutionary Neuroscience*, 4, 1-13.

Comins JA, Russ BE, Humbert KA, Hauser MD (2011). Innovative coconut-opening in a semi free-ranging rhesus monkeys (*Macaca mulatta*): a case report on behavioral propensities. *Journal of Ethology*, 29, 187-189.

Hauser MD, Comins JA, Pytka LM, Cahill DP, Velez-Calderon S (2011). What experimental experience affects dogs' comprehension of human communicative actions? *Behavioural Processes*, 86, 7-20.

Comins JA, Gentner TQ (2010). Working memory for patterned sequences of auditory objects. *Cognition*, 117, 38-53.

Russ BE, Comins JA, Smith R, Hauser MD (2010). Recognizing and respecting claims over resources in free-ranging rhesus monkeys, *Macaca mulatta. Animal Behaviour*, 80, 563-569.

ABSTRACT OF THE DISSERTATION

Perceptual mechanisms of pattern generalization in songbirds

by

Jordan Alexander Comins

Doctor of Philosophy in Psychology

University of California, San Diego, 2014

Professor Timothy Gentner, Chair

The adaptive value of learning is constrained by generalization. Whether an organism successfully generalizes knowledge carries lofty implications for such things as navigating new paths, employing foraging innovations and identifying conspecifics. Generalizing knowledge of *patterned relationships*, however, is considered a hallmark of human faculties, such as music, analogical reasoning and language. In the auditory domain, generalizing knowledge of sound-based patterns is thought to underlie important computational processes necessary for language acquisition. In this dissertation, we take a comparative approach to consider the scope and constraints of auditory pattern generalization in a non-human species, the European starling songbird. We begin by establishing the ecological validity of studying pattern generalization in starlings by showing that the sequential patterning of components (motifs) in male starling songs carries information about likely relative fitness. Crucially, female starlings are sensitive to the sequential patterning of motifs in male song, evinced by strong preferences for some pattern types over others. From here, we provide definitive evidence that starlings can learn and generalize complex patterns built from natural auditory communication signals in controlled operant settings. We go on to identify key constraints underlying their ability to generalize patterns. These findings offer important insights into the relationship between the structure of a pattern and its constituents in real-world signals and how this relationship impacts the ability to generalize. We propose that the pattern generalization abilities of starlings resemble critical processes observed during language acquisition in human infants, such as auditory rule learning. We conclude that starlings constitute an important model organism for understanding how language-relevant learning modulates sensory systems to support language acquisition. We discuss our findings in light of recent advances in the behavioral and neural basis of temporal and syntactic processing of natural auditory communication signals in the starling.

CHAPTER ONE

Identifying targets for a comparative neurobiology of language

While most species communicate, only humans possess language. In recent years, biologists and psychologists have fiercely debated the evolutionary origins of language – did it emerge as an adaptation in response to selection pressures or as an unintended byproduct, or exaptation, of other cognitive abilities? To address these questions, empirical work on language evolution has largely focused on whether language exists along some continuum with other communication systems, or is categorically distinct (Hauser et al, 2002; Fitch & Hauser, 2004; Fitch et al, 2005; Pinker & Jackendoff, 2005; Jackendoff & Pinker, 2005; Margoliash & Nusbaum, 2009; Berwick et al, 2011; Terrace, 2011). The result has been a large-scale dichotomization of animal cognition into processes that are or are not "human-like" (Pinker & Jackendoff, 2005; Jackendoff & Pinker, 2005).

Rather than identifying putatively unique, language-relevant abilities and asking whether non-human animals show evidence for them, we aim to guide the conversation towards more fundamental auditory and memory challenges that many vocal communication systems share. By studying these more basic processes in fuller detail, from neurons to behavior, we can begin to construct an initial substrate for an ultimately more complete evolution and neurobiology of language. In other words, from a comparative perspective, we suggest that there is much more progress to be gained in terms of understanding the neurobiology and evolution of language by studying the mechanisms that are shared across taxa, than those that are putatively unique to humans.

In this dissertation, we consider how several fundamental aspects of auditory cognition such as serial expertise, categorization and abstraction, inform our knowledge of the comparative psychology of language. We focus our discussion of these processes in songbirds because this system provides the most well-developed model for the neurobiology of vocal learning. In this introduction, we use one of these aspects of auditory cognition, serial expertise, as a case study in how the comparative approach can highlight important similarities and, therefore, targets for future neurobiological investigations related to language.

Serial expertise in humans

Sensitivity to the ordering of linguistic units across time is vital to language comprehension. Indeed, in many languages, word order plays a vital component in assessing grammaticality (e.g., English) while in others syllable order serves an important role in phonology as in determining stress (e.g., Polish). As such, a requisite capacity for language is knowledge of the serial order of events occurring at multiple timescales within a signal. Thus, a fundamental challenge to linguistic processing is monitoring not only which elements occur in a sequence, but also where they occur. A system capable of linguistic processing must have at its disposal sufficient memory to store multiple items after a signal fades and to represent the serial arrangement of those items. Understanding how temporal pattern information is encoded, otherwise known as the 'problem of serial order' by Lashley (1951), has been of longstanding importance to psychology since Ebbinghaus' early models of the serial position effect (i.e., primacy and recency; Ebbinghaus, 1913). Here, we briefly discuss the two most prominent behavioral accounts of sequence-encoding: chaining and positional models.

Chaining models emerged from the classic stimulus-response theories of serial behavior championed by Watson (1920), Washburn (1916) and Skinner (1934). These models propose that a given element's location in a sequence is encoded by association with both the preceding and succeeding element. Accordingly, the sequence ABCD would be encoded (most simply) as a sequence of pairwise associations, such as A-B, B-C, C-D, where the recall of a single item initiates the recall of a subsequent item. Positional models, on the other, suggest that items are encoded on the basis of their position in the sequence. The limitations in understanding serial behavior and learning using only strict associative chaining theories, even in regards to language, have long been known (e.g. Lashley, 1951) and a considerable amount of research has been undertaken to demonstrate that sequence learning in humans and nonhumans additionally relies on positional information. Further, while both models posit potential psychological accounts to encode serial order, positional models do a better job accounting for common errors in human serial recall. For example, intrusion errors occur when trying to recall one of two lists, such as ABCD and EFGH. Such errors occur when an item from the second list is mistakenly presented during recall of the first list, but in its canonical position assignment (i.e., AFCD). Chaining models would predict that an error in the second position would trigger a cascade of incorrect responses, such as AFGH, a type of mistake not commonly seen during erroneous sequence recall (Henson 1998). Given that human errors during serial recall suggest an encoding method that can incorporate positional cues, one might ask how these positional representations are instantiated. The most compelling behavioral account suggests that positions are assigned relative to certain perceptual anchors. Specifically, Henson's start-end model (1998) states that positional assignments are encoded relative to sequence edges (Henson 2001; Henson 1998; Henson et al. 2000; Henson et al. 1996).

Further evidence supporting an edge-based serial processing system in human adults comes from the study of artificial grammar learning. For instance, Endress and colleagues (2005) demonstrated that repetition-based rule structures are only generalized if those repetitions occur on the edges of a sequence. For example, syllable sequences of the form ABCDDEF were not generalized by subjects. If, however, the repetitions were edge-based, such as ABCDEFF, subjects did effectively generalize knowledge of the repetition rule. Importantly, subjects could still discriminate grammatical and ungrammatical sequences with either internal or edge-based repetitions, suggesting differences in generalization ability were not based on a failure to detect internal repetitions. Thus, the authors conclude that subjects relied on both detecting a repetition and processing where the repetition occurred, thereby paralleling the claim of Henson that items in initial and final positions are more robustly encoded than those at inner positions of a sequence.

The use of positional information to recall serial order is not purely an artifact of sequence memory or artificial grammar learning tasks (Endress et al. 2005). Rather, natural languages possess many grammatical and phonological properties that rely on the positional assignment of elements within a sequence. For example, affixation rules most often alter the beginning or end positions of a sequence, rather than middle positions (Endress et al. 2009). Additionally, prosodic components of language similarly rely on positional information, as in allocating syllabic stress (e.g., in Hungarian, the first syllable of a word is stressed; in Macedonian, the antepenultimate; in Polish, the penultimate; in French the ultimate). Thus, the convergence of positional encoding strategies for serial recall, as well as their prevalence in natural languages, suggests that such serial expertise relying on positional, as well as chaining, strategies is a necessary cognitive ability used for linguistic communication.

Serial expertise in non-human animals

To have an appropriate animal model for grammatical processing, a species must demonstrate similar working memory constraints as humans for sequence recall as well as similar encoding strategies. While several studies of serial behaviors in pigeons and mammals are easily explained exclusively by chaining strategies (Balleine et al. 1995; Weisman et al. 1980), others are not (D'Amato and Colombo 1988; Endress et al. 2009; Straub and Terrace 1982; Terrace 2005; Terrace 1987). Some of these latter experiments demonstrate reliance on positional information during serial recall via the "simultaneous chaining" procedure (Terrace, 2005). In this task, animals learn to touch images in a specific order, receiving reinforcement only if the entire sequence is recalled correctly. Crucially, the images are presented simultaneously but in a random location for each trial. Consequently, unlike successive chain tasks, where serial learning was assessed with a specified spatial configuration (i.e., in a maze), the simultaneous chain paradigm forces subjects to acquire a representation of stimulus serial order independent of learning a fixed set of motor responses. This procedure has provided strong support for the use of positional information to learn and recall serially ordered visual stimuli (Chen et al. 1997; D'Amato and Colombo 1988; Terrace 1987). In addition, a more recent experiment undertaken by Endress et al (Endress et al. 2010) found that both chimpanzees and humans similarly utilize edge-based positional information in an auditory artificial grammar habituation/dishabituation task.

Amongst vocal learning species, songbirds have been the most extensively studied regarding their serial processing capacities. Through a direct comparison with human subjects, one species of songbird, the European starling (*Sturnus vulgaris*), has demonstrably similar auditory memory capacity limits and decay functions for short-term store (Zokoll et al. 2008). Recently, an experiment by Comins and Gentner (2010) explored the sequence-encoding strategies of this same species of songbird using an operant conditioning procedure. Here, subjects encountered a string of eight species-specific song motifs arranged in either a structured or random order. When positional information was available, subjects successfully learned to classify structured from random strings of these motifs. When absolute position information was removed, subjects' classification performance could persist with only relative position cues, albeit much more modestly. Thus, these results demonstrate starlings can rely, at least partially, on absolute and relative position information and suggest that recognition of structured vocal signals in this species could additionally utilize positional information for representing serial order (Comins and Gentner 2010).

This body of work indicates that animals, like humans, can use a host of positional cues in the absence of associative chains. Unlike other non-human species, only the songbirds undergo a vocal learning procedure with many notable similarities to human infants acquiring knowledge of language (Brainard and Doupe 2002). Thus, the songbird may extend previous animal models of serial order processing and recall into a natural behavioral context tied to vocal communication and serve as a suitable window for similar processes in linguistically-endowed humans.

Neural mechanisms of temporal order in songbirds

In this section we revisit Lashley's 'problem of the serial order' from a neurobiological perspective. At its core, this task requires the nervous system to enhance or suppress responses to a stimulus based on its temporal context: did stimulus event B correctly follow event A? In many communication systems, this "temporal context" can be defined on multiple levels of a hierarchy. Consider the case of speech and language, where the ordering of phonemes is crucial to the emergence of a word, and still the ordering of words can be vital to the emergence of an expression. Thus, the neural systems responsible for this behavior must integrate contextual information across a large range of timescales of stimulus events, from milliseconds to perhaps several seconds. Here, we review findings on how serial order for hierarchically organized elements of natural communication is, at least partially, represented in the zebra finch (*Taeniopygia guttata*).

Zebra finch males sing stereotyped songs to court females, who do not sing. These songs are characterized by their short length and hierarchical organization. On the most fundamental level, the song is a series of short sound units called notes. These notes are combined in packages referred to as song syllables. The mature zebra finch song further has a canonical syllable progression within a song bout. Thus, zebra finch song proceeds along a sequence of changing syllables, where each syllable represents a complex auditory event.

Sensitivity to such serial ordering of song elements has been investigated physiologically across different levels of the avian telencephalon, namely field L and HVC (Lewicki and Arthur 1996; Lewicki and Konishi 1995; Margoliash 1983;

Margoliash and Fortune 1992). Field L, analogous to the mammalian primary auditory cortex, receives thalamic afferents from the nucleus ovoidalis (Kelley and Nottebohm 1979), while HVC is an upstream projection target of the higher subregions of field L, L1 and L3, as well as the nonprimary auditory area CLM which projects to the HVC shelf. In these studies, an anesthetized male subject is exposed to several variants of his own pre-recorded song, a stimulus known to selectively drive neuronal responses particularly in HVC (Margoliash 1986). By presenting the subject with renditions of his own song occurring in normal, reversed, syllable reversed and sub-syllable reversed orders, researchers have successfully identified classes of neurons sensitive to the progression of syllabic or sub-syllabic features or both (Margoliash & Fortune, 1992; Lewicki & Arthur, 1996). In addition, Lewicki & Arther (1996) showed strong convergence between the anatomical projections from field L to HVC and the sensitivity of neurons therein to higher-order temporal contexts. Specifically, recordings from the primary thalamorecipient zones of field L, L2a and L2b, were only sensitive to temporal differences between normal and reversed song. However, HVC projecting regions L1 and L3 showed a modest percentage of temporal context sensitive cells for sub-syllable manipulations and L3 alone responded to differences in syllable order, while HVC shows an even higher proportion of neurons tuned to these stimulus properties.

Though these studies clearly show temporal sensitivity in a percentage of field L and HVC neurons, they do not directly address how these cells gate their responses to a given stimulus event as a function of its temporal context. Extracellular responses do, however, rule out simple facilitation as a potential mechanism of syllable-order sensitivity (Lewicki & Konishi, 1995; Lewicki & Arthur, 1996). Further, given that strong extracellular responses to specific syllables in forward song were nearly entirely eradicated in distorted temporal contexts, Lewicki & Konishi (1995) employed intracellular methods in HVC neurons. Their results suggest that temporal context sensitivity is modulated by inhibition and bursting following the appropriate sequential order. Computational models of such responses have thus predicted that zebra finch song sequencing information is organized in a chain-like manner, where nodes on the chain are responsible for variable context-sensitivities (Drew and Abbott 2003; but see Nishikawa et al. 2008 for an alternative model in Bengalese finches).

The responses of temporal context sensitive cells in the zebra finch auditory system are highly tuned to the local and global structure of serial order of the bird's own song. Many properties of serial order representation in songbirds, however, remain to be tested. In the auditory system, the hierarchy of context sensitivity has only been studied at the level of field L and HVC, while the contribution of other auditory areas, such as CM and NCM, which show behaviorally relevant modification of song selectively (Gentner and Margoliash 2003; Thompson and Gentner 2010), remains unknown. The role of NCM is particularly important in understanding temporal context sensitivity, as this is a well-established area of experience-dependent decision-making for mate decisions based on male song features in European starlings (Gentner et al. 2001; Gentner and Hulse 2000; Sockman et al. 2005; Sockman et al. 2002), whose songs are largely characterized by their motif-structure (Eens et al. 1988).

Additionally, neurophysiological explorations have yet to dissociate chaining from positional representations of sequentially arranged stimuli. For example, consider a cell that shows sensitivity to the sequence AB. At present, it is not entirely possible to isolate whether the subject is responding to B given the information provided by the association of A to B, or by B's position in the sequence relative to A (i.e., the second motif). One possible way to parse apart these types of temporal information would be to create stimuli for a subject that combine motifs across different variations of autogenous songs. Thus, if a bird sings two songs, one beginning with motif sequence AB and another beginning CD, a relative position-encoding model might be robust to order violations in the sequence CB compared to BB or BC. The reason being that motif B, though presented in a non-canonical transition from C to B is still located in the correct relative position of the sequence. Such a design would dissociate between the encoding of positional versus transitional sequence information as outlined above.

Finally, the role of non-auditory areas, such as the basal ganglia, might provide important contributions to serial order representation that remain to be explored. While the songbird anterior forebrain has been intensively studied as premotor circuit contributing to song sequence production (for a review, see Brainard & Doupe, 2002), it has only recently been suggested as an important region of syllable-level syntax discrimination. Abe & Watanabe (2011) tested syntax discrimination abilities of Bengalese finches. Using an immunocytochemical technique, the authors stained Zenk protein, an immediate early gene upregulated during exposure to conspecific song (Mello et al. 1992; Mello and Ribeiro 1998), to localize areas responding strongest to violations of a familiarized temporal syllable order. With this method, it was shown that neurons in the lateral magnocellular nucleus of the anterior nidopallium (LMAN) showed heightened activation to temporal orders (Abe and Watanabe 2011). LMAN, along with two other regions in the anterior forebrain pathway, Area X (which receives projection from HVC) and the dorsal lateral nucleus of the medial thalamus (DLM), comprise an analogue to the human basal ganglia (Mello et al, 1992). Thus, these results suggest an additional important contribution of non-auditory structures in serial order learning of natural communication sounds.

Closing remarks

The above case represents but one of the panoply of behaviors encapsulated by language. While the whole of human language is unique, here we suggest that many of the requisite cognitive capacities that underlie it, such as object recognition, categorization, and pattern generalization, are shared by other species and can therefore offer important insights into the biological basis of such language-relevant behaviors. In this dissertation, we offer evidence that the European starling can serve as a model system to examine the biological basis of auditory pattern generalization, whereby knowledge of a sound-based pattern is dissociable from the sounds comprising the pattern itself, a crucial ability underlying language acquisition (Marcus et al, 1999).

We begin by establishing the ecological validity of studying pattern recognition in European starlings. To do so, we show that the naturally occurring sequential patterning of vocal components (motifs) in each male's songs can carry information about his likely relative fitness (chapter two). We then show that female starlings are sensitive to the sequential patterning of motifs in male song, evinced by strong preferences for some pattern types over others. Together, these results implicate a role for auditory pattern learning and generalization in starling mate choice.

After establishing the ecological relevance of motif patterning in starlings, we assess the cognitive mechanisms underlying their ability to learn and generalize auditory patterns. In chapter three, we demonstrate conclusively that abstract motif patterning rules can be acquired by starlings from the use trial-unique pattern stimuli. Next, in chapter four, using a bottom-up perspective, we consider how the perceptual organization of a pattern's underlying constituent elements constrains pattern learning and generalization. Then, in chapter five, we take the reverse approach by asking how the structure of a pattern itself effects the starling's understanding of the perceptual organization of constituents. The results from these chapters clearly highlight the rich interaction between the structure of a pattern and the constituent elements in real-world communication signals and further showcase how knowledge of serial expertise, categorization and pattern learning work together to achieve successful generalization.

Using our understanding of these constraints, we demonstrate in chapter six that the starlings' knowledge of the rules describing patterned sequences, and the constituent elements that instantiate those patterns, are dissociable. Finally, in chapter seven, we conclude by reviewing these advances in the behavioral expression of pattern generalization and connect these findings to recent work on the neural basis of temporal processing of natural auditory communication signals. We put forth the starling as a model organism for understanding how language-relevant pattern recognition shapes and modulates sensory systems, which informs our understanding of an initial neurobiology of language.

Acknowledgements

This chapter is, in part, is material as it appears in Frontiers in Evolutionary Neuroscience, 2012. Kiggins, Justin K.; Comins, Jordan A.; Gentner, Timothy Q., Frontiers 2012. The material included only represents those portions of the article of which the dissertation author was the primary author.

References

Abe K, Watanabe D (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. Nature Neuroscience 14:1067–1074. doi: 10.1038/nn.2869

Balleine BW, Garner C, Gonzalez F, Dickinson A (1995) Motivational control of heterogeneous instrumental chains. Journal of Experimental Psychology: Animal Behavior Processes 21:203–217. doi: 10.1037//0097-7403.21.3.203

Berwick R, Okanoya K, Beckers G, Bolhuis J (2011) Songs to syntax: the linguistics of birdsong. Trends in Cognitive Sciences 15:113–121.

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

Chen S, Swartz KB, Terrace HS (1997) Knowledge of the orginal position of list items in rhesus monkeys. Psychological Science 8:80–86.

Comins JA, Gentner TQ (2010) Working memory for patterned sequences of auditory objects in a songbird. Cognition 117:38–53.

D'Amato MR, Colombo M (1988) Representation of serial order in monkeys (Cebus apella). Journal of experimental psychology Animal behavior processes 14:131–9.

Drew PJ, Abbott LF (2003) Model of song selectivity and sequence generation in area HVc of the songbird. Journal of neurophysiology 89:2697–706. doi: 10.1152/jn.00801.2002

Ebbinghaus H (1913) Memory: A Contribution to Experimental Psychology. Teachers College, Columbia University., New York

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86.

Endress AD, Carden S, Versace E, Hauser MD (2010) The apes' edge: positional learning in chimpanzees and humans. Animal Cognition 13:483–495.

Endress AD, Nespor M, Mehler J (2009) Perceptual and memory constraints on language acquisition. Trends In Cognitive Sciences 13:348–353.

Endress AD, Scholl BJ, Mehler J (2005) The role of salience in the extraction of algebraic rules. Journal of experimental psychology General 134:406–19. doi: 10.1037/0096-3445.134.3.406

Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. Science (New York, NY) 303:377–80. doi: 10.1126/science.1089401

Fitch WT, Hauser MD, Chomsky N (2005) The evolution of the language faculty: clarifications and implications. Cognition 97:179–210; discussion 211–25. doi: 10.1016/j.cognition.2005.02.005

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Animal Behaviour 59:443–458.

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. Journal of neurobiology 46:48–58.

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674.

Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–79. doi: 10.1126/science.298.5598.1569

Henson RN (1998) Short-term memory for serial order: the Start-End Model. Cognitive psychology 36:73–137.

Henson RN (2001) Serial order in short-term memory. The Psychologist 14:70–73. doi: 10.1037/a0018540

Henson RNA, Burgess N, Frith CD (2000) Recoding , storage , rehearsal and grouping in verbal short-term memory : an fMRI study. Neuropsychologia 38:426–440.

Henson RNA, Norris DG, Page MPA, Baddeley AD (1996) Unchained Memory: Error Patterns Rule out Chaining Models of Immediate Serial Recall. The Quarterly Journal of Experimental Psychology A 49:80–115. doi: 10.1080/027249896392810

Jackendoff R, Pinker S (2005) The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). Cognition 97:211–225. doi: 10.1016/j.cognition.2005.04.006

Kelley DB, Nottebohm F (1979) Projections of a telencephalic auditory nucleus-field L-in the canary. The Journal of comparative neurology 183:455–69. doi: 10.1002/cne.901830302

Lashley K (1951) The problem of serial order in behavior. In: Jeffress L (ed) Cerebral Mechanisms in Behavior. Wiley, pp 112–137

Lewicki MS, Arthur BJ (1996) Hierarchical organization of auditory temporal context sensitivity. The Journal of neuroscience : the official journal of the Society for Neuroscience 16:6987–98.

Lewicki MS, Konishi M (1995) Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. Proceedings of the National Academy of Sciences of the United States of America 92:5582–6.

Margoliash D (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. The Journal of neuroscience : the official journal of the Society for Neuroscience 3:1039–57.

Margoliash D (1986) Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. Journal of Neuroscience 6:1643–1661.

Margoliash D, Fortune S (1992) Temporal and Harmonic Combination-Sensitive Finch ' s HVc Neurons in the Zebra. The Journal of Neuroscience 12:

Margoliash D, Nusbaum HC (2009) Language: the perspective from organismal biology. Trends In Cognitive Sciences 13:505–510.

Mello C V, Ribeiro S (1998) ZENK protein regulation by song in the brain of songbirds. The Journal of comparative neurology 393:426–38.

Mello C V, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proceedings of the National Academy of Sciences of the United States of America 89:6818–22.

Nishikawa J, Okada M, Okanoya K (2008) Population coding of song element sequence in the Bengalese finch HVC. The European journal of neuroscience 27:3273–83. doi: 10.1111/j.1460-9568.2008.06291.x

Pinker S, Jackendoff R (2005) The faculty of language: what's special about it? Cognition 95:201–36. doi: 10.1016/j.cognition.2004.08.004

Skinner B (1934) The extinction of chained reflexes. Proceedings of the National Academy of Sciences 20:234–237.

Sockman KW, Gentner TQ, Ball GF (2005) Complementary neural systems for the experience-dependent integration of mate-choice cues in European starlings. Journal of neurobiology 62:72–81.

Sockman KW, Gentner TQ, Ball GF (2002) Recent experience modulates forebrain gene-expression in response to mate-choice cues in European starlings. Proceedings Biological sciences / The Royal Society 269:2479–85. doi: 10.1098/rspb.2002.2180

Straub RO, Terrace HS (1982) Generalization of serial learning in the pigeon. Animal learning and Behavior 9:454–468.

Terrace HS (2011) Characterizing Consciousness: From Cognition to the Clinic? In: Dehaene S, Christen Y (eds) Research and Perspectives in Neurosciences - Characterizing Consciousness: From Cognition to the Clinic? Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–25

Terrace HS (2005) The simultaneous chain: a new approach to serial learning. Trends In Cognitive Sciences 9:202–210.

Terrace HS (1987) Chunking by a pigeon in a serial learning task. Nature 325:149–151.

Thompson J V, Gentner TQ (2010) Song Recognition Learning and Stimulus-Specific Weakening of Neural Responses in the Avian Auditory Forebrain. Journal Of Neurophysiology 103:1785–1797.

Washburn MF (1916) Movement and Mental Imagery: Outlines of a Motor Theory of the Complexer Mental Processes. The Philosophical Review. doi: 10.2307/2178292

Watson J (1920) Is thinking merely the action of language mechanisms? British Journal of Psychology 11:87–104.

Weisman RG, Wasserman E a., Dodd PW, Larew MB (1980) Representation and retention of two-event sequences in pigeons. Journal of Experimental Psychology: Animal Behavior Processes 6:312–325. doi: 10.1037/0097-7403.6.4.312

Zokoll MA, Naue N, Herrmann CS, Langemann U (2008) Auditory memory: A comparison between humans and starlings. Brain Research 1220:33–46.

CHAPTER TWO

Male song syntax controls preference in female songbirds

Sexual selection pressures, i.e. differences in reproductive success caused by competition over mates, help determine which characteristics persist in the population (Darwin 1871). Across species, selection occurs when females choose superior mates, assessing male quality via morphological and behavioral characteristics. Many male oscine birds use acoustically conspicuous songs to court females, and the elaborate structure of these vocalizations is understood to be the product of sexual selection (Searcy and Nowicki 2005). In European starlings, male song consists of a series of short, stereotyped sound patterns (called motifs) arranged into long sequences (bouts). Male starlings continue to add motifs to their vocal repertoires as they age (Mountjoy and Lemon 1995), which in turn increases average song bout length (Eens et al. 1991). Because older males enjoy greater reproductive success than younger males, bout length and repertoire size are thought to provide honest cues for female choice (Eens et al. 1991; Mountjoy 1996; Gentner and Hulse 2000). Consistent with this, female starlings prefer to copulate with males who sing longer song bouts (Komdeur et al. 2005), and relative bout length can influence female behavior in the laboratory (Gentner and Hulse 2000) and neuronal responses (Gentner et al. 2001). Surprisingly, female preferences for longer songs do not require exposure to complete song bouts or full motif repertoires (Eens et al. 1988; Gentner and Hulse 1998), and the explicit features of song that directly control female mating preferences are not known. Here, we investigate the relationship between male song syntax and female preferences.

We first asked whether the sequential organization of motifs in natural song (i.e. song syntax) could carry information about the potential fitness of different males. We analyzed sequential organization in the song bouts of several adult male starlings (N = 5;

for an example song, see Figure 1A). Using an information theoretic approach (Shannon and Weaver 1949), we quantified variability in the statistical reliability of transitions between motifs in the songs of the different males (Figure 1B; Supplementary Materials). As reported previously in starlings *(10)* and other songbirds (Briefer et al. 2010), the songs of all singers showed a characteristic organization in which most of the information about the temporal sequence of motifs was carried in ordered pairs of motifs. That is, the uncertainty (Shannon entropy) of observing a given motif at time t was, in all males, greatly reduced by taking into account the immediately preceding motif in the bout (i.e. the motif at time $t-1$), and reduced still more by also considering the next preceding motif (at time $t-2$; Figure 1B). We then quantified the sequential information contained in each male's song repertoire as the total drop in uncertainty (mutual information) provided by the temporal ordering of motifs, which we term syntactic reliability (Supplementary Materials). Surprisingly, the syntactic reliability of each male's song was a very strong predictor of their song bout length (SBL; Pearson's correlation, $r(3) = 0.9183$, $p = 0.0277$; Figure 1C). In other words, males that sing longer song bouts, tend to sequence the motifs in their bouts in more predictable ways than males that sing shorter bouts. Female starlings' sensitivity to this syntactic reliability could explain mate choice preferences based on bout length (Searcy and Nowicki 2005), and would provide a means of obtaining honest fitness information from the songs of different males.

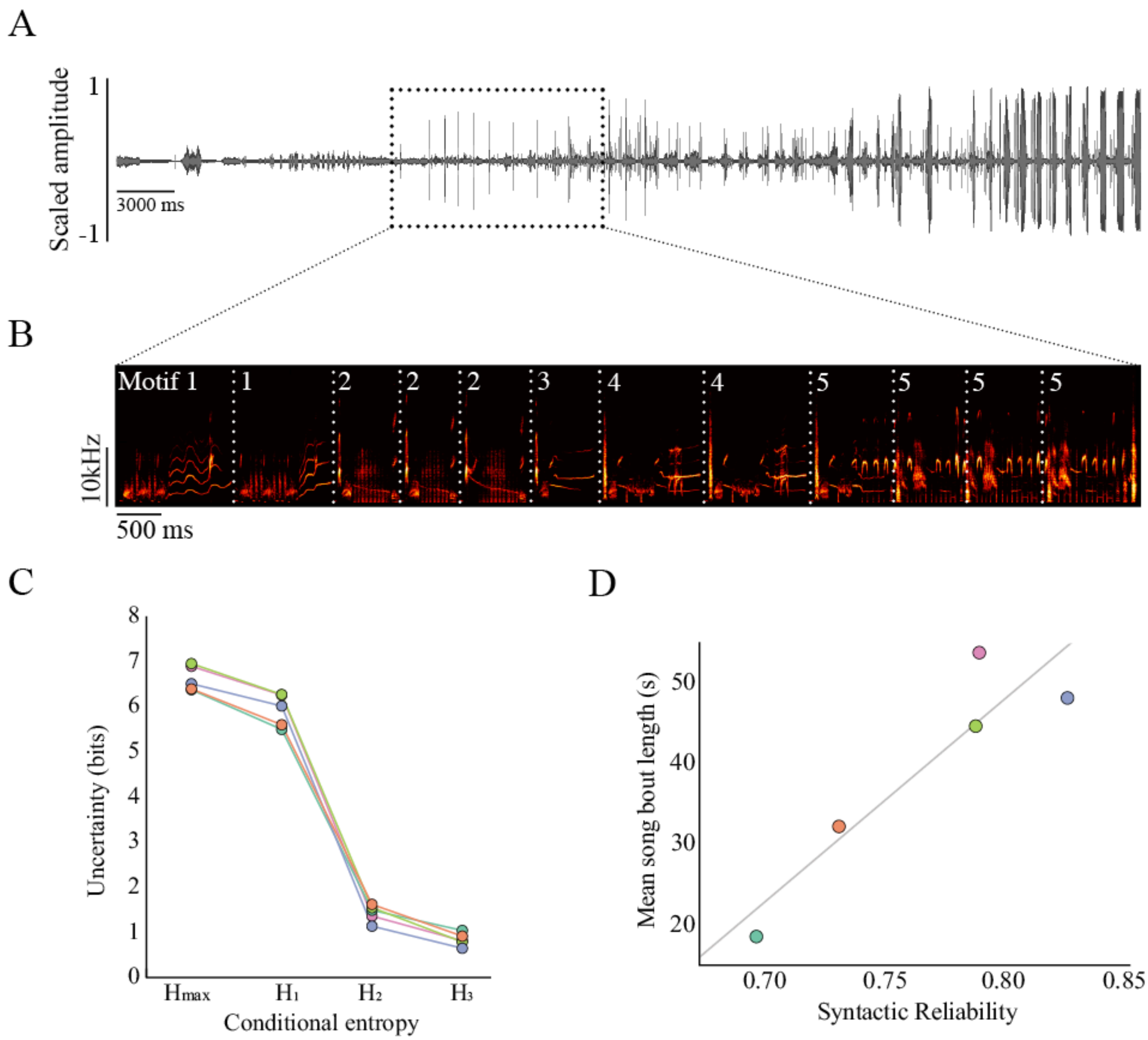

Figure 2.1. Male song syntax predicts bout length. (A) Sound pressure waveform of one song bout from a male starling and (B) spectrogram from a portion of the bout showing a sub-sequence of motifs. Motif boundaries are marked by dotted lines. Different numbers overlaid on the spectrogram denote unique motifs. Motif sequences within songs are not random. (C) For all birds recorded (each represented with a unique color), the total possible uncertainty about the identity of a given motif (H_{max}; methods) at any point in the song bout, decreases systematically when one accounts for the overall frequency at which each motif occurs (H_1), the frequency of order pairs of motifs (H_2), and the frequency of ordered triplets (H_3). The large drop in uncertainty between H_1 and H_2 reveals that most of the information about motif occurrence derives from incorporating knowledge of the immediately preceding motif. (D) The proportion of total motif entropy explained by motif sequences, syntactic reliability (methods), for each singer is a significant predictor of the mean length of the bouts that bird sings.

If the syntactic structure of natural songs provides a salient cue, then female starlings should prefer motif sequences with a predictable syntax over those without such organization. To test this hypothesis, we measured the preferences of female (N=6) and male (N=8) starlings for artificial motif sequences with identical low-level acoustics but different sequential statistics. We extracted eight unique warble motifs from the song repertoire of a single male starling unfamiliar to all subjects (Supplementary Materials), and arranged these motifs into different sequences by applying one of two generative rules (Figure 2A). We generated sequences with reliable syntax by setting the transition probability between eight pairs of motifs to be 0.93, such that each motif accurately predicted the next motif in the sequence 93% of the time (Figure 2B). We generated sequences without reliable syntax in a similar way, except that all transitions between motifs were equally likely, and thus the occurrence of a specific motif at any specific time could not be predicted from the presence (or absence) of any other motif (Figure 2C; Supplementary Materials). We then measured behavioral preferences for these two classes of sequences using an operant apparatus containing three nest boxes, each outfitted with an internal speaker and an infra-red wired perch (Figure 2D; Supplementary Materials). When the subject alighted on the perch in front of a given nest box, motif sequences following one of the two syntactic rules played from the speaker inside that nest box. When the bird left the perch, the motif sequence terminated. Subjects lived in the apparatus for four days. On each day, one perch always triggered playback of syntactically reliable sequences, one perch triggered unreliable sequences (according to the two generative rules), and the third perch generated no song. We refer to these locations as the syntactic, random, and silent perches, respectively. Across successive days, the syntactic and random perches switched, and the silent perch stayed fixed. Across subjects, we counter-balanced the initial assignment of the syntactic sequences to perches (Supplementary Materials).

We quantified behavioral preference by comparing the proportion of time each subject spent on the syntactic and random perches each day (Supplementary Materials). Consistent with the structure of natural song, female starlings showed a significant preference for the syntactically reliable motif sequences on the first day (rmANOVA on ranks, stimulus class*day interaction, $F_{3,35} = 5.09$, $p = 0.0049$; planned contrast between stimulus classes on day one, $p = 0.0134$; Figure 2E). During this first day, the females made more visits on average to the syntactic than the random perch (mean over all females: 496 vs. 139 visits), and spent more time overall on the syntactic perch than on the random perch (mean over all females: 3067 vs. 675 seconds). The preference for syntactically reliable songs on day one was observed in 5 of the 6 females tested (Supplemental Figure S1), and analyses of alternative measures of behavior in the apparatus yield similar results (Supplemental Figure S2).

The song-driven preference present on day one transformed into a stable position preference on subsequent days. On days 2-4, females continued to spend more time on the perch that had been associated with the syntactically reliable songs on day one (rmANOVA on ranks $F_{1,25} = 11.609$, $p = 0.0022$; Figure 2E). This position preference coincided with a general decrease in the attractiveness of song, with the average duration of a visit to either perch declining significantly across days 2-4 compared to day one (rmANOVA on ranks, stimulus class*day interaction, $F_{3,35} = 8.26$, $p = 0.0003$; Tukey HSD post-hoc tests). Across the whole 4-day session, however, females spent more time on the perches that played song than on the silent perch (rmANOVA on ranks, $F_{(1,35)} = 19.35$; $p<0.0001$) and this difference was consistent across days (rmANOVA on ranks, song-silence*day interaction, $F_{(3,35)} = 0.356$; $p = 0.785$). We conclude that the syntactic organization of conspecific male song is perceived by females, and can drive differential female behavior at putative nest sites.

In contrast to females, male starlings did not show a preference for either motif sequence (rmANOVA on ranks, stimulus class*day interaction, $F_{3, 49} = 1.18$, $p = 0.33$; Supplemental Figure S3). Indeed, the pattern of male behavior in the apparatus differed significantly from that of the females (rmANOVA on ranks, stimulus class*day*sex interaction, $F_{3, 84} = 3.4192$, $p = 0.0210$).

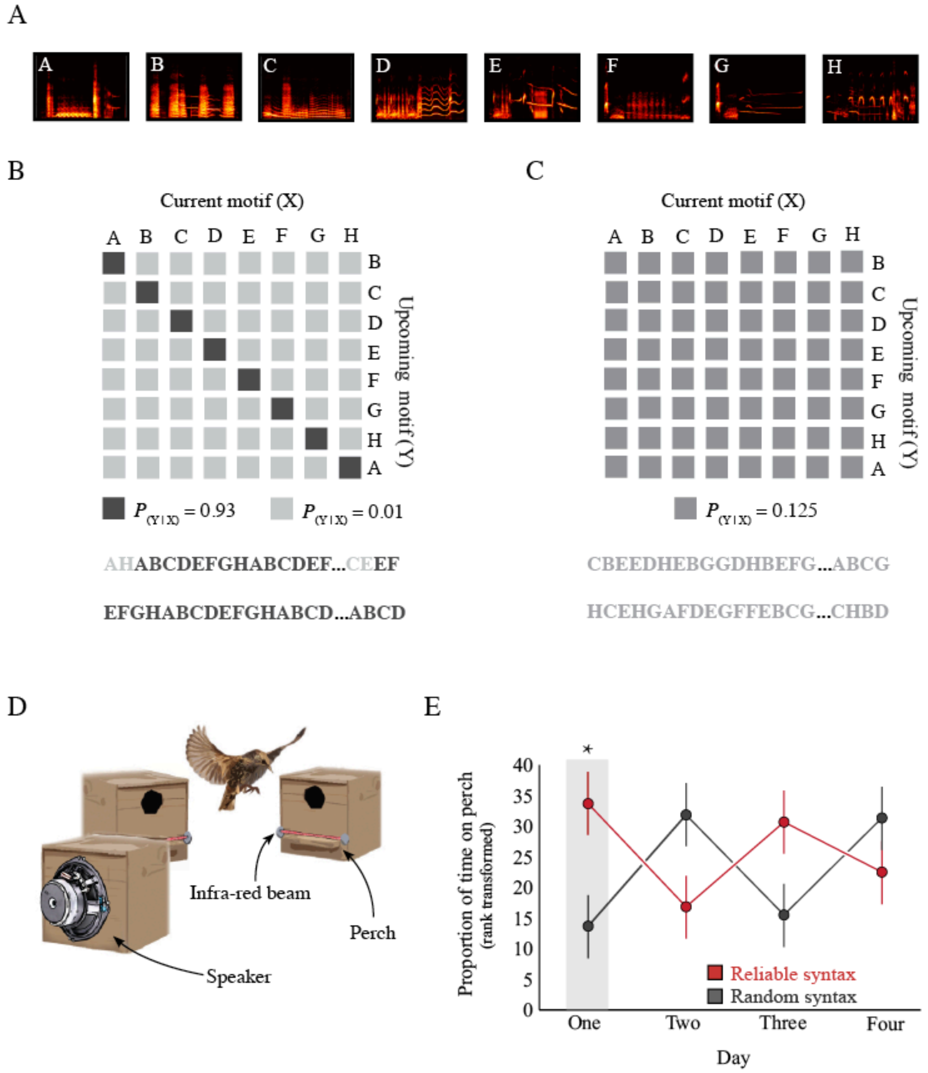

Figure 2.2. Sequential syntax driven behavioral preferences in female starlings. (A) Spectrograms of eight motifs used to examine spontaneous preference for motif sequences with reliable syntax versus sequences with random syntax. (B) Sequences with reliable syntax were built following a transitional matrix that heavily biased 8 out of all possible motif transitions (methods). Demonstrative stimulus sequences with reliable syntax are seen below. (C) To construct sequences with random syntax, all possible transitions between our 8 motifs were given equal probability of occurrence (again, see below for demonstrative sequences). (D) Schematic of song preference apparatus. Subject generated the playback of a randomly generated synthetic motif sequence, by landing on an IR-wired perch in front of a nest box. The sequence continued to play up to 40 motifs (8 unique) as long as the bird remained on that perch. One perch played songs with reliable syntax another play songs with random syntax and one played nothing. On alternating days, the song sequences assigned to perches switched. (E) Females showed a significant preference to spend time on the perch in front of the nest box associated with a syntactically structured song during the first day (graph displays mean ranked proportion ± pooled standard error). They maintained their preference over the next 4 days of data collection, despite the stimulus associated with that perch changing each day.

The observed preferences for syntactically reliable songs may reflect pre-specified, innate biases (Chiandetti and Vallortigara 2011), or may emerge as a consequence of prior song experience. Because the subjects we tested were all wild-caught, our knowledge of their experience is limited, and if experience does play a role, it would require preference-relevant memories to endure (at minimum) over the several months that the animals were held in captivity prior to testing. To address the role of experience in song-driven preferences, we trained another group of starlings (N = 4) using a go-nogo operant procedure to classify two sets of conspecific song motifs according to simple syntactic cues, i.e. whether specific motifs occurred in a pair or alone (Figure 3A, B; supplemental materials). Subjects rapidly learned to peck at a defined location following the presentation of some motifs and motif pairs (S+), and to withhold pecks following other motifs and motif pairs (S-). If subjects responded correctly to the S+ stimuli they received a brief opportunity to feed. If they responded (incorrectly) to the S- stimuli they were punished with a short time-out during which the apparatus went dark and food remained inaccessible. Correctly withholding pecks to the S- stimuli never led directly to food. We trained the birds for an average of 98 ± 21.8 100-trial blocks, at which point the S+ and S- stimuli evoked significantly different behavioral responses, compared to initial performance (rmANOVA $F_{(1,41)}= 11.6$, $p = 0.0015$; Figure 3C). We then moved the subjects from the operant apparatus to a large flight-cage where they had free access to food and water. After a minimum of 211 days (range: 211-299), we tested the subjects' preference for the trained S+ and S- motif stimuli (Figure 3A) in the nest box apparatus used for the song syntax experiments. Our motivation was to assess whether the learned appetitive and aversive valences for the S+ and S- stimuli, respectively, would influence subsequent nest box preferences. Consistent with this idea, the operantly trained starlings preferred to spend more time perched at the nest box that triggered S+ stimuli, than S- stimuli, on day one (rmANOVA on ranks, stimulus class*day interaction, $F_{3, 21} = 8.01$, $p = 0.001$; planned contrast between stimulus classes on day one, $p = 0.011$; Figure 3D, and Supplementary Figure S4A). As with the syntactically reliable motifs sequences, the S+/S- nest box preferences observed on the first day of testing also gave way to a position preference across subsequent days (rmANOVA on ranks, $F_{(1,15)} = 6.61$, $p = 0.0213$). These results are important for two reasons. First, the direction of the response differences toward songs with a strong positive valence (i.e. those associated with food), validates the underlying assumption that our apparatus is measuring differences in appetitive value of stimuli associated with the nest boxes. Indeed the strength of the initial learning, as measured in a generalization test (Supplementary Materials), correlated significantly with the number of visits to the S+ perch on day one (Supplementary Figure S4C). Second, while these results do not rule out the possibility that preferences for syntactically reliable songs have a significant genetic component (Deacon 2010; Kagawa et al. 2012), they show that prior experience can clearly modulate song preferences across extended time periods and diverse behavioral contexts. Thus, an innate preference is not essential.

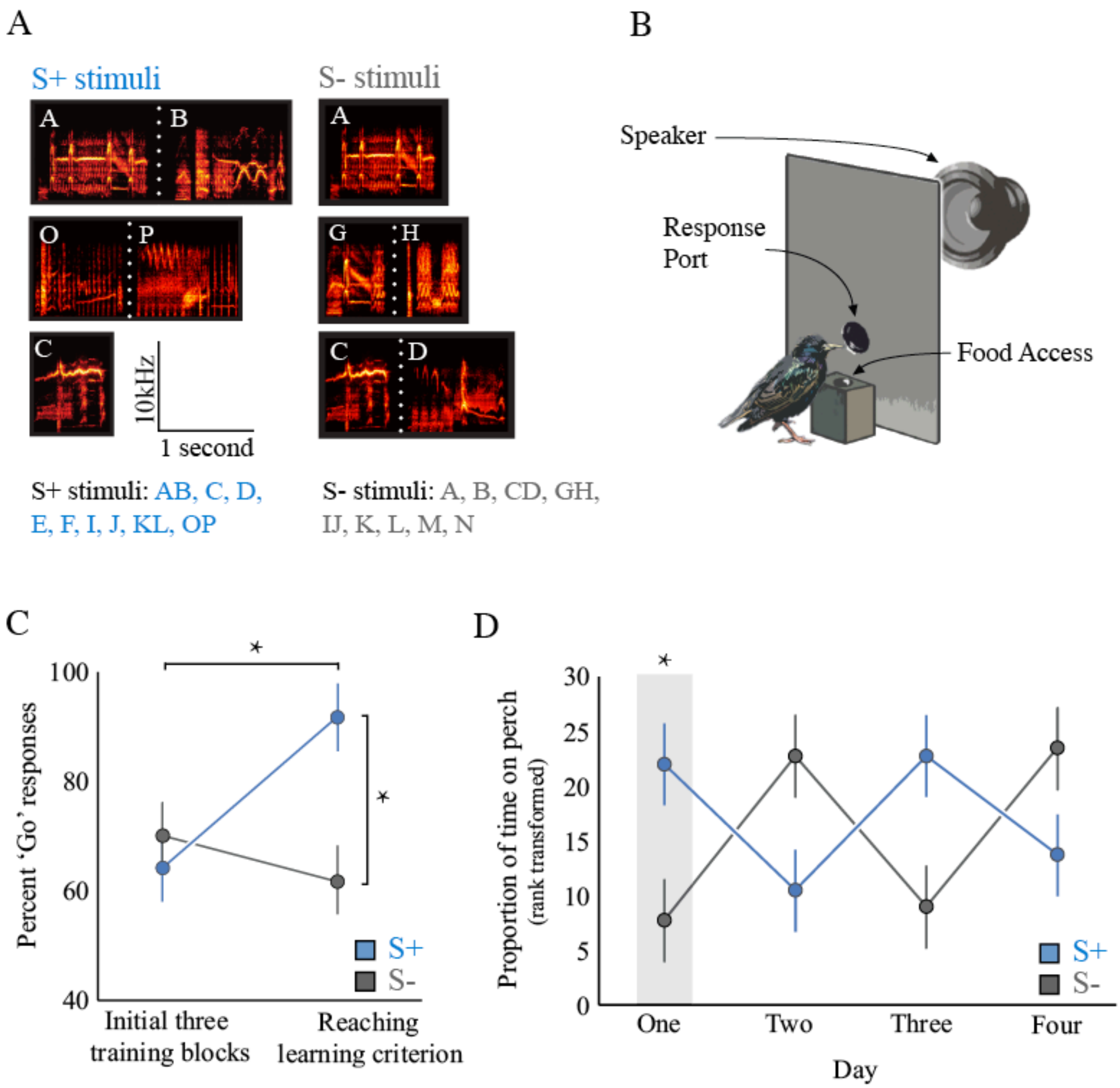

Figure 2.3. Preference behaviors elicited from associatively learned behavioral goals. (A) Spectrograms of motif stimuli used as S+ and S- stimuli in a go-nogo paradigm. (B) Schematic of the operant apparatus. Subjects start a trial by pecking the center response port. After the stimulus ends, the subject either pecks the center response port again or withholds any response depending on the class from which the stimulus was drawn. Correct responses yield a food reward. Incorrect responses lead to the house light being extinguished and food being inaccessible. (C) Percentage of 'go' responses produced by subjects during the initial three 100-trial blocks of training compared with the three 100-trial blocks at our learning criterion (see supplemental methods; graph displays mean ranked proportion ± pooled standard error). Following training, subjects were placed in a flight-cage for at least 6 months with free access to food and water and no further exposure to these motif stimuli. (D) Following this period, subjects were placed for 4 days in the preference apparatus shown in fig 2, D. All subjects spent more time on the perch of the nest box associated with the former S+ stimuli on the first day of data collection and, as for the other birds, this nestbox preference persisted across the 4 days of data collection, regardless of the stimulus associated with that perch changing each day (graph displays mean ranked proportion ± pooled standard error).

Individual variation in the syntactic structure of male starling song carries behaviorally relevant information, and this syntactic information can bias the behavioral preferences of females. These behavioral preferences can be shaped more generally by experience and retained over many months. The precise mechanisms that underlie the incorporation of syntactic information into song, and its subsequent perception require further experiments. For production, it may be that singing a specific motif before or after another is somehow less costly or difficult motorically, leading ultimately to a canalization of motor output. Although the production dynamics for large repertoire song birds like starlings are poorly understood, higher level, supra-motif organization has been noted in several species (Honda and Okanoya 1999). For perception, it is clear that female starlings are sensitive to conditional probabilities between ordered pairs of motifs. Although the ultimate benefits to this capacity lie in increasing the probability that females mate with higher-quality males, the proximate mechanisms may be tied to encoding, retention, and associative biases imparted by the more regular sequences. The syntactic sensitivity described here resembles statistical learning in human infants (Saffran et al. 1996), where variability in behavior is linked to variation in the conditional probabilities of phoneme sequences (Aslin et al. 1998), and which has been proposed as a domain-general process integral to the emergence of language precursors in early hominids (Conway and Christiansen 2001; Christiansen and Chater 2008). Consistent with this idea, our results reveal a pathway for the adaptation of sequential learning and processing, and the conservation of syntactic complexity in vocal communication systems, without the necessity for combinatorial syntactic ability.

Supplementary Materials:

Syntactic reliability

We analyzed the statistical structure of motif sequences in the songs of five adult male starlings using information theoretic measures. Details of the capture and song recording procedures are published elsewhere (Gentner & Hulse 1998). For each singer, we parsed a minimum of 0.5 h of song bouts into their constituent motifs, then counted the total number of unique motifs, and the number of times that each single motif, sequential pair, and sequential triplet of motifs occurred in that bird's song repertiore. With these values for each singer, and their corresponding probabilities over the singer's repertoire, we computed the following measures of Shannon information:

(1) $$H_{\max} = \log_2 n,$$

where n is total number of unique motifs in the repertoire.

(2) $$H_1 = H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i),$$

where $p(x_i)$ is the probability that motif x_i will appear in any bout (i.e., the ratio of the frequency of a given motif to the total number of motifs in the repertoire).

(3) $$H_2 = H(Y \mid X) = -\sum_{i,j=1}^{n} p(x_i, y_j)\log_2 p(y_j \mid x_i),$$

where $p(x_i,y_j)$ is the joint probability of any sequential pair of motifs x_i and y_j, and $p(y_j|x_i)$ is the conditional probability that motif y_j immediately follows motif x_i. The equation for H_3, or $H(z|xy)$, has the same form as H_2, except that the probabilities for each sequential pair of motifs, $p(x,y)$ and $p(y|x)$, are replaced by the joint and conditional probabilities for each sequential triplet of motifs, $p(xyz)$ and $p(z|xy)$, respectively.

To measure the syntactic reliability, *SR*, of each male's songs, we first compute the reduction in entropy of each motif, *H*(X), that is obtained by taking into account the motifs preceding in time, *H*(Z|XY). This is the mutual information, given by:

(4) $$MI = H(X) - H(Z \mid XY)$$

To obtain *SR*, we then divide *MI* by the total entropy of each singer's songs, $H_{\max}$, which by eq. (1) is a function of their motif repertoire size (*n*). Thus, *SR* reflects the proportion of the total entropy in the songs of each singer that is explained by motif order. Although the measure SR isolates the effects of motif order cleanly and completely up to motif triplets, other measures of motif ordering show similar correlations with song bout length to that depicted in Figure 1.

Sequential syntax driven behavioral preferences in female starlings Methods

Subjects

Fourteen (N=14) wild-caught European starlings (*Sturnus vulgaris*) from southern California served as subjects. From the time of capture until use in this study, subjects were housed in a large mixed sex, conspecific aviary with *ad libitum* access to food and water.

Procedure

Prior to behavioral testing of sequential syntax versus no syntax preference, subjects were subcutaneously implanted with a 10mm silastic tube containing crystalline 17β-estradiol in their left flank which modulates song sensitivity in seasonal songbirds like starlings *(1)* and replicates natural circulating levels of estradiol observed during active periods of mate choice. Subjects were then held in social isolation for 48-72 hours before being moved to the preference apparatus. Sex was determined immediately following testing via laparotomy in all but one case, where gender was established based on identification of secondary sexual characteristics by two researchers. Estradiol implants were also recovered during this procedure, all of which still contained estradiol.

Apparatus

We constructed within a sound attenuation chamber an 18-inch cubical cage that housed on each of its back walls a nest box with a perch. Each nest box had a 2-inch hole drilled out of it, covered in the back with plastic mesh netting so as to not allow the bird to actually crawl through the hole. Behind each hole was a speaker to play various acoustic stimuli. Because of the small dimensions of the cage, each speaker was wrapped with acoustic foam so that its sound was channeled primarily through the nest box hole and rearward reflections were minimized. The walls of the sound attenuation chamber were covered with acoustic foam to further attenuate any reflections. On the front of each nestbox was a perch fitted with an infrared detector that, once triggered, allowed a computer to play a given stimulus. Subjects had *ad libitum* access to feed and drink from the front of the cage.

When landing on the perch, subjects must stay for a short length of time (0.5-1.5s) to initiate the auditory stimulus. This short delay was used to ensure that data collected is in fact from the bird landing instead of a “fly by” that can trip the sensor. Each perch is assigned a specific stimulus class that changes daily and is countered balanced across subjects for initial stimulus class assignments. The silent perch was always located at the center perch.

Stimuli

Sequences were built using motifs from the song recordings of one adult male starlings *(2)*. We began by extracting eight unique (non-repetitious) motifs. Using these 8 motifs, we constructed two types of motif-sequences (reliable and random). Playbacks of a randomly generated synthetic motif sequence occurred when subjects landed on an IR-wired perch in front of a nest box. The sequence continued to play for up to 40 motifs (8 unique) as long as the bird remained on that perch. Thus, in total, sequences from either reliable or random syntax conditions contained exactly 40 stimulus occurrences. For both stimulus classes the initial motif was chosen at random. However, for reliable sequences, 8 specific motif transitions (A-B, B-C, C-D, D-E, E-F, F-G, G-H and H-A) had a very high probability of occurrence (0.93) while all remaining transitions were much less likely (probability of 0.01). All motif transitions in the random sequences were equally likely to occur throughout the sequence.

Data collection and analysis

Subjects were housed in the operant preference apparatus for minimally 4 days (range 4-9). Our analysis included only those data generated from the first 4 days to ensure each subject contributed equally to our subsequent analyses. Data was converted to proportions for each of the two stimulus perches as follows:

> Proportion stimulus perch A = time on stimulus perch A perch /(time on stimulus perch A + time on stimulus perch B + time on silent perch).

Given that our data violated the assumptions of normality (Shapiro-Wilk Test for female data: W = .90, $p = 0.0008$; Shapiro-Wilk Test for male data: W = .88, $p < 0.0001$), results were rank transformed to run nonparametric tests *(17)*.

Initially, 16 adult starlings were run in this experiment, but one male and one female (N=2) were dropped from analyses either because a) they did not sample the random transition stimulus perch at all on Day 1 or b) because the subject was not experimentally naïve (if this female was included in our rmANOVA on ranks, results remain significant: $F_{(3, 42)} = 5.54$, $p = 0.0027$). Our analysis was a full-factorial general linear model, which included subjects as a random factor, stimulus class (reliable or random) and day (1, 2, 3 and 4) as fixed factors as well as the interaction between these fixed factors (stimulus class*day). The interactions between our random factor and each fixed factor were not included as these interactions did not explain any of the overall variance (subject*stimulus class; subject*day).

Preference behaviors elicited from associatively learned behavioral goals.

Subjects

Twelve (N=12) wild-caught European starlings (*Sturnus vulgaris*) from southern California served as subjects. We controlled for neither age nor sex, though all subjects had full adult plumage when acquired and, thus, were at least one year old. From the time of capture until use in this study, subjects were housed in a large mixed sex, conspecific aviary with *ad libitum* access to food and water.

Training Apparatus

The operant apparatus where starlings learned to classify the patterned stimuli is portrayed in main text (Fig 3,B). Each subject was held in a small weld-wire cage containing an operant panel. On the panel, a centrally located response port was a PVC-housed opening. Inside of this opening was an IR receiver and transmitter used to detect when the bird broke the plane of the response port with its beak. This ‘poke-hole’ design allows starlings to probe the apparatus with their beaks, in a manner akin to their natural appetitive foraging behavior. A remotely controlled hopper, positioned behind the panel, moved the food within and beyond the subject’s reach beneath the opening. Acoustic stimuli were delivered through a small full-range audio speaker mounted behind the panel and out of the subject’s view. The sound pressure level inside all chambers was calibrated to the same standard broadband noise signal. Full details for all the mechanical components of the apparatus, audio interface, digital I/O control hardware, and custom software are available upon request.

The song stimuli used throughout this study were built using 16 unique (non-repetitious) motifs from the song recordings of one adult starling. During this task, 6 isolate motifs and 3 paired motifs served as S+ stimuli, while a different set of 6 isolate and 3 paired complements served as S- stimuli. During training, several isolate or pairs of motifs remained reserved for probe testing. The design of this task allowed us to ask if starlings in this paradigm spontaneously acquire knowledge of the patterning rule that is *not* tied, and in fact contradicts, the perceptual associations between the stimuli, or whether they rely on strategies constrained to learned associations between the stimuli and their consequences. If, for example, the stimuli A and B were not reinforced, then the compound AB (or BA) would be. Contrarily, if stimuli C and D were reinforced, then the compound stimulus CD (or DC) would not be reinforced. Crucially, the training set included certain uncomplemented stimuli such as E+ and F+, GH- and HG-. Following learning, researchers tested how individuals performed on stimuli such as EF, FE, G and H. Following learning, we tested how starlings performed on stimuli such as EF, FE, G and H. As opposed to utilizing a rule-based strategy, starlings relied strongly on associative strengths to generalize task knowledge (single sample t-test, where chance of d' = 0: t = -7.31, df = 3, *p* = 0.005). Interestingly, the strength of these associations, as measured by d', correlate with the number of times subjects visit the S+ perch on the first day of preference testing (more than 6 months later; supplemental fig 8).

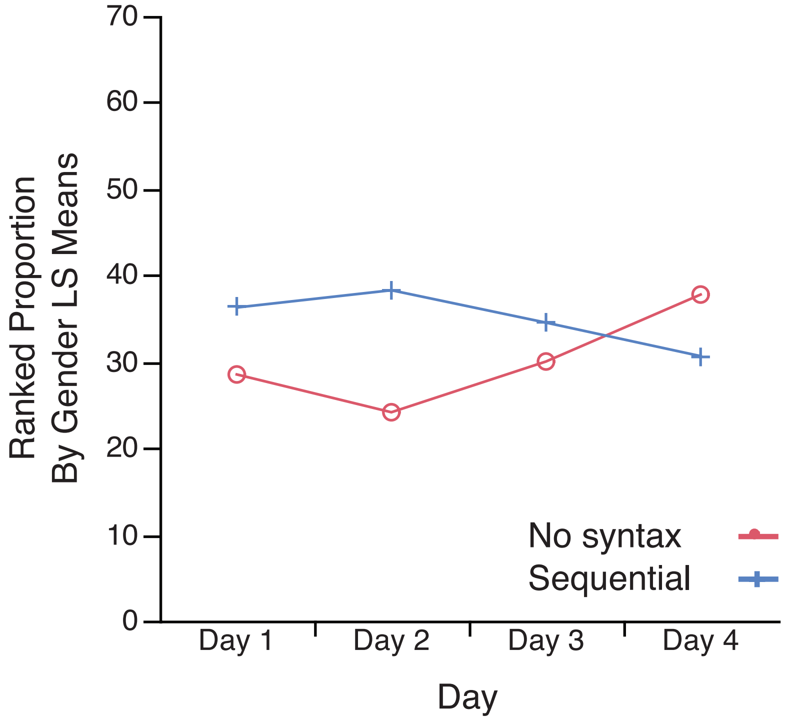

Figure 2.4. Rank transformed proportions for males on sequential syntax versus no syntax preference task. Males did not reveal any demonstrable preference for one stimulus compared with the other (F(3, 49) = 1.178, *p* = 0.3277).

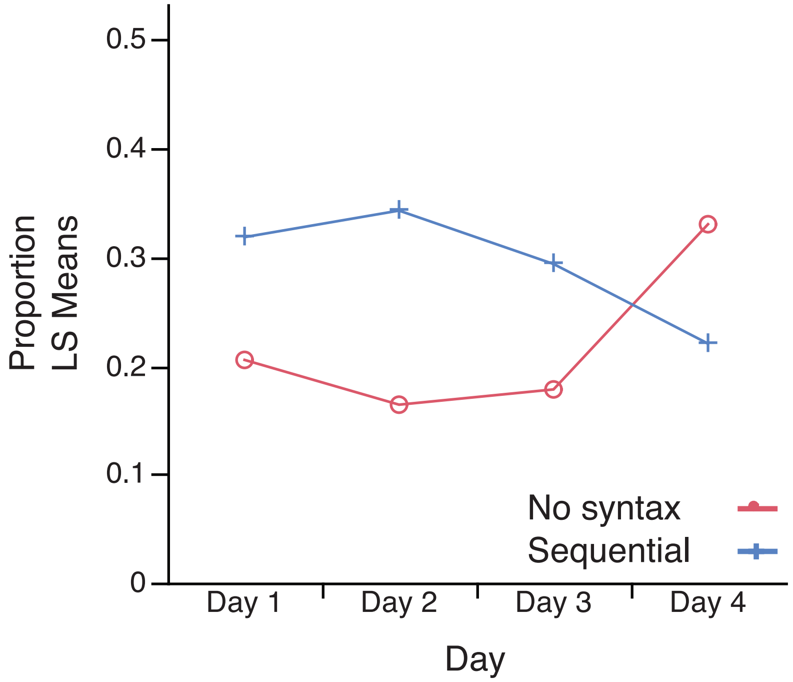

Figure 2.5. Raw proportions for males on sequential syntax versus no syntax preference task. Males did not reveal any demonstrable preference for one stimulus compared with the other (F(3, 49) = 1.2615, *p* = 0.2979).

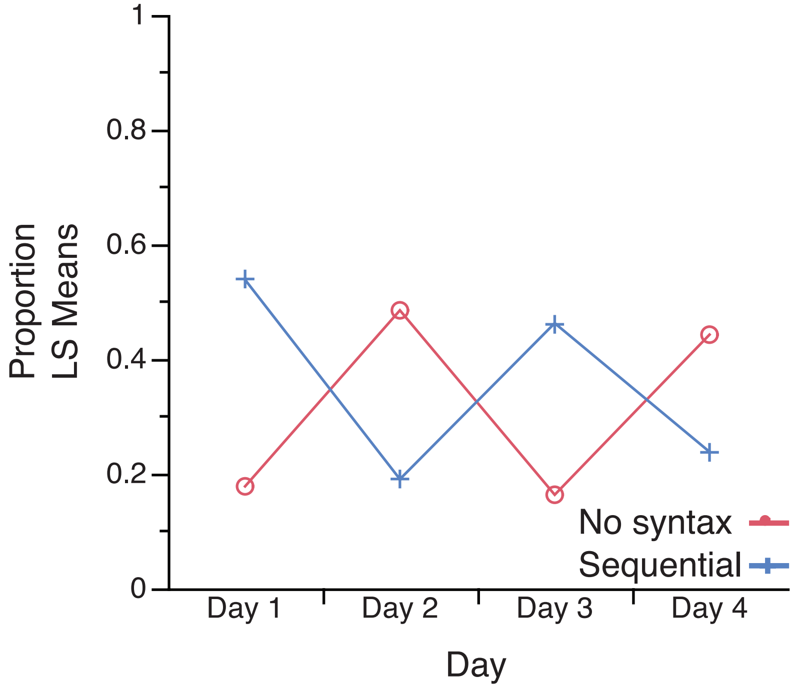

Figure 2.6. Raw proportions for females on sequential syntax versus no syntax preference task. Females showed a strong preference for the nestbox initially associated with the sequential syntax (F(3, 42) = 6.3872, *p* = 0.0012).

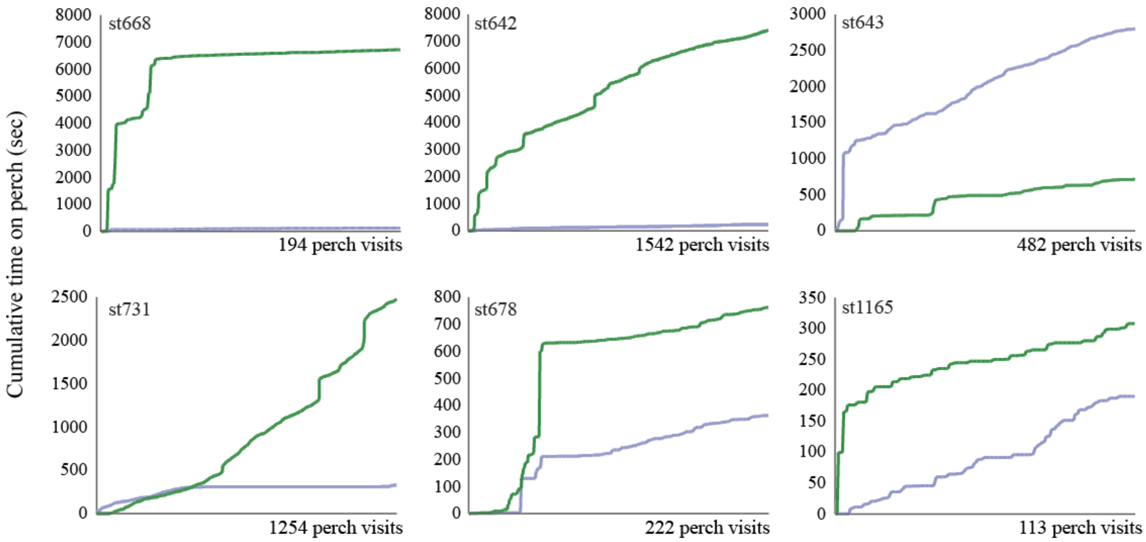

Figure 2.7. Cumulative time spent of the sequential syntax (green) versus no syntax (purple) perches across the first day. Data for females only.

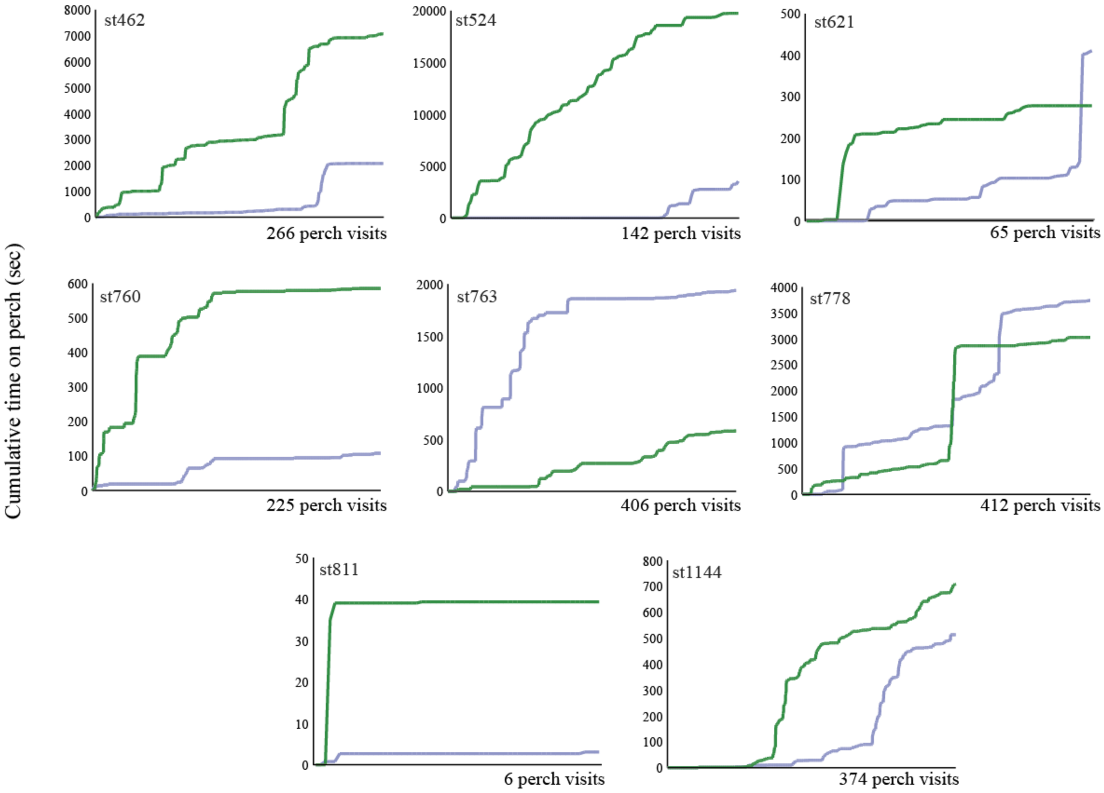

Figure 2.8. Cumulative time spent of the sequential syntax (green) versus no syntax (purple) perches across the first day. Data for males only.

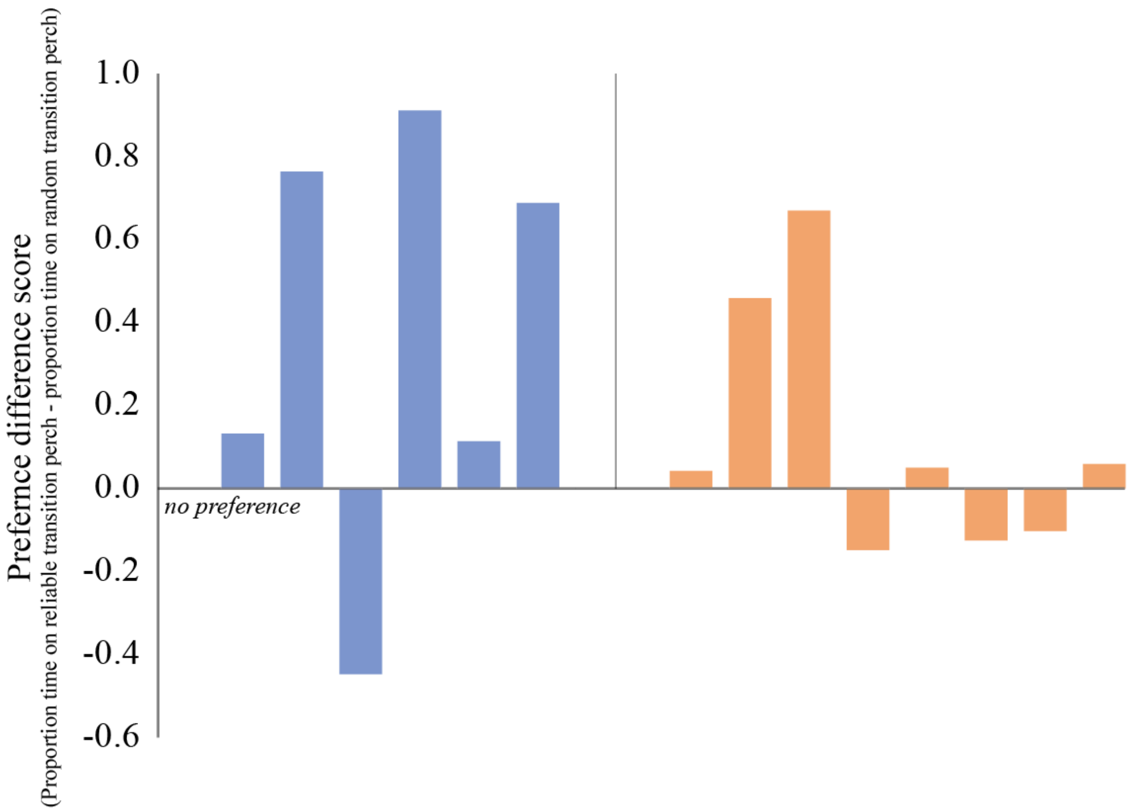

Figure 2.9. Difference scores representing time spent on the nestbox associated with sequential syntax versus no syntax after the first day. Data for females in purple, males in orange.

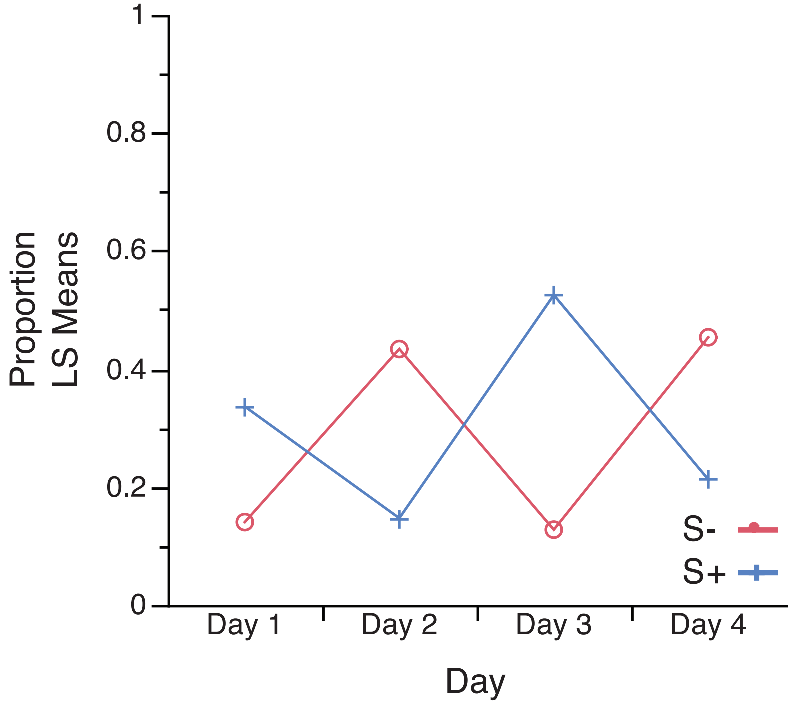

Figure 2.10. Raw proportions on S+ versus S- preference task. The nest box initially associated with S+ stimuli was preferred (F(3, 21) = 5.0310, *p* = 0.0088).

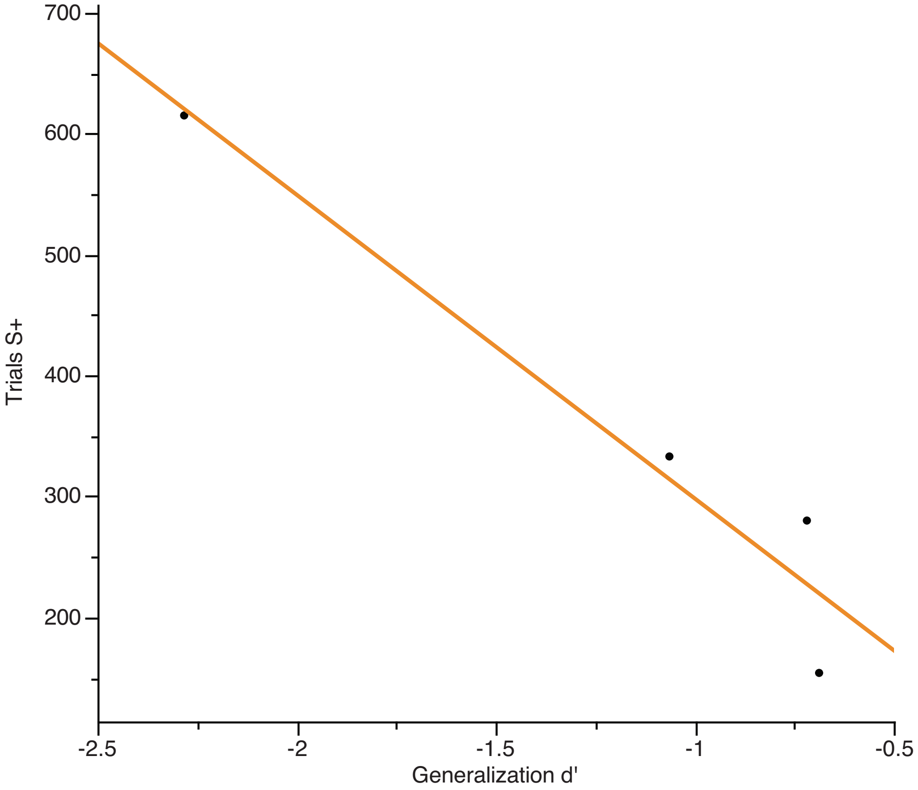

Figure 2.11. Associative strength on the go-nogo task, as measured by generalization testing, correlated with the number of perch visits paid to the nestbox associated with S+ stimuli on day 1.

Acknowledgements

This chapter, in full, is currently being prepared for submission for publication of the material. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

References

Aslin RN, Saffran JR, Newport EL (1998) Computation of Conditional Probability Statistics by 8-Month-Old Infants. Psychological Science 9:321–324. doi: 10.1111/1467-9280.00063

Ball GF, Ketterson ED (2008) Sex differences in the response to environmental cues regulating seasonal reproduction in birds. Philosophical transactions of the Royal Society of London Series B, Biological sciences 363:231–46. doi: 10.1098/rstb.2007.2137

Briefer E, Osiejuk TS, Rybak F, Aubin T (2010) Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. Journal of theoretical biology 262:151–64. doi: 10.1016/j.jtbi.2009.09.020

Chiandetti C, Vallortigara G (2011) Chicks like consonant music. Psychological science 22:1270–3. doi: 10.1177/0956797611418244

Christiansen MH, Chater N (2008) Language as shaped by the brain. The Behavioral and brain sciences 31:489–508; discussion 509–58. doi: 10.1017/S0140525X08004998

Conover W, Iman R (1981) Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics. The American Statistician 35:124–129.

Conway C, Christiansen M (2001) Sequential learning in non-human primates. Trends in cognitive Sciences 5:539–546.

Darwin C (1871) The descent of man and selection in relation to sex. 17:1223.

Deacon TW (2010) Colloquium paper: a role for relaxed selection in the evolution of the language capacity. Proceedings of the National Academy of Sciences of the United States of America 107 Suppl :9000–6. doi: 10.1073/pnas.0914624107

Eens M, Pinxten R, Verheyen RF (1991) Male Song as a Cue for Mate Choice in the European Starling. Behaviour 116:210–238.

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86.

Gentner T, Hulse S (1998) Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Animal Behaviour 56:579–594.

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Animal Behaviour 59:443–458.

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. Journal of neurobiology 46:48–58.

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 16:319–326.

Kagawa H, Yamada H, Lin R. (2012) Ecological correlates of song complexity in white-rumped munias: The implication of relaxation of selection as a cause for signal variation in birdsong. Interaction Studies 13:263–284. doi: 10.1075/is.13.2.05kag

Komdeur J, Oorebeek M, van Overveld T, Cuthill IC (2005) Mutual ornamentation, age, and reproductive performance in the European starling. Behavioral Ecology 16:805–817. doi: 10.1093/beheco/ari059

Maney DL, Goode CT, Lange HS. (2008) Estradiol modulates neural responses to song in a seasonal songbird. The Journal of comparative neurology 511:173–86. doi: 10.1002/cne.21830

Mountjoy D (1996) Female choice for complex song in the European starling: a field experiment. Behavioral Ecology and Sociobiology

Mountjoy JD, Lemon RE (1995) Extended song learning in wild European starlings. Animal Behaviour 49:357–366. doi: 10.1006/anbe.1995.0048

Saffran J, Aslin R, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274:1926–1928.

Searcy WA, Nowicki S (2005) Song and mate choice in birds: how development of behavior helps us understand function. The Auk 122:1–14.

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois

CHAPTER THREE

Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition

Comins, JA and Gentner, TQ

As it appears in

Animal Cognition

2014

Digital Object Identifier (DOI): 10.1007/s10071-014-0732-5

Anim Cogn
DOI 10.1007/s10071-014-0732-5

ORIGINAL PAPER

Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition

Jordan A. Comins · Timothy Q. Gentner

Received: 3 September 2013 / Revised: 21 December 2013 / Accepted: 3 February 2014

Abstract The sequential patterning of complex acoustic elements is a salient feature of bird song and other forms of vocal communication. For European starlings (*Sturnus vulgaris*), a songbird species, individual vocal recognition is improved when the temporal organization of song components (called motifs) follows the normal patterns of each singer. This sensitivity to natural motif sequences may underlie observations that starlings can also learn more complex, unnatural motif patterns. Alternatively, it has been proposed that the apparent acquisition of abstract motif patterning rules instead reflects idiosyncrasies of the training conditions used in prior experiments. That is, that motif patterns are learned not by recognizing differences in temporal structures between patterns, but by identifying serendipitous features (e.g., acoustical cues) in the small sets of training and testing stimuli used. Here, we investigate this possibility, by asking whether starlings can learn to discriminate between two arbitrary motif patterns, when unique examples of each pattern are presented on every trial. Our results demonstrate that abstract motif patterning rules can be acquired from trial-unique stimuli and suggest that such training leads to better pattern generalization compared with training with much smaller stimulus subsets.

Keywords Vocal recognition · Pattern learning · Auditory cognition · Trial-unique stimuli

Introduction

Recognizing familiar individuals is essential for adaptive social behavior. In songbirds, vocalizations serve as primary sensory signals used to identify others. This role is reflected by the sensitivity of neurons throughout the avian forebrain to conspecific songs and their associated behavioral goals (Mello et al. 1992). For species with acoustically complex songs, numerous components of the signal can carry information about individual identity (Knudsen and Gentner 2010). In the case of European starlings (*Sturnus vulgaris*), for instance, males sing elaborate temporally patterned songs built from stereotyped units called motifs, where each motif itself is a patterned arrangement of notes. Thus, starling song unfolds as a sequence of changing complex auditory events (Meliza et al. 2010). Starlings rely on identification of singer-specific motifs and their serial arrangement for successful vocal recognition (Gentner and Hulse 1998). While the behavioral and electrophysiological mechanisms supporting auditory object recognition are well established (Gentner and Margoliash 2003; Jeanne et al. 2011, 2013; Knudsen and Gentner 2013; Meliza et al. 2010; Meliza and Margoliash 2012; Thompson and Gentner 2010), less is known about how the nervous system of songbirds (or any other animal) represents behaviorally relevant patterns in vocal

J. A. Comins · T. Q. Gentner (✉)
Department of Psychology, University of California,
San Diego (UCSD), 9500 Gilman Drive MC0109, La Jolla,
CA 92092-0109, USA
e-mail: tgentner@ucsd.edu

J. A. Comins
Center for Academic Research and Training in Anthropogeny,
La Jolla, CA, USA

T. Q. Gentner
Neurosciences Program, University of California,
San Diego (UCSD), La Jolla, CA, USA

T. Q. Gentner
Kavli Institute for Brain and Mind, La Jolla, CA, USA

sequences (Kiggins et al. 2012). We suggest that understanding what information is acquired implicitly from such patterns during learning provides the best guide to understanding how complex vocal sequences are represented in neural circuits.

Recent behavioral results suggest that starlings recognize motif sequences by learning their underlying pattern structures (Comins and Gentner 2013; Gentner et al. 2006). Others, however, caution against such conclusions—suggesting that certain alternative methods for stimulus discrimination have not been rejected (Beckers et al. 2012; Corballis 2009). Specifically, experimental designs using repeated exposure to small sets of pattern exemplars might introduce unintended acoustic cues that subjects could use to distinguish pattern types, rather than differences in abstract structures (ten Cate and Okanoya 2012; van Heijningen et al. 2009). As a result, conclusions of the form "subjects differentiate pattern type *X* and pattern type *Y*" may be stated more accurately as "subjects differentiate only specific sequences from pattern type *X* and *Y*". Here, we examine the necessity of the forgoing solution strategy by determining whether European starlings can learn to recognize different patterns through exposure to very large sets of pattern exemplars.

Starlings are expert auditory sequence learners under conditions of both natural song development (Eens et al. 1988; Eens and Pinxten 1992; Mountjoy and Lemon 1995) and in the laboratory (Comins and Gentner 2010; Gentner and Hulse 1998, 2000; Knudsen et al. 2010). Most germane to the current work, starlings can accurately classify artificial motif patterns that take the forms AABB and ABAB (Gentner et al. 2006), where A and B represent sets of ethologically determined motif categories (Eens et al. 1988) termed warbles and rattles. These pattern recognition capabilities persist for even more complicated pattern arrangements. Starlings recognize the patterning forms XXYY and XYXY, where on any given trial an X could represent a motif from set A or set B (and vice versa for Y) and therefore distinguish AABB and BBAA patterns from ABAB and BABA (Comins and Gentner 2013). Use of the XXYY/XYXY patterns precludes a number of simpler solution strategies available in the AABB/ABAB task where, for instance, the animal might rely on information at a single sequence position (e.g., the second element) to classify 4-motif sequences (Comins and Gentner 2013; Gentner et al. 2006). In XXYY/XYXY patterns, every element can occur at every location, and thus the animal is *minimally* required make a decision on the basis of the relationship between *at least* two or more motifs. In the prior study, however, only 32 XXYY and XYXY stimuli (16/pattern) out of 16,384 possible patterned stimuli were used for training. Even when considering the larger subset of patterns used to test generalization abilities ($N = 500$) in this earlier study, this set constitutes a very small percentage of the possible patterns (<3.5%).

Here, we used an operant conditioning procedure to train subjects to classify patterns of the form XXYY and XYXY. Unlike in previous reports of pattern recognition in this species, every animal was presented with nearly the entire population of pattern combinations. Each subject encountered at least 16,300 out of a possible 16,384 patterns during classification training, and each individual pattern no more than twice over the entire experiment (excluding correction trials—see "Methods" section). This experimental design maximizes stimulus diversity and minimizes repetition within the set of well-formed patterns, and thereby ensures that if subjects learn to classify XXYY and XYXY patterns correctly, it is only by recognizing differences in the abstract structures governing the temporal organization of motifs. Our results are consistent with this interpretation, and contradict the notion that the learning of abstract temporal patterns can be explained by attention to acoustic cues serendipitously found in restricted stimulus subsets.

Methods

Subjects

Three European starlings, wild-caught near the Los Angeles International Airport (LAX) in California, served as subjects. Prior to being tested, subjects were entirely naïve to the motifs used to generate patterned stimuli in this experiment. All subjects were at least 1 year old as indicated by their adult plumage (Feare 1984, 1996); the sex of subjects was not controlled in this study. From the time of capture until use in this study, subjects were housed in a large mixed sex, conspecific aviary with ad libitum access to food and water. The photoperiod in the aviary and the operant chambers followed the seasonal variation in local sunrise and sunset times.

Stimuli

Patterned stimuli were constructed using motifs from the song recordings of one adult male starling that was captured near Baltimore, Maryland (see Gentner and Hulse 1998). We extracted 16 distinct motifs (non-repetitious; see Meliza 2011) and separated these motifs into two sets, labeled A and B. Motif membership in set A and B was based on natural acoustic (perceptual) category boundaries (i.e., 8 warbles for set A; 8 rattles for set B). Whereas warbles and rattles both possess a heterogeneous acoustic structure, rattles are also characterized by the occurrence of a broadband click-train. Crucially, these motif categories

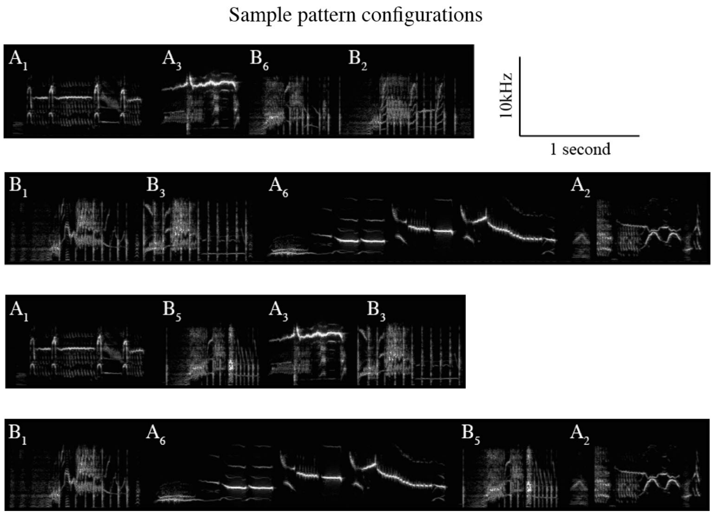

Fig. 1 Example motif sequence stimuli from the AABB and ABAB pattern classes. *Letters* denote motif classes (A: warbles, B: rattles) and *subscript numbers* (1–8) denote unique motifs within each class used to generate patterned sequences. A total of 16,300 (out of 16,384 possible) XXYY and XYXY sequences were used (4,075 AABB, 4,075 BBAA, 4,075 ABAB and 4,075 BABA)

preserve ethologically relevant and psychologically salient boundaries (Braaten 2000; Eens et al. 1988) as the integrity of these categorical boundaries is argued to be necessary to demonstrate recognition of abstract pattern structures (Comins and Gentner 2013).

Using these two classes of motifs, we made patterns of motif sequences following the forms XXYY or XYXY. We generated all possible combinations of XXYY and XYXY patterned stimuli using eight warble (hereafter, motif set A) and eight rattle (hereafter, motif set B) motifs, where on any given trial an X could represent a motif from either set A or set B (and vice versa for Y), but never both. Therefore, subjects were required to distinguish AABB and BBAA patterns from ABAB and BABA yielding a total of 16,384 sequences (4,096 AABB–4,096 BBAA and 4,096 ABAB–4,096 BABA; Fig. 1—further details about these pattern constructions can be found in Comins and Gentner 2013).

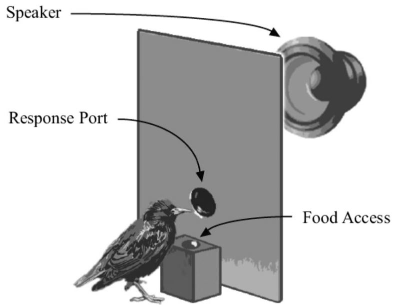

Fig. 2 Schematic of the operant apparatus. Subjects start a trial by pecking the center response port. After the motif pattern stimulus ends, the subject either pecks the center response port again or withholds any response depending on the class from which the stimulus was drawn. Correct responses yield a food reward. Incorrect responses lead to the house light being extinguished and food being inaccessible

Apparatus

The custom-built operant apparatus where starlings learned to classify the patterned stimuli is portrayed in Fig. 2. Each subject was held in a small weld-wire cage (41 × 41 × 35 cm) with a 30 × 30 cm operant panel mounted on one wall, mounted inside a 61 × 81 × 56 cm ID sound attenuation chamber (Acoustic Systems). The operant panel contained a centrally located, PVC-lined response port, roughly 14 cm off the floor of the cage. Inside the opening of the response port was an IR emitter–receiver that enabled detection of precise times when the bird broke the plane of the response port with its beak. This "poke-hole" design allowed starlings to probe the apparatus with their beaks, in a manner akin to their natural appetitive foraging behavior. Directly below the response port, in the section of cage floor immediately adjacent to the panel, another PVC-lined opening provided access to food. A remotely controlled hopper, positioned behind the panel, moved the food within and beyond the subject's reach beneath the opening. Acoustic stimuli were delivered through a small full-range audio speaker mounted behind the panel and out of the subject's view. The sound pressure level inside all operant chambers was calibrated to a standard broadband noise signal. Custom hardware and software monitored the subjects' responses, controlled the delivery of stimuli (16 bit resolution, 44.1-kHz sample rate), access to food, and lighting inside the chamber according to experimental contingencies.

Shaping procedure

Upon initially entering the operant chamber, we provided each subject with unrestricted access to the food hopper. Following acclimation to eating from this device, the hopper was lowered beyond the subject's reach. Next subjects were placed on an autoshaping routine (Brown and Jenkins 1968). Between two and three times per minute, we presented the subject with a blinking LED in the center response port followed by 2-s access to food. This process recurred until the subject acquired a key-peck response. From then, the lowered food hopper would only be engaged if the subject pecked the blinking LED in the center port. Subjects repeated this behavior for 100 trials. After the completion of this phase, the center LED ceased blinking, requiring subjects to peck at the darkened center port to raise the food hopper. Following 100 such trials, pecking the center port initiated the playback of an acoustic stimulus where subjects earned food rewards in accordance with standard go/no-go training procedures described below.

Training procedure

We trained subjects to classify the S+ and S− stimuli using a standard go/no-go operant conditioning paradigm. Two subjects had one set of patterned stimuli serving as the S+ (XXYY), while the remaining subject had the other set of patterned stimuli (XYXY) serving as the S+. Subjects initiated a trial by pecking a small response port to start playback of a stimulus (see Fig. 1). For half of the training stimuli (S+), the subject was trained to peck the response port after playback completed to obtain a 2-s access to food. For the other half of the training stimuli (S−), the subject was trained to withhold pecks to the response port to avoid a mild punishment (extinguished house lights for 2 or more seconds). Correctly withholding pecks was not reinforced with food. False alarms (pecking to S− stimuli) initiated a correction sequence in which the same stimulus was repeated on subsequent trials until the subject correctly withheld a response. Only data from non-correction trials were analyzed here.

Given that data were to be analyzed in blocks of 100 trials, each subject was initially presented 16,300 of all possible 16,384 patterns. This first round of training ensured that each pattern was only encountered one time per bird (excluding correction trials). However, due to relatively poor performance at this point (see "Results" section), we generated and presented subjects with another 16,300 patterns of these same motifs. Thus, by the end of data collection, every bird heard at least 97.5 % of all possible patterns, wherein no specific sequence was ever encountered more than two times (again, excluding correction trials). In total, subjects received 17,879 ± 3,433 (μ ± SE) correction trials over the course of training.

Analysis

Percent correct served as our metric to quantify learning (sum of correct "Go" responses to "Go" stimuli and correct "NoGo" responses to "NoGo" stimuli divided by the total number of responses). All data were analyzed in blocks of 100 trials. We assessed learning with overall performance accuracy for different pattern classes using matched-samples t tests and repeated measures ANOVA.

Results

Pattern learning

During the initial 163 100-trial blocks of training, performance for all three subjects was poor. By the last 10 blocks on this first round of training (blocks no. 154–163), none of the subjects were performing significantly better than on the first 10 blocks of training (matched-pairs t test: in all cases $df = 9$, $P > 0.05$) and only one subject was classifying the pattern stimuli better than expected by chance (single-sample t test against chance of 50 %: subject B851 $df = 9$, $P = 0.03$; both others: $df = 9$, $P > 0.05$). Thus, we presented subjects with another random arrangement of 16,300 XXYY and XYXY patterned motif sequences.

Our main findings are shown in Fig. 3. By the final 10 blocks of the second round of training (blocks no. 317–326), all three subjects were classifying the patterned stimuli significantly better than expected by chance (single-sample t test against chance of 50% correct: in all cases, $df = 9$, $P < 0.0001$; Subject 851's performance 68.2 ± 1.41; Subject 852's performance 64.9 ± 1.92; Subject 877's performance 67.4 ± 2.7; mean ± SE). In addition, performance for all subjects during the final 10 blocks of the second round of training was significantly better than performance during the first 10 blocks of training (matched-pairs t test: in all cases, $df = 9$, $P < 0.0004$; Fig. 3). Finally, hit and false alarm rates for all three subjects changed in consistent ways along with performance. Hits and false alarms were high during initial training initially, and over the course of learning the false alarm rate slowly reduced for each animal. This is

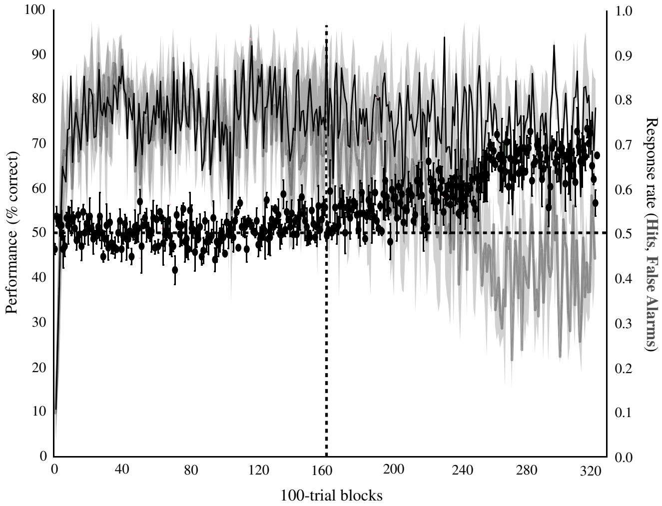

Fig. 3 XXYY and XYXY pattern classification performance averaged across subjects. Mean (±SE) percent correct (*black circles*) increases gradually across blocks 164 through 326. Mean (± SE) hit rate (*black line*) and false alarm rate (*gray line*) diverge as the overall percent correct increases. The *dotted horizontal line* denotes chance performance; the *vertical dotted line* denotes the point where we began to cycle through our trial-unique stimuli for a second time

consistent with all the subjects learning to discriminate between the XXYY and XYXY patterns in similar ways.

Pattern learning strategies

One possible explanation for the learning is that subjects focused on a subset of the elements in the sequence. Several such strategies are possible. For instance, subjects might: (1) determine whether the first two motifs are in the same category, (2) whether the last two motifs are in the same category, (3) whether the first and third motifs are in the same category or (4) whether the second and fourth motifs are in the same category.

To examine the likelihood of these "pattern subset" strategies, we took advantage of the trial-unique stimulus design, where each trial is in effect a generalization test with novel stimuli. Because each of the pattern subset strategies involves comparisons between pairs of elements, we reasoned that starlings' classification should improve when the two elements being compared are identical motifs. For example, if subjects were comparing the first two elements (strategy 1), then the sequence $A_1A_1B_3B_4$, where the first two elements are the same motif, should be more accurately classified than $A_1A_2B_3B_4$, where the first two elements are not identical acoustically. This is because the perceptual task of deciding whether A_1 and A_1 are in the same category is trivial contrasted with A_1 and A_2. We searched for evidence supporting the use of any of these strategies across the final 10 blocks of training in each subject. To examine strategy 1, for instance, we took all XXYY patterns during these final 10 blocks and compared performance on those trials where the same motif occurred in the first and second position of the pattern to those XXYY trials where the first two motifs differed. The process was repeated for strategies 2–4 using XXYY, XYXY and XYXY patterns, respectively. For each subject, we failed to detect a significant advantage in response accuracy based on any of the pattern subset strategies (all Pearson's Chi-squared tests; see Table 1). These results are

Table 1 Comparison of performance for each subject across a variety of stimulus arrangements

Subject	Repetition	No repetition	Chi-square test
Strategy 1: initial pair			
B851	37/66 correct	277/447 correct	$\chi^2 = 0.845$, $P = 0.356$
B852	23/53 correct	223/460 correct	$\chi^2 = 0.492$, $P = 0.483$
B877	55/70 correct	360/445 correct	$\chi^2 = 0.209$, $P = 0.647$
Strategy 2: final pair			
B851	34/58 correct	280/455 correct	$\chi^2 = 0.184$, $P = 0.668$
B852	30/60 correct	216/453 correct	$\chi^2 = 0.114$, $P = 0.736$
B877	58/67 correct	357/448 correct	$\chi^2 = 1.763$, $P = 0.184$
Strategy 3: first and third			
B851	55/69 correct	313/418 correct	$\chi^2 = 0.748$, $P = 0.387$
B852	45/57 correct	358/430 correct	$\chi^2 = 0.655$, $P = 0.419$
B877	37/76 correct	222/409 correct	$\chi^2 = 0.806$, $P = 0.369$
Strategy 4: second and fourth			
B851	56/70 correct	312/417 correct	$\chi^2 = 0.871$, $P = 0.351$
B852	45/59 correct	358/428 correct	$\chi^2 = 1.975$, $P = 0.16$
B877	32/66 correct	227/419 correct	$\chi^2 = 0.742$, $P = 0.389$

None of the subjects showed better performance when patterns consisted of motif repeats in any configuration consistent with the conclusion that subjects use similarities and differences between three or more elements to classify each patterned sequence.

Another possible solution strategy would be to use all the elements, but focus only on a subset of the sequences. For example, subjects might achieve above chance performance by accurately discriminating only the sequences that begin with A or only those that begin with B. To test this, we compared performance on AABB versus BBAA patterns as well as ABAB versus BABA patterns. In general, these comparisons did not indicate that starlings performed better on patterns that began with an A or B category of motif (Subjects B851: AABB vs. BBAA $\chi^2 = 2.028$, $P = 0.15$, ABAB vs. BABA $\chi^2 = 0.36$, $P = 0.549$; Subject B852: AABB vs. BBAA $\chi^2 = 2.3$, $P = 0.129$, ABAB vs. BABA $\chi^2 = 0.174$, $P = 0.676$; Subject B877: AABB vs. BBAA $\chi^2 = 1.92$, $P = 0.166$, ABAB vs. BABA $\chi^2 = 7.296$, $P = 0.007$). Thus, with the exception of ABAB versus BABA patterns for subject B877, it appears all pattern types were equally well learned by all subjects. Finally, we examined the possibility that subjects relied on an alternation strategy by detecting transitions between X and Y elements. To test this, we reasoned that patterns of the form XXYY would be more difficult to classify than XYXY patterns, because the latter have more X–Y transitions. Inconsistent with this hypothesis, the error rates for classifying XXYY and XYXY patterns were not significantly different (matched-pairs *t* test, $P = 0.74$).

Comparison with prior pattern generalization results

One potential consequence of using a training regimen that samples thoroughly from the full set of potential pattern stimuli is the improved recognition of the underlying abstract structure common to all sequences. To test this hypothesis, we compared performance of the three subjects used in the current experiment to starlings trained to recognize these same XXYY/XYXY patterns using only a small subset of all possible sequence combinations. In that experiment, four subjects were initially trained to distinguish 16 XXYY and 16 XYXY patterns. These animals were then presented with 500 novel XXYY and XYXY patterned sequences to measure recognition of the underlying pattern structure (Comins and Gentner 2013).

First, we compared the acquisition rates for subjects trained to recognize patterns using the 32 exemplar stimulus set to those in the present study. The birds trained with 32 of the possible 16,132 patterned motif sequences learned significantly faster than those trained with the full set of sequences (nested rmANOVA; training regimen*training block interaction, $F_{(1,1562)} = 141.6$, $P < 0.001$). Moreover, the birds trained on the restricted stimulus set reached a significantly higher mean (±SEM) level of performance (74.9 ± 2.48, averaged over the last five blocks of training), than the subjects in the current study (64.7 ± 1.23, average over the last 500 trials; $t = 3.28$, $df = 5$, $P = 0.02$; Fig. 4). When tested on a set of 500 trial-unique patterned sequences, however, the mean percent correct of the birds trained on restricted sets dropped to 57.25 ± 0.718, which, although above chance, is significantly *below* that of the subjects in the present study over the last 500 trials ($t = -5.54$, $df = 5$, $P = 0.002$; Fig. 4). This is consistent with the interpretation that training with trial-unique stimuli enhances pattern generalization.

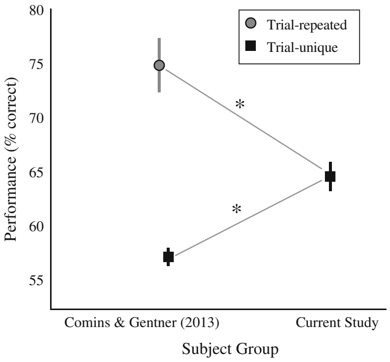

Fig. 4 Mean (±SE) accuracy (% correct) for the classification of the final 500 trial-unique (*black squares*) XXYY/XYXY patterned stimuli for subjects in the current study (*right*) and previous work (*left*) where subjects were trained to a high level of performance with a small set of trail-repeated (*gray circle*) pattern exemplars. *Lines* indicate comparisons between conditions, $*P < 0.05$, see text for statistics

Discussion

The results of the current study demonstrate that knowledge of abstract sequential patterning among acoustic categories can be acquired from large trial-unique stimulus sets with minimal repetition. Although the explicit patterning rule acquired in the present case cannot be unequivocally stated, our analyses indicate that it involves comparisons among three or more element classes independent of their absolute position in the sequence. The simple demonstration that patterning rules can be acquired through training with trial-unique stimuli is inconsistent with the idea that pattern recognition in starlings is driven by serendipitous acoustic features in restricted subsets of patterned stimuli (ten Cate and Okanoya 2012), and lends further support the conclusions of prior pattern recognition studies (Comins and Gentner 2013; Gentner et al. 2006).

Understanding the implicit patterning rules acquired by animals through operant training requires the analysis of performance on generalization trials. Because in the present study each stimulus exemplar was only presented twice, with each presentation separated by several thousands of trials, the opportunity to learn explicit stimulus–response associations was minimal. Thus, each trial can be considered a kind of generalization test of the animal's experience up to that point, and successful classification of motif patterns over a series of such trials indicates that *some* rule consistent with the patterns has been acquired. By making a series of post hoc comparisons between responses to specific subsets of motif sequences during the last 1,000 trials, we were able to exclude a number of potential rules linked to pair-wise comparisons among elements at specific sequence locations. Although recent reports indicate that zebra finches (*Taeniopygia guttata*), a related species of songbird, rely heavily on element repetition to solve similar pattern discrimination tasks (Van Heijningen et al. 2012), this or other similar rules do not explain the current results.

Our results also highlight the important role that training can play in generalization performance. In the present case, we observed remarkably robust discrimination of novel patterned motif sequences at the end of training that was significantly higher than that observed in an earlier study, where the subjects were trained with much smaller stimulus sets. This matches the well-known, but poorly studied, observation that training with more exemplars yields broader generalization in category discrimination (Wasserman and Astley 1994), and suggest that such effects extend to abstract features of categories as well. It seems reasonable to hypothesize that this enhanced generalization derives from a less explicit association between the pattern and the underlying acoustics of each element. Such associations are likely to be reinforced by training with restricted stimulus sets, or when training involves a piecemeal introduction of stimuli (Van Heijningen et al. 2009). Our results suggest that both of these factors likely impact pattern learning and subsequent generalization performance, though more detailed and well-controlled comparisons of pattern generalization need to be conducted.

Neurophysiology of pattern recognition

Given the starlings' persistent pattern recognition behavior across experimental designs (Comins and Gentner 2010, 2013; Gentner et al. 2006; Gentner and Hulse 1998), a key future direction is to investigate the instantiation of these learning mechanisms physiologically. It is already clear that associative learning in starlings modifies the response properties of neurons throughout the avian telencephalon. Firing rates of single neurons and populations of single neurons in several forebrain regions analogous mammalian auditory cortices (e.g., caudomedial and caudolateral mesopallium, and caudomedial nidopallium, NCM), are modulated by the behavioral relevance of song motifs (Chew et al. 1995; Gentner and Margoliash 2003; Jeanne et al. 2011, 2013; Knudsen and Gentner 2013; Meliza et al. 2010; Meliza and Margoliash 2012; Thompson et al. 2012; Thompson and Gentner 2010).

A recent characterization of neuronal selectivity and tolerance across six primary and non-primary auditory areas in starlings to learned and unlearned songs offers new clues as to the representation of learned sequences (Meliza and Margoliash 2012). The stimuli used in these experiments were natural songs, which contain multiple renditions of the same motif occurring in several temporal positions. It was demonstrated that NCM, while highly selective in its response profile (a normal feature of non-primary sensory processing areas), showed almost no "tolerance" for the same motif occurring in different positions of song. In other words, the response of a neuron in NCM varied considerably to the same stimulus occurring in multiple temporal positions of the song. Rather than indicating neuronal tolerance per se, the cells in NCM might, as Meliza and Margoliash (2012) suggest, be highly sensitive to the global temporal context in which motifs occur. Whether such contextual modulation is tied to learning is not known, but NCM is an attractive target for future investigations of sequence learning in starlings.

Conclusions

We conclude that sophisticated acoustic recognition abilities of starlings include the capacity to learn abstract rules governing the temporal patterning of song elements. These results will be informative for future work examining the neurobiology of auditory sequence learning in songbirds, and the evolution of these pattern recognition mechanisms across species. Indeed humans too are extraordinary auditory pattern learners. Only hours after birth, infants are capable of detecting statistical regularities from sequences of speech sounds (Teinonen et al. 2009), and by 7 months of age can extract underlying abstract rules governing auditory patterns (Marcus et al. 1999). These powerful learning mechanisms are thought to lay the groundwork for acquiring knowledge of uniquely human faculties such as music (Hannon and Trehub 2005; Saffran et al. 1999) and language (Marcus et al. 2007; Marcus 2000). Our findings illuminate potential parallels with human work showing that sequence learning is buttressed by both knowledge of lower-level perceptual organization of pattern elements

(Emberson et al. 2013) and an understanding of abstract structures (Marcus et al. 1999, 2007). These parallels suggest that starlings can serve as a non-human model system to examine the neurobiological implementation of pattern recognition at cellular and circuit levels.

Acknowledgments This work supported by NSF Graduate Research Fellowship 2011122846 to JAC and NIH DC008358 to TQG. We thank Patrick Plummer and Stephanie Carmack for comments on an earlier draft and Scott Freeman for technical expertise.

References

Beckers GJL, Bolhuis JJ, Okanoya K, Berwick RC (2012) Birdsong neurolinguistics: songbird context-free grammar claim is premature. NeuroReport 23:139–145. doi:10.1097/WNR.0b013e32834f1765

Braaten RF (2000) Multiple levels of representation of song by European starlings (*Sturnus vulgaris*): open-ended categorization of starling sing types and differential forgetting of song categories and exemplars. J Comp Psychol 114:61–72. doi:10.1037//0735-7036.114.1.61

Brown P, Jenkins H (1968) Auto-shaping of the pigeon's keypeck. J Exp Anal Behav 11:1–8

Chew SJ, Mello C, Nottebohm F et al (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. Proc Nat Acad Sci USA 92:3406–3410

Comins JA, Gentner TQ (2010) Working memory for patterned sequences of auditory objects in a songbird. Cognition 117:38–53

Comins JA, Gentner TQ (2013) Perceptual categories enable pattern generalization in songbirds. Cognition 128:113–118. doi:10.1016/j.cognition.2013.03.014

Corballis MC (2009) Do rats learn rules? Anim Behav 78:e1–e2

Eens M, Pinxten R (1992) No overlap in song repertoire size between yearling and older starlings *Sturnus vulgaris*. Ibis 134:72–76

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86

Emberson LL, Liu R, Zevin JD (2013) Is statistical learning constrained by lower level perceptual organization? Cognition 128:82–102. doi:10.1016/j.cognition.2012.12.006

Feare C (1984) The starling. Oxford University Press, London

Feare C (1996) Studies of West Palearctic birds 196. Common starling *Sturnus vulgaris*. Br Birds 89:549–568

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

Gentner TQ, Hulse SH (2000) Perceptual classification based on the component structure of song in European starlings. J Acoust Soc Am 107:3369–3381

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440:1204–1207

Hannon EE, Trehub SE (2005) Tuning in to musical rhythms: infants learn more readily than adults. Proc Nat Acad Sci USA 102:12639–12643. doi:10.1073/pnas.0504254102

Jeanne JM, Thompson JV, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31:2595–2606. doi:10.1523/JNEUROSCI.3930-10.2011

Jeanne JM, Sharpee TO, Gentner TQ (2013) Associative learning enhances population coding by inverting interneuronal correlation patterns. Neuron 78:352–363. doi:10.1016/j.neuron.2013.02.023

Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a comparative neurobiology of language. Front Evol Neurosci 4:1–13. doi:10.3389/fnevo.2012.00006

Knudsen DP, Gentner TQ (2010) Mechanisms of song perception in oscine birds. Brain Lang 115:59–68

Knudsen DP, Gentner TQ (2013) Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. J Neurophysiol 0109:1690–1703. doi:10.1152/jn.0.0461.2012

Knudsen D, Thompson J, Gentner T (2010) Distributed recognition of natural songs by European starlings. Learn Mot 102:1763–1778

Marcus GF (2000) Pabiku and Ga Ti Ga: two mechanisms infants use to learn about the world. Curr Dir Psychol Sci 9:145–147. doi:10.1111/1467-8721.00080

Marcus G, Vijayan S, Brandi Rao S, Vishton PM (1999) Rule learning by seven-month-old infants. Science 283:77–80

Marcus GF, Fernandes KJ, Johnson SP (2007) Infant rule learning facilitated by speech. Psychol Sci 18:387–391. doi:10.1111/j.1467-9280.2007.01910.x

Meliza CD (2011) Effects of auditory recognition learning on the perception of vocal features in European starlings (*Sturnus vulgaris*). J Acoust Soc Am 130:3115–3123. doi:10.1121/1.3641420

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci 32:15158–15168. doi:10.1523/JNEUROSCI.0845-12.2012

Meliza CD, Chi Z, Margoliash D (2010) Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. J Neurophysiol 103:1195–1208

Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proc Nat Acad Sci USA 89:6818–6822

Mountjoy JD, Lemon RE (1995) Extended song learning in wild European starlings. Anim Behav 49:357–366. doi:10.1006/anbe.1995.0048

Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. Cognition 70:27–52

Teinonen T, Fellman V, Näätänen R et al (2009) Statistical language learning in neonates revealed by event-related brain potentials. BMC Neurosci 10:21. doi:10.1186/1471-2202-10-21

ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Phil Trans R Soc B Biol Sci 367:1984–1994. doi:10.1098/rstb.2012.0055

Thompson JV, Gentner TQ (2010) Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. J Neurophysiol 103:1785–1797

Thompson JV, Jeanne JM, Gentner TQ (2012) Local inhibition modulates learning-dependent song encoding in the songbird auditory cortex. J Neurophysiol 109:721–733. doi:10.1152/jn.0.0262.2012

Van Heijningen CAA, De Visser J, Zuidema W, ten Cate C (2009) Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proc Nat Acad Sci USA 106:20538–20543

Van Heijningen C, Chen J, Van Laatum I, van der Hulst B, ten Cate C (2012) Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim Cogn 16:165–175. doi:10.1007/s10071-012-0559-x

Wasserman E, Astley S (1994) A behavioral analysis of concepts: its application to pigeons and children. Psychol Learn Mot 31:73–132

CHAPTER FOUR

Perceptual categories enable pattern generalization in songbirds

Comins, JA and Gentner, TQ

As it appears in

Cognition

2013

Volume 128, Pages113-118

Digital Object Identifier (DOI): http://dx.doi.org/10.1016/j.cognition.2013.03.014

Cognition 128 (2013) 113–118

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Perceptual categories enable pattern generalization in songbirds

Jordan A. Comins [a], Timothy Q. Gentner [a,b,c,*]

[a] Department of Psychology, University of California, San Diego, United States
[b] Neurosciences Program, University of California, San Diego, United States
[c] Kavli Institute for Brain and Mind, La Jolla, United States

ARTICLE INFO

Article history:
Received 27 November 2012
Revised 9 March 2013
Accepted 25 March 2013

Keywords:
Pattern recognition
Artificial grammar learning
Songbird
Auditory processing
Rule learning

ABSTRACT

Since Chomsky's pioneering work on syntactic structures, comparative psychologists interested in the study of language evolution have targeted pattern complexity, using formal mathematical grammars, as the key to organizing language-relevant cognitive processes across species. This focus on formal syntactic complexity, however, often disregards the close interaction in real-world signals between the structure of a pattern and its constituent elements. Whether such features of natural auditory signals shape pattern generalization is unknown. In the present paper, we train birds to recognize differently patterned strings of natural signals (song motifs). Instead of focusing on the complexity of the overtly reinforced patterns, we ask how the perceptual groupings of pattern elements influence the generalization pattern knowledge. We find that learning and perception of training patterns is agnostic to the perceptual features of underlying elements. Surprisingly, however, these same features constrain the generalization of pattern knowledge, and thus its broader use. Our results demonstrate that the restricted focus of comparative language research on formal models of syntactic complexity is, at best, insufficient to understand pattern use.

1. Introduction

Theories of syntax are central to understanding language. As a result, many comparative psychologists interested in the study of language evolution have targeted pattern complexity, using formal mathematical grammars, as the key to organizing language-relevant cognitive processes across species (Berwick, Beckers, Okanoya, & Bolhuis, 2012; Berwick, Okanoya, Beckers, & Bolhuis, 2011). In identifying similarities and differences between human and nonhuman animals in relation to pattern recognition and production, some researchers have concluded that comparative models are inadequate because "human language sentences are potentially unbounded in length and structure, limited only by extraneous factors, such as short-term memory or lung capacity" (Berwick et al., 2011). This focus on formal syntactic complexity, however, disregards the close interaction in real-world signals between the structure of a pattern and its constituent elements as well as core biological and cognitive constraints intrinsic to temporal processing and, therefore, language. Others have argued that comparative studies are essential to the study of language precisely because they showcase how biological and cognitive mechanisms interact with dynamic real-world signals to tune pattern perception mechanisms crucial to aspects of language (Kiggins, Comins, & Gentner, 2012; Margoliash & Nusbaum, 2009). The latter perspective proposes to study language and its evolution in the context of the principles of organismal biology (Margoliash & Nusbaum, 2009), whereas the former posits these questions in the domain of mathematical formalisms specifically unburdened by such restrictions (Berwick et al., 2011, 2012).

---

* Corresponding author. Address: UCSD, Department of Psychology, 9500 Gilman Drive MC0109, La Jolla, CA 92092-0109, United States. Tel.: +1 858 822 6763.

E-mail address: tgentner@ucsd.edu (T.Q. Gentner).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.03.014

What role does syntactic pattern complexity play in language origin? Is the capacity for humans to perceive *mildly context-sensitive language* constructions, such as "John Mary Peter Jane lets help teach swim" (see Berwick et al., 2011), or even *context-free language* constructions, like "the starling the cats want was tired" (again, Berwick et al., 2011) what marks the divide between a nervous system that can support language from one that cannot? Or are these grades of temporal pattern complexity *extraneous* to the minimally sufficient requirements for language? Regardless of formal computational models that might be instantiated to perceive a given pattern, the broader utility of any pattern is constrained by its generalizability. That is, can knowledge of a pattern gained in one context be employed to recognize the same pattern in other contexts.

To explore this idea, we examined how perceptual constraints shape pattern generalization in European starlings, *Sturnus vulgaris*, a species of songbird. Starlings are expert auditory sequence learners both in the laboratory and during natural song development. Acoustically, starling song appears as a patterned sequence of distinct motifs, where each 200–1000 ms long motif comprises multiple shorter notes presented in a stereotyped pattern. Although motif repertoires between starlings are largely unique, all motifs can be grouped into four general classes by their broad acoustic characteristics (Eens, Pinxten, & Verheyen, 1988). Perceptually, these classes form open-ended natural categories for starlings (Braaten, 2000), and individual motifs within categories can be readily differentiated (Meliza, 2011). Starlings attend to the temporal patterning of notes within motifs (Gentner, 2008) and to the temporal patterning of motifs in songs (Gentner & Hulse, 1998). Further, starlings accurately classify and generalize artificial motif patterns that take the forms AABB and ABAB (Gentner, Fenn, Margoliash, & Nusbaum, 2006), where sets A and B represent two of four natural motif categories (rattles and warbles) in starling song.

2. Methods

2.1. Subjects

Eight ($N = 8$) European starlings (*S. vulgaris*) wild-caught in southern California served as subjects. We controlled for neither age nor sex, though all subjects had full adult plumage when acquired and, thus, were at least one year old. From the time of capture until use in this study, subjects were housed in a large mixed sex, conspecific aviary with *ad libitum* access to food and water. The photoperiod in the aviary and the testing chambers followed the seasonal variation in local sunrise and sunset times.

2.2. Stimuli

Patterned stimuli were made using motifs from the song recordings of one adult male starling captured near Baltimore, Maryland (for details, see Gentner & Hulse, 1998). Motifs can be broadly classified into four classes based on their spectrotemporal features: whistles, warbles, rattles and high frequency events (Adret-Hausberger & Jenkins, 1988). We extracted eight unique (non-repetitious) warble and rattle motifs from the songs of one male. Both warbles and rattles possess a variable acoustic structure, with rattles further characterized by a broadband click train. The rattle and warble motif categories follow ethologically-relevant and psychologically-salient boundaries (Braaten, 2000; Eens et al., 1988). Using these two classes of motifs, we built two patterns of motif sequences of the form XXYY or XYXY. While the pattern constituents and groupings (i.e. motifs and motifs classes) are ethologically-relevant, the overall pattern forms XXYY/XYXY are not.

To construct XXYY and XYXY patterned stimuli, we divided the motifs into two sets, labeled A and B. For half of the subjects ($n = 4$), motif membership in set A and B preserved natural acoustic (perceptual) category boundaries (i.e., 8 warbles for set A; 8 rattles for set B; see Fig. 1). For the other half of subjects ($n = 4$) motif membership in sets A and B conflicted with the natural perceptual boundaries (i.e., 4 warbles, 4 rattles forset A; 4 warbles, 4 rattles forset B; see Fig. 2C and D).

Subjects were trained using 32 (out of a possible 16,384) XXYY and XYXY patterned stimuli (8 of the explicit form AABB, 8 BBAA, 8 ABAB and 8 BABA). Within each subset of 8 explicit training patterns (e.g. AABB), each motif appeared exactly once in each possible position, and no motif appeared twice in the same sequence. No differences were observed between the mean duration of warbles and rattles (independent samples *t*-test: $t = 0.98$; df = 14; $p = 0.34$), or in the durations of XXYY and XYXY patterns in the group where the natural perceptual boundaries A/B motif were preserved (independent samples *t*-test: $t = 0.0$; df = 30; $p = 1.00$), or pseudo-randomized (independent samples *t*-test: $t = 0.0$; df = 30; $p = 1.00$). Thus, stimulus length cannot be used to recognize patterns.

2.3. Apparatus

Fig. 2A illustrates the operant apparatus where starlings learned to classify the training stimuli. Each subject was held in a small weld-wire cage containing an operant panel. On the panel, a centrally located response port was a PVC-housed opening. Inside of this opening was an IR receiver and transmitter used to detect when the bird broke the plane of the response port with its beak. This 'poke-hole' design allows starlings to probe the apparatus with their beaks, in a manner akin to their natural appetitive foraging behavior. A remotely controlled hopper, positioned behind the panel, moved the food within and beyond the subject's reach beneath the opening. Acoustic stimuli were delivered through a small full-range audio speaker mounted behind the panel and out of the subject's view. The sound pressure level inside all chambers was calibrated to the same standard broadband noise signal. Full details for all the mechanical components of the apparatus, audio interface, digital I/O control hardware, and custom software are available upon request.

2.4. Shaping

Upon initially entering the operant chamber, we provided each subject with unrestricted access to the food

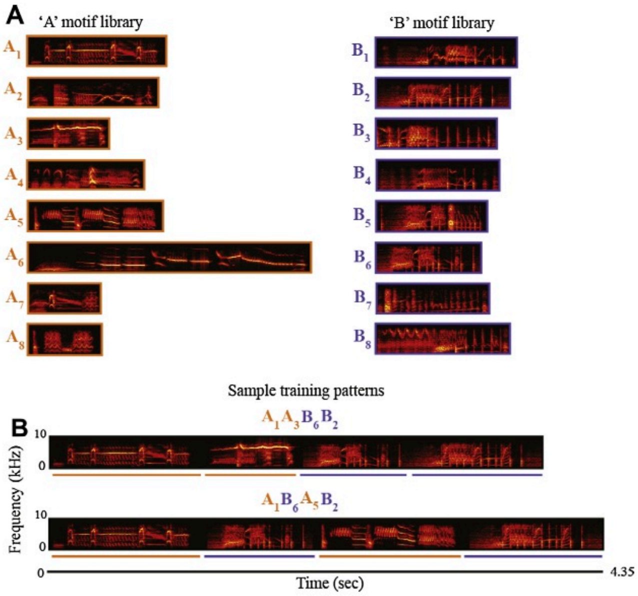

Fig. 1. (A) The 8 warbles and 8 rattles used to construct XXYY and XYXY pattern stimuli. While both warbles (left – orange) and rattles (right – dark blue) are heterogeneous in their overall spectrotemporal features, rattles additionally possess broad-band click trains not found in warbles. (B) Two out a total of 32 training patterns built from these motif libraries are shown below. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

hopper. Following acclimation to eating from this device, the hopper was lowered beyond the subject's reach. Next subjects were placed on an autoshaping routine (Brown & Jenkins, 1968). Several times per minute, we presented the subject with a blinking LED in the center response port followed by 2-s access to the food hopper. This process recurred until the subject acquired a key-peck response. From then, the lowered food hopper would only be engaged if the subject pecked the blinking LED in the center port. Subjects repeated this behavior for 100 trials. After the completion of this phase, the center LED ceased blinking, requiring subjects to peck at the darkened center port to raise the food hopper. Following 100 such trials, pecking the center port initiated the playback of an acoustic stimulus where subjects earned food rewards in accordance with standard go/no-go training procedures described below.

2.5. Classification training

Fig. 2B demonstrates the operant procedure. We trained subjects to classify sets of starling song stimuli abiding by two different patterning rules. To initiate a trial, subjects pecked a darkened center port which triggered the playback of a training stimulus. Once stimulus presentation concluded, the animal could either peck the darkened center port once more (termed a 'go' response) or not (termed a 'no-go' response). A peck in response to one set of stimuli (S+ training set) was positively reinforced with 2-s access to the food hopper. A peck in response to the other set of training stimuli (S-training set) was punished by extinguishing the operant box light at least 10 s, thereby prohibiting food hopper access. An absence of a response to either the S+ or S− stimuli yielded no operant consequence. Correction trials were available in the experiment during only training phases. For these trials, a key-peck response within the response window following an S− stimulus not only extinguished the house lights as usual, but the program would also continue to present the same S− stimulus until the subject abstained from giving a key-peck response. Correction trials were discontinued for this experiment once the animal achieved our criterion for successful classification performance during the initial training phase (criterion: 3 successive $d' \geqslant 1.0$). The stimulus exemplar presented on any given non-correction trial was sampled randomly with replacement from the pool of all stimuli the animal was learning to classify. To evaluate classification accuracy, we labeled a response to an S+ stimulus and the withholding of a response to an S− stimulus as correct. Conversely, withholding a response to an S+ stimulus and responding to an S- stimulus were labeled incorrect. Subjects could freely peck at the center response port throughout stimulus presentation, but only the first response within a 2-s response window beginning at stimulus offset triggered reinforcement or punishment. Subjects were on a closed economy during training, with daily ses-

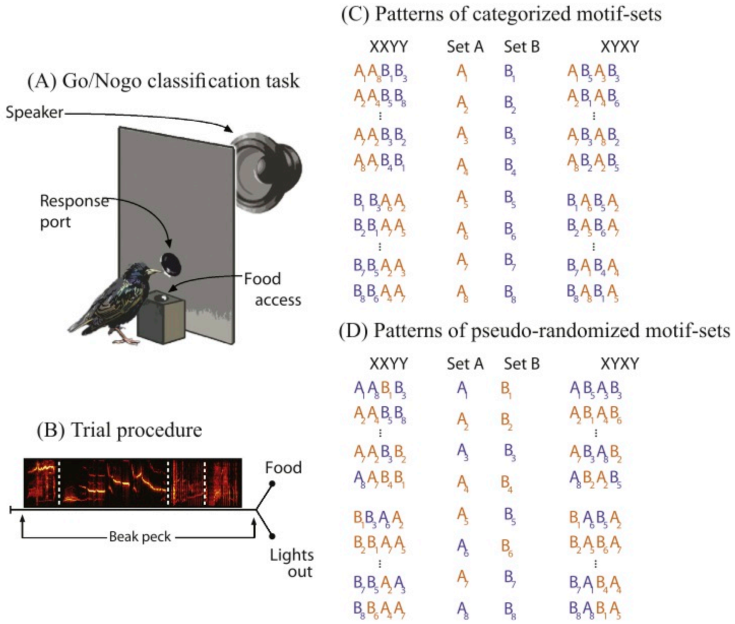

Fig. 2. (A) Schematic of operant training apparatus and (B) procedure. Visualization of pattern stimuli built from 2 sets of 8 motifs that either preserve natural acoustic categorical boundaries (C) or do not (D). Orange letters designate warbles, while dark blue designates rattles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sions lasting from sunrise to sunset, and each subject could run as few or as many trials as they were able. Food intake was monitored daily to ensure each subject's well-being. The inter-trial interval was a minimum of 2-s. Water was always available. All procedures were approved by the UCSD institutional animal care and use committee.

2.6. Procedure

To quantify learning during training we used a performance criterion of 3 consecutive 100-trial blocks where d' scores exceeded 1.0. Subjects remained on training stimuli after reaching this learning threshold to ensure stable behavioral performance prior to generalization testing. We assessed the subjects' ability to generalize knowledge of the XXYY and XYXY patterns presented during initial training using an abrupt and complete transfer to 500 novel motif sequences that followed the training patterns. Each of the 500 test sequences was built using the training motifs in sets A and B assigned to that subject, and was presented only once. The entire generalization test was exactly 500 trials. During transfer, we maintained all reinforcement contingencies as in training, except that no correction trials were used.

We used *d*-prime (d') to measure bias-free sensitivity of baseline and test pattern classification as follows: $d' = z(H - \varepsilon) - z(F + \varepsilon)$, where $z(H - \varepsilon)$and $z(F + \varepsilon)$ are the *z*-scores of the proportions of responses to S+ and S− stimuli adjusted by a factor of $\varepsilon = 0.001$ to set response rate bounds (0, 1), respectively. To compare generalization performance between the two groups, we normalized d' by dividing every subject's d' on each of their five 100-trial transfer blocks by their mean performance across the five 100-trial blocks preceding transfer.

3. Results

Initial pattern learning was similar for both groups. We observed no significant differences between groups in acquisition rate (number of 100-trial blocks until $d' > 1.0$ for 3-consecutive blocks; $t = -2.14$, df = 6, $p = 0.076$; Fig. 3A). More importantly, all subjects learned to classify the training patterns with accuracy significantly above chance ($p < 0.02$ all cases, *t*-test per bird for last 5 100-trial blocks of training, where chance $d' = 0$; Fig. 3B).

Surprisingly, pattern generalization differed qualitatively between groups. Subjects that learned patterns of categorically-organized motifs maintained accurate classification across the 500 novel patterns ($p < 0.05$ all cases, *t*-test per bird relative to chance; Fig. 3C). In contrast, none of the subjects that learned patterns with pseudo-randomly grouped motifs performed above chance during the transfer (*t*-test for each bird; in three cases: $p > 0.3$; one subject below chance; $p = 0.044$; Fig. 3C). Consistent with these differences, normalized generalization perfor-

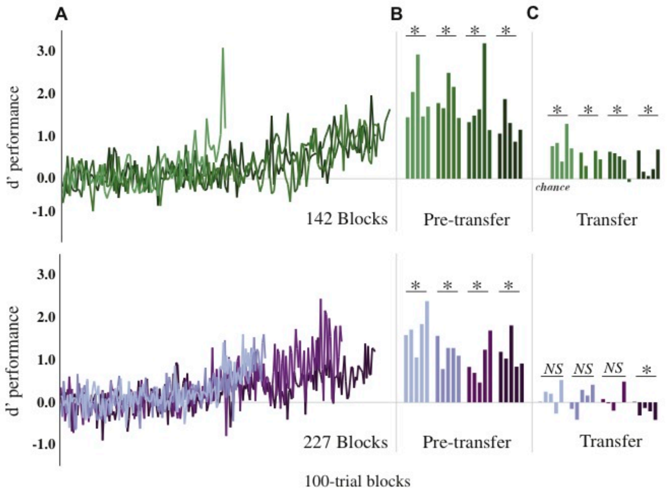

Fig. 3. (A) Training performance (d') on XXYY/XYXY patterns, where $d'=0$ represents chance (each subject is a different saturation). (B) Each bar shows performance of one subject across the final five 100-trial blocks of training; green (top) shows subjects trained with categorically organized motifs, purple (bottom) shows data for subjects trained with pseudorandomly organized motifs. Here, every subject was well-above chance performance, indicating reliable learning of XXYY and XYXY patterns. (C) Generalization performance across the five 100-trial blocks of novel XXYY and XYXY patterns differed dramatically between groups, with every subject trained on categorically organized motif patterns displaying above chance generalization and none of the subjects classifying patterns of pseudo-randomly organized motifs doing so. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mance (see methods) for subjects trained with the categorically organized motifs was significantly better than that for subjects trained with the pseudo-randomly organized motifs ($F(1,6) = 13.67$, $p = 0.01$; nested rmANOVA; Fig. 4).

4. Discussion

The results of this simple experiment are clear. Reshuffling pattern elements across the boundaries of natural acoustic categories leaves pattern learning intact, but fully blocks pattern generalization. While pattern learning and perception in a restricted context (i.e., explicit training sequences) appears agnostic to the perceptual grouping of underlying elements, such grouping constrains generalization of pattern knowledge. Thus, regularity in the perceptual structure of the pattern elements is a salient feature of pattern perception in non-humans and is essential to the broader use pattern knowledge.

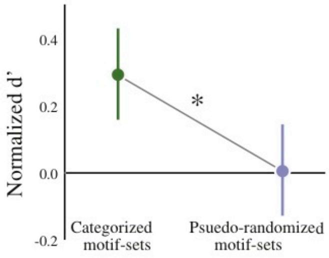

Fig. 4. Mean (±95% CI) normalized d' (methods), for subjects in the categorically-organized (green) and pseudo-randomly organized (purple) groups.$^*p < 0.05$, see text for statistics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Although this experiment was not designed to test the perceptual strategies that each group of subjects used to acquire pattern knowledge during baseline training, the results are instructive nonetheless. Given that subjects trained with the pseudo-randomized motif groups failed to generalize to novel pattern exemplars, the most parsimonious strategy for their recognition of the baseline sequences is rote memorization. In contrast, rote memorization cannot account for generalization to the novel patterns by the group working with the categorically well-defined motifs. Although we observed no statistically significant difference in the acquisition rates between subjects in the two groups, there was a trend for those learning patterns with the pseudo-category motifs to take longer, and it is likely that with larger groups sizes these differences would reach statistical significance, consistent with a different learning strategy. Understanding the learning strategy that enables successful generalization will require further experiments, but it clearly involves sensitivity to the perceptual grouping of the patterned elements.

It is unclear whether successful pattern generalization was enabled because the perceptual features of motifs aligned with already known natural categories of song ele-

ments i.e. warble and rattle, or because the categorical boundaries were more easily acquired during training due to greater within-class perceptual similarities, or both. Our choice to use of species-typical song elements was motivated by recent demonstrations that such stimuli, compared to simple tones, enhance task performance on other auditory tasks (Bregman, Patel, & Gentner, 2012). We suspect that starlings would perform similarly on this same pattern task using other auditory objects so long as they are sensitive to the perceptual categories of those stimuli – this remains, however, an open question for future research.

The present results may help in understanding mechanisms and constraints for pattern generalization in young infants. Specifically, Marcus, Fernandes, and Johnson (2007) report that infants fail to generalize patterning rules across sequences of tones, animal vocalizations or different musical timbres unless they are first exposed to those same patterning rules instantiated over speech sounds. Our results demonstrate a close interaction between pattern knowledge and pattern components, and support the idea that the abstraction of patterning information is closely constrained by the categorical structure of the constituent elements. Thus the ease with which children learn patterning rules instantiated over speech signals may reflect both familiarity and, perhaps more importantly, the perceptual structure of the speech sounds. This is consistent with several results suggesting that human infants use acoustic categories embedded in constituent elements of natural speech (e.g., words) as perceptual scaffolding for later acquiring knowledge of more abstract categories of language (Frigo & McDonald, 1998; Kelly, 1992; Monaghan, Christiansen, & Chater, 2007; Shi, Morgan, & Allopenna, 1998; Shi, Werker, & Morgan, 1999).

Our results indicate that the restricted focus of comparative language research on formal models of syntactic complexity is, at best, insufficient to understand how patterns are used. The abilities of non-human animals to perceive specific, formally defined patterns do not predict pattern generalization. Because reinterpreting patterns independent of specific perceptual events is essential for human language acquisition, we contend there is much to learn about the evolution of language faculties by understanding the benefits and constraints that learning any pattern exerts over multiple levels of abstraction.

Acknowledgements

Work supported by NSF Graduate Research Fellowship 2011122846 to JAC and NIH DC008358 to TQG. We thank Dan Knudsen, Amy Skerry and Jon Mahlandt for comments.

References

Adret-Hausberger, M., & Jenkins, P. (1988). Complex organization of the warbling song in the European starling *sturnus vulgaris*. *Behaviour, 107*(3), 138–156.

Berwick, R. C., Beckers, G. J. L., Okanoya, K., & Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Frontiers in Evolutionary Neuroscience, 4*(April), 5. http://dx.doi.org/10.3389/fnevo.2012.00005.

Berwick, R., Okanoya, K., Beckers, G., & Bolhuis, J. (2011). Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences, 15*, 113–121.

Braaten, R. F. (2000). Multiple levels of representation of song by European starlings (*Sturnus vulgaris*): Open-ended categorization of starling sing types and differential forgetting of song categories and exemplars. *Journal of Comparative Psychology, 114*(1), 61–72. http://dx.doi.org/10.1037//0735-7036.114.1.61.

Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2012). Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition, 122*(1), 51–60. http://dx.doi.org/10.1016/j.cognition.2011.08.008.

Brown, P., & Jenkins, H. (1968). Auto-shaping of the pigeon's keypeck. *Journal of the experimental analysis of behavior, 11*(I), 1–8.

Eens, M., Pinxten, R., & Verheyen, R. F. (1988). Temporal and sequential organisation of song bouts in the starling. *Ardea, 77*, 75–86.

Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language, 39*(2), 218–245. http://dx.doi.org/10.1006/jmla.1998.2569.

Gentner, T. Q. (2008). Temporal scales of auditory objects underlying birdsong vocal recognition. *The Journal of the Acoustical Society of America, 124*(2), 1350–1359.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204–1207.

Gentner, T. Q., & Hulse, S. H. (1998). Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Animal Behaviour, 56*, 579–594.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*(2), 349–364.

Kiggins, J. T., Comins, J. A., & Gentner, T. Q. (2012). Targets for a comparative neurobiology of language. *Frontiers in Evolutionary Neuroscience, 4*(April), 1–13. http://dx.doi.org/10.3389/fnevo.2012.00006.

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*(5), 387–391. http://dx.doi.org/10.1111/j.1467-9280.2007.01910.x.

Margoliash, D., & Nusbaum, H. C. (2009). Language: the perspective from organismal biology. *Trends in Cognitive Sciences, 13*(12), 505–510.

Meliza, C. D. (2011). Effects of auditory recognition learning on the perception of vocal features in European starlings (*Sturnus vulgaris*). *The Journal of the Acoustical Society of America, 130*(5), 3115–3123. http://dx.doi.org/10.1121/1.3641420.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology, 55*(4), 259–305. http://dx.doi.org/10.1016/j.cogpsych.2006.12.001.

Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language, 25*(1), 169–201.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*(2), B11–21.

CHAPTER FIVE

Pattern-induced perceptual learning in songbirds

Introduction

Language and its acquisition involve numerous perceptual and computational systems, whose coordinated activity is considered unique to humans (Hauser et al. 2002). Despite the uniqueness of language, the evolutionary roots of at least some of the capacities supporting language are thought to be shared with non-human animals (Bloomfield et al. 2011; Hockett 1960; Margoliash and Nusbaum 2009). Traditionally, the comparative approach to language has focused on sensorimotor systems involved in vocal control and production (Brainard and Doupe 2000; Eliades and Wang 2008; Schmidt and Konishi 1998). However, developmental psychologists have long shown that language modulates perceptual systems to enhance how infants process linguistic input (Aslin and Newport 2012; Kuhl 2004; Werker and Tees 1999). To understand the biology of such language-relevant perceptual changes, we need an animal model amenable to invasive neuroscience techniques (Kiggins et al. 2012).

One well-studied perceptual change, phonemic categorization, occurs early in human development. While newborn infants come into the world with the impressive ability to distinguish the vast array of phonemes comprising all human languages, the range of these perceptual abilities quickly narrows. Indeed, by 10-months, infants show adult-like categorization of the phonemes present in their linguistic environment and lose the ability to detect those that are not (Werker and Tees 1999; Werker and Lalonde 1988). This modulation of perceptual expertise in infants is experience-dependent, often exemplified by the categorical perception of differences between /r/ and /l/ sounds by English speakers and an inability to distinguish these same sounds by Japanese speakers (Miyawaki et al. 1975). Consistent with this, we know that infants are sensitive to the frequency distribution of speech sounds along phonemic continua (Maye et al. 2002; Werker et al. 2007). It remains unclear, however, precisely which aspects of early experience carry this information and lead to subsequent changes in categorical perception during natural language exposure. One possibility is that simple exposure to speech sounds at different distributional frequencies is enough to alter perception. Alternatively it may be that changes in categorical perception require the relevant speech sounds to be patterned in informative ways. Dissociating these two possible learning pathways is difficult in humans at it requires limiting natural language exposure.

Here we use songbirds to explore whether learning patterns instantiated over auditory categories covertly enhances categorical perception. During their lifetime, both male and female starlings learn to sing songs that are a hierarchically structured (Eens et al. 1988; Mountjoy and Lemon 1995; Pavlova et al. 2005). Acoustically, starling song appears as a patterned sequence of distinct motifs, where each motif comprises multiple shorter notes presented in a stereotyped pattern roughly 200 - 1000ms long. Although motif repertoires between birds are largely unique, all motifs can be grouped into four general classes by their broad acoustic characteristics (Eens et al. 1988). Perceptually, these classes form open-ended natural categories for starlings (Braaten 2000), and individual motifs within categories can be readily differentiated (Meliza 2011; Meliza et al. 2010). Motif patterning underlies both successful individual recognition (Gentner and Hulse 1998) and mate selection (Gentner and Hulse 2000). In controlled operant settings, starlings can accurately classify and generalize artificial motif patterns that take the forms AABB and ABAB (Gentner et al. 2006), where sets A and B represent two of four natural motif categories (warbles and rattles, respectively). Further, similarly to humans, their ability to generalize these patterns is constrained by the integrity of these motif categorical boundaries when assigning motifs to sets A and B (Comins and Gentner 2013; Emberson et al. 2013).

In the current experiment, we asked how prior experience with motif categories in pattern or non-pattern sequences affects later motif categorization. First, one group of subjects was trained to differentiate complex auditory patterns built from warble and rattle motifs. In brief, these patterns were of the form AABB, BBAA, ABAB and BABA and the subject's task was to perform one response (e.g., peck) following an AABB or BBAA pattern, and a different response (e.g., withhold peck) to either an ABAB or BABA pattern. For a second group of birds (sequence controls), some subset of all pattern types (again, AABB, BBAA, ABAB and BABA) required one type of response (e.g., peck) and another subset of all pattern types required a different response (e.g., no peck). Thus, these subjects were presented similarly patterned stimuli, but the patterning rule was non-informative for stimulus classification (see Table 1).

We find that starlings operantly trained to recognize patterns built from motif categories are better able to differentiate between these categories. Control subjects demonstrate that this effect is not due to familiarity with sequences of motif categories, but is specifically attributable to the recognition of an abstract pattern relation among them. Our findings highlight the close interaction in real-world signals between the structure of a pattern and its constituent elements.

Methods

Subjects

Twelve (N=12) wild-caught European starlings (*Sturnus vulgaris*) of both sex served as subjects. All subjects had full adult plumage when acquired and, thus, were at least one year old. From the time of capture until use in this study, subjects were housed in a large mixed sex, conspecific aviary with *ad libitum* access to food and water. Figure 1A illustrates the operant apparatus where starlings learned to classify the training stimuli (for full description, see (Comins and Gentner 2010)). Full details for all the mechanical components of the apparatus, audio interface, digital I/O control hardware, and custom software are available upon request.

Stimuli

To construct the patterns used in this study, we first extracted sixteen acoustically distinct warble and rattle motifs (eight motifs per class) from the recorded songs of one adult male starling. Warbles and rattles are acoustically complex, with rattles characterized by a broadband click train overlying more narrow-band spectro-temporal features (Fig. 1B). Categorical differences between warble and rattle motifs are ethologically-relevant (Eens et al. 1988) and perceptually salient (Braaten 2000; Eens et al. 1988). Using the eight motifs in each of these two classes, we constructed motif sequences that followed the patterning rules AABB, BBAA, ABAB and BABA.

To construct patterned sequences, we labeled the sets of warble and rattle motifs ‘A’ and ‘B’, where A and B denote any of the eight possible warble or rattle motifs, respectively. We then assembled individual warble and rattle motifs into 4-motif sequences of the form AABB, BBAA, ABAB, and BABA. For successful pattern classification, subjects must produce one response to AABB and BBAA stimuli (e.g., peck) and a different response to ABAB and BABA stimuli (e.g., no peck).

We trained one group of subjects (pattern-trained; N = 4) using 32 (out of a possible 16,384) patterned stimuli (8 of the explicit form AABB, 8 BBAA, 8 ABAB and 8 BABA). In each subset of 8 explicit training patterns (e.g. AABB), each motif appeared exactly once in each possible position, and no motif appeared twice in the same sequence (Table 1). A second group of birds (sequence-controls; N = 4) served as motif and sequence exposure controls. These subjects also encountered patterned stimuli in the operant apparatus. Unlike pattern-trained birds, however, stimulus response associations for this second group were not tied to the AABB and BBAA versus ABAB and BABA pattern distinction. Instead, for these animals we assigned 4 AABB, 4 BBAA, 4 ABAB and 4 BABA sequences as S+ and another 4 AABB, 4 BBAA, 4 ABAB and 4 BABA sequences as S- (Table 1).

Stimuli for the motif categorization portion of the study were the eight warble and eight rattle motifs used to construct the patterned sequence stimuli, with a single motif presented on each trial.

Procedures

Details of the operant training procedures used in the current study can be found elsewhere (Comins and Gentner 2013). In brief, starlings learned to peck a central response port to initiate the playback of a stimulus. Following the stimulus, the animal either pecked the port again (a ‘go’ response) or withheld a peck to the port (a ‘no-go’ response). Go responses to one set of stimuli (S+) were positively reinforced with 2-s access to a food hopper below the response port. Go responses to the other set of training stimuli (S-) were punished by extinguishing the light in the operant apparatus for at least 10 seconds. Food was only available following a go response to an S+ stimulus; the no-go response to either the S+ or S- stimuli yielded no operant consequence. To minimize non-specific responses biases unrelated to the training stimuli we used correction trials during initial training. There were no correction trials during generalization testing.

We divided the subjects into three groups, and eventually trained all the birds in each group to distinguish between the eight warble and eight rattle motifs. Prior to this motif classification task, we trained one group of subjects to differentiate patterned acoustic stimuli built from the motifs in each category. We trained a second group of subjects to classify similarly patterned stimuli, except that the patterning rule was non-informative for stimulus classification (Table 1). The third group of subjects remained experimentally naïve until the motif category classification task.

We tested pattern generalization for the pattern learning and sequence control groups by presenting subjects with 500 novel motif sequences built from the same motifs used in training. All procedures were approved by the UCSD institutional animal care and use committee.

Analysis

For group analyses, we compared percent correct scores across groups (between-subjects factor) and block number (within-subjects factor) using analysis of variance. We analyzed individual subject data using binomial tests comparing raw number of correct responses in a given number of trials where chance likelihood of a correct response was 0.5. As our primary focus concerned the learnability of motif categories (warbles and rattles), data were analyzed until any group reached our criterion threshold of learning. In this case, every subject needed to exceed chance performance on the binomial analysis.

This occurred for our pattern-trained birds in the sixth 100-trial block of testing – therefore data across groups was only compared for these 600 trials.

Results

We hypothesized that experience with behaviorally-relevant patterned sequences would improve perceptual expertise for lower-level acoustic categorization. If true, then those subjects pre-trained with the patterned motif sequences, should show advantages in motif categorization over naive birds and over those for whom the patterned motif sequences were familiar but not behaviorally relevant (sequence controls). We first describe the effects of our different pre-training regimens on each group, and then discuss the results of the motif categorization training common to all groups.

Pre-training performance:

The pattern pre-trained subjects learned to classify AABB and BBAA from ABAB and BABA patterns. By 10,000 trials, mean performance (percentage of correct responses) for the group was well-above values expected by chance (single sample t-test; $t = 11.09$; $p = 0.008$ against chance value of 50%). Moreover, mean classification accuracy stayed significantly above chance (single sample t-test; $t = 3.9$; $p = 0.0298$ against chance value of 50%; Fig. 1C) even when tested with 500 novel motif patterns built from the same warbles and rattles used for training. This pattern generalization effect is observed at the individual level for 3 out of 4 subjects (Bird 681: $p < 0.0001$, Bird 716: $p < 0.002$, Bird 827: $p < 0.0001$, Bird 828: $p = 0.227$; binomial tests where chance is 0.5). This corroborates previous results suggesting starlings recognize auditory patterns of motif categories based on their underlying temporal structures(Comins and Gentner 2013; Gentner et al. 2006).

For subjects in the second pre-training group, who served as controls for sequence exposure, performance never exceeded chance thresholds. By design, the number of operant trials executed by sequence control subjects exceeded that of their pattern-trained counterparts (sequence controls: 159.25 + 21.47; pattern-trained: 119.25 + 27.59 blocks of 100-trials; matched pairs t-test t = 3.22; *p* = 0.0487). To properly ensure matched experimental history, sequence control birds were also given 500 dummy pattern generalization trials, where they encountered the same generalization test stimuli as pattern-trained birds. As with their training stimuli, however, there was no fixed relationship between pattern and reward and therefore performance remained at chance levels (AABB, BBAA, ABAB and BABA stimuli randomly assigned as S+ and S-).

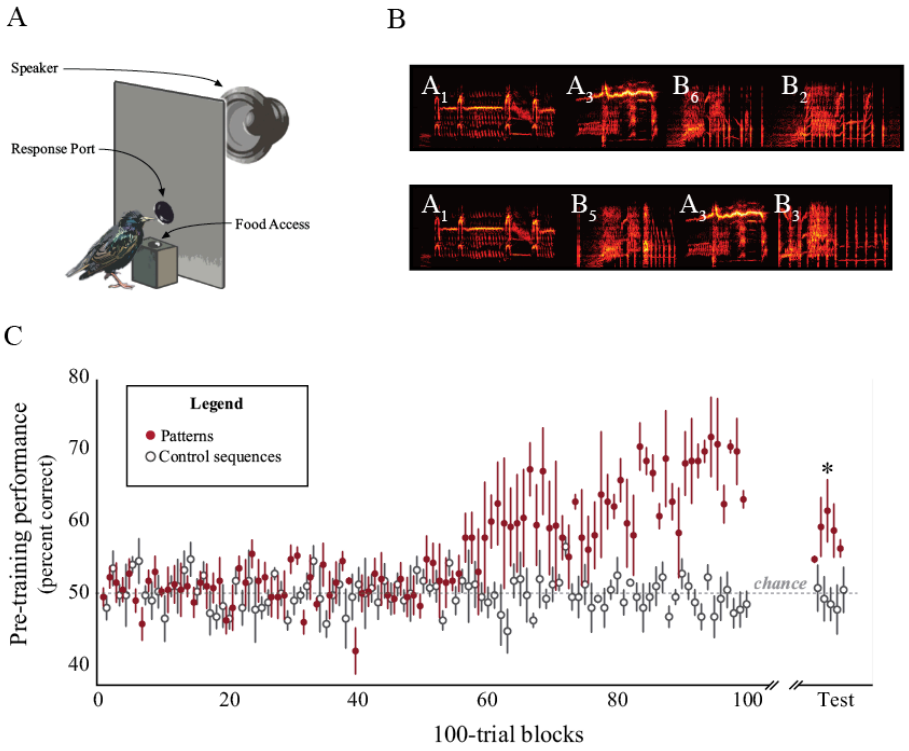

Figure 5.1. (A) Schematic of the operant apparatus. (B) Spectrograms typifying stimuli from the AABB and ABAB pattern classes (As denote warbles, Bs denote rattles). In total, eight warbles and eight rattles were used to generate all types of patterned sequences in this experiment: AABB and BBAA as well as ABAB and BABA. (C) Mean (± SEM) performance (% correct) over the course of motif pattern pre-training and subsequent generalization (labeled “test”) to novel exemplars of the patterns.

Motif-category learning

Next, we transferred the pattern-trained and sequence-control birds to a go-nogo task to assess their discrimination of the warbles and rattles used in pre-training stimuli. In addition, a set of naïve birds was also placed on the same task. Performance on this task was measured for 600 trials (see Methods – Analysis). Pattern pre-trained birds learned to classify the song motifs significantly faster than both the sequence and naïve control groups. We found a main effect of training regime on category discrimination (nested rmANOVA $F_{(2,9)} = 9.96$; $p = 0.0052$; post-hoc tests of pattern-trained versus sequence-controls $p = 0.0121$ and pattern-trained versus naïve controls $p = 0.0019$) and also a significant interaction across the 6 100-trial blocks (nested rmANOVA $F_{(10,45)} = 3.551$; $p = 0.0016$). Post-hoc analyses comparing group performance for each of the 6 blocks of 100-trials reveal significant differences emerging by blocks 5 and 6 (Bonferroni-corrected $\alpha = 0.0083$; $p = 0.002$ and $p < 0.0001$, respectively; Fig. 2). On the individual subject level, by the fifth block of the category task, 2 of 4 pattern-trained birds were significantly better than expected by chance (binomial test where chance is 0.5, in two cases $p < 0.05$). By the sixth block, all of the pattern-trained birds were significantly above chance performance (binomial test where chance is 0.5, in all four cases $p < 0.05$). Across all 6 blocks, each pattern-trained bird performed better than expected from chance, (binomial test where chance is 0.5, in all four cases $p < 0.05$). In contrast, none of the naïve birds were above chance and only a single control-sequence bird performed above chance across the six blocks. We conclude that auditory pattern learning, but not exposure to or rote memorization of acoustic sequences, enhances the perceptual mechanisms that underlie acoustic categorization.

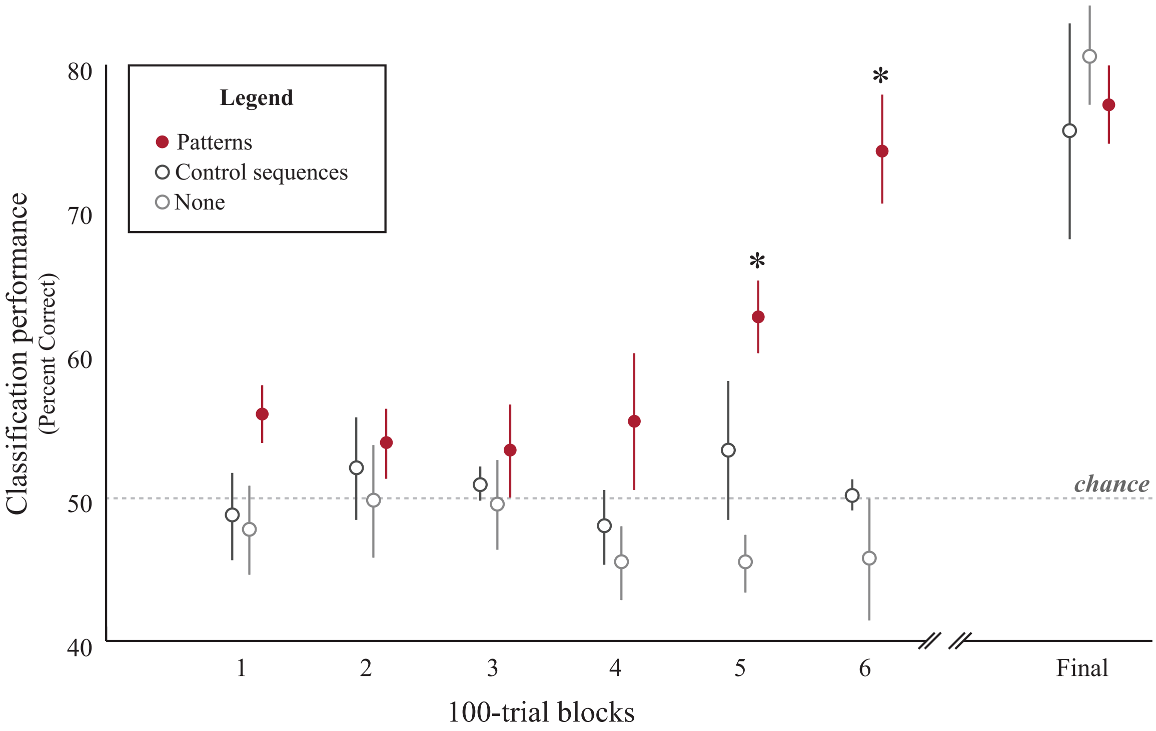

Figure 5.2. Performance of pattern-trained, sequence-controls and naïve birds of motif categorization task. Across 600 trials, performance of previously trained pattern birds far exceeds that of either the control-sequence or experimentally naïve subjects, with each individual pattern bird performing better than chance on the sixth block of training. Eventually all groups converge to comparable performance levels.

To ensure that subjects from these two control groups could indeed learn to categorize warbles and rattles, all subjects remained classifying these warbles and rattles for a varying number of blocks *after* these initial 6 100-trial blocks. If we look at each subject's *final* 100-trial block, motif categorization performance is roughly equivalent (ANOVA $p = 0.7608$; Fig. 2). This shows that for all subjects these categories are *learnable* (though the number of training blocks represented here varied considerably across groups: pattern-trained – 12.25 ± 1.5 blocks; sequence controls – 23.75 ± 3.8 blocks; naïve controls – 27.75 ± 6.2 blocks; μ ± SE).

Table 5.1. Configuration of warbles (As) and rattles (B) for pattern group stimuli (top) and the sequence-controls (bottom).

Stimulus class	One	Two
XXYY/XYXY patterns	$A_1A_8B_1B_3$	$A_1B_5A_3B_3$
	$A_2A_4B_5B_8$	$A_2B_1A_4B_6$
	$A_3A_6B_7B_6$	$A_3B_7A_6B_8$
	$A_4A_5B_8B_5$	$A_4B_4A_5B_1$
	$A_5A_1B_6B_4$	$A_5B_6A_1B_4$
	$A_6A_3B_2B_7$	$A_6B_8A_7B_7$
	$A_7A_2B_3B_2$	$A_7B_3A_8B_2$
	$A_8A_7B_4B_1$	$A_8B_2A_2B_5$
	$B_1B_3A_6A_2$	$B_1A_6B_5A_2$
	$B_2B_1A_7A_5$	$B_2A_5B_6A_7$
	$B_3B_4A_1A_4$	$B_3A_7B_8A_3$
	$B_4B_7A_3A_8$	$B_4A_3B_3A_8$
	$B_5B_2A_5A_6$	$B_5A_2B_2A_6$
	$B_6B_8A_8A_1$	$B_6A_4B_7A_1$
	$B_7B_5A_2A_3$	$B_7A_1B_4A_4$
	$B_8B_6A_4A_7$	$B_8A_8B_1A_5$
XXYY/XYXY control sequences	$A_1A_5B_4B_8$	$A_5A_1B_8B_4$
	$A_2A_6B_3B_7$	$A_6A_2B_7B_3$
	$A_3A_7B_2B_6$	$A_7A_3B_6B_2$
	$A_4A_8B_1B_5$	$A_8A_4B_5B_1$
	$A_5B_6A_2B_3$	$A_1B_2A_6B_7$
	$A_6B_7A_3B_4$	$A_2B_3A_7B_8$
	$A_7B_8A_4B_1$	$A_3B_4A_8B_5$
	$A_8B_5A_1B_2$	$A_4B_1A_5B_6$
	$B_1B_4A_5A_8$	$B_5B_8A_1A_4$
	$B_2B_3A_6A_7$	$B_6B_7A_2A_3$
	$B_3B_2A_7A_6$	$B_7B_6A_3A_2$
	$B_4B_1A_8A_5$	$B_8B_5A_4A_1$
	$B_5A_2B_6A_3$	$B_1A_6B_2A_6$
	$B_6A_1B_5A_4$	$B_2A_5B_1A_5$
	$B_7A_4B_8A_1$	$B_3A_8B_4A_8$
	$B_8A_3B_7A_2$	$B_4A_7B_3A_7$

While these results suggest that pattern learning can improve lower-level categorization, the failure of the sequence control group to exceed chance performance on their pre-training task raises an important concern. Namely, it is possible that the enhanced motif categorization observed in pattern pre-trained starlings arises from their having learned something about the mechanics of the operant procedure rather than motif categories. If this is correct, then the individual variation in performance on the pattern pre-training and generalization should correlate with later performance on motif categorization. While pattern pre-training performance across the final 6 blocks of training strongly correlated with pattern generalization performance (Pearson's correlation, r = 0.9674, $p = 0.0326$) neither pattern pre-training nor pattern generalization correlated with motif categorization (Pearson's correlation of pattern training and motif categorization, r = -0.0356, $p = 0.9644$; Pearson's correlation of pattern generalization and motif categorization, r = -0.1432, $p = 0.8568$). Thus, enhanced motif categorization in pattern pre-trained birds is not driven by expertise with operant procedures.

To confirm that the sequence control group did indeed learn how to use the operant apparatus equally well compared with the pattern pre-trained birds, we compared perseverance rates over time between these two groups of subjects. In the context of the current study, perseverance is the length (in number of trials) of a given correction trial sequence. For instance, if a subject incorrectly performs a 'go' response to a 'no-go' stimulus then the next trial, a correction trial, will present the same 'no-go' stimulus to the subject. Until the subject correctly withholds a pecking response, each subsequent trial will continue to present the same 'no-go' stimulus. Over time, animals develop strategies in response to correction trials that indicate learning about the mechanics of the operant apparatus, such that perseverance, or the number of times a given correction trial has to be repeated, decreases. In our study, we analyzed the perseverance for all subjects during the pre-training phase. All subjects, regardless of experimental group, show significant decreases in perseverance over time (Pearson's correlation: in all 8 cases, $p < 0.05$). Most importantly, the rates at which perseverance decreased (i.e., magnitude of the slope from these correlations) across our pattern pre-trained and sequence control groups were not significantly different (unmatched t-test $t = -1.86$; $p = 0.152$). This analysis suggests that despite failing to learn to successfully classify S+ and S- stimuli in the pre-training phase, our experiment control sequence subjects did indeed learn about the mechanics of the operant procedure they were using.

Discussion

We have demonstrated that learning to classify patterned arrangements of species-specific categories enhances perceptual abilities to differentiate between those categories in the starling. Crucially, this result is not driven by experience with sequences of these categories, as we provided a control group with even more overall operant experience and familiarity with such sequences. Thus, pattern recognition in songbirds tunes perceptual systems towards lower-level category distinctions when they are consistent with behavioral goals.

Our results bear a striking similarity to certain perceptual changes that occur during the first year of human development. Indeed, human infants rely on speech input to learn their native language's sound-based (i.e., phonological) organization. Initially 2-months-old infants are equally sensitive to differences between speech sounds present and absent in their language environment. By 10-months-old, however, they develop adult-like categorical expertise (Maye et al. 2002) thereby suggesting that experience with natural speech input shapes the representations of relevant phonemic dimensions.

Until now, the main hypothesis behind phonemic boundary acquisition in infancy derives from monitoring the frequency distributions along pertinent phonemic distinctions (Maye et al. 2002; Werker et al. 2007). One important proposal from our data, however, is that any reliable patterning of sounds induces top-down modulation of the relevant perceptual space. In other words, the temporal patterns occurring in everyday speech might facilitate the acquisition of the relevant perceptual (i.e., phonemic) structures of a given language. Such ideas are consistent with data showing top-down influences on phoneme perception in human adults, where ambiguous speech sounds are resolved perceptually based on the subject's knowledge of a word (Samuel 2001). For example, classic psychological experiments (Ganong 1980) show that if a sound located in the middle of the /d/ - /t/ phonetic continuum precedes "_ask", listeners will report hearing the word "task" as opposed to the non-word "dask." Contrarily, if the same stimulus precedes "_ash" subjects report hearing the word "dash" over the non-word "tash." Still, whether similar top-down influences over perception play a role in phonemic category learning during development remains an important question.

Our findings also share important similarities to previous reports in humans suggesting that higher-level or conceptual learning augments lower-level perceptual expertise of complex objects in the visual system(Goldstone 1998). Most traditional models of complex object recognition characterize information flow in a feed-forward hierarchical fashion, where increasingly complex perceptual representations unilaterally support high-level stimulus representations (Riesenhuber and Poggio 1999). More recent experimentation, however, strongly suggests an exchange of information across multiple levels object representation (Ahissar and Hochstein 2004). In one illustrative example of this effect, human subjects were first trained to categorize squares as being from either category A or category B. These squares varied along only two dimensions: size and brightness. Crucially, only one of these dimensions was relevant for evaluating category membership. For instance, subjects might be trained that the relevant dimension for determining category membership was size (with brightness then being irrelevant for category assignment). Afterwards, subjects were transferred to a psychophysical task and asked to assess whether two squares were the same or different. Discrimination ability was vastly improved when the squares in this task varied along the relevant category dimension (in this example, size) as opposed to the irrelevant category. Further, same/different performance was even better along the relevant dimension at the categorical boundary (Goldstone 1994). These findings strongly suggest that learned information at higher-levels of stimulus representation back-propagate to influence relevant lower-level perceptual tuning.

The starlings' behavior here is to our knowledge the first demonstration of a similar role of high-level stimulus learning shaping relevant lower-level perceptual representations in an organism amenable to invasive neuroscience techniques. Starlings already serve as an important model species to investigate how experience alters the response properties of sensory neurons throughout the avian forebrain (Gentner and Margoliash 2003; Jeanne et al. 2013; Jeanne et al. 2011; Knudsen and Gentner 2013; Meliza et al. 2010; Meliza and Margoliash 2012; Thompson and Gentner 2010). Our results here showcase potential parallels with human work of high-level object or pattern knowledge dynamically interacting with knowledge of lower-level perceptual expertise (Emberson et al. 2013; Goldstone 1994). Thus our findings suggest that this species might serve as a suitable nonhuman model system to examine the biological instantiation of certain perceptual changes underlying language at a cellular level.

Acknowledgements

This chapter, in full, is currently material under review. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

References

Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. Trends in cognitive sciences 8:457–64. doi: 10.1016/j.tics.2004.08.011

Aslin RN, Newport EL (2012) Statistical Learning: From Acquiring Specific Items to Forming General Rules. Current Directions in Psychological Science 21:170–176. doi: 10.1177/0963721412436806

Bloomfield TC, Gentner TQ, Margoliash D (2011) What birds have to say about language. Nature neuroscience 14:947–8. doi: 10.1038/nn.2884

Braaten RF (2000) Multiple levels of representation of song by European starlings (Sturnus vulgaris): Open-ended categorization of starling sing types and differential forgetting of song categories and exemplars. Journal of Comparative Psychology 114:61–72. doi: 10.1037//0735-7036.114.1.61

Brainard M, Doupe A (2000) Auditory feedback in learning and maintenance of vocal behaviour. Nature Reviews Neuroscience 1:1–10.

Comins JA, Gentner TQ (2010) Working memory for patterned sequences of auditory objects in a songbird. Cognition 117:38–53.

Comins JA, Gentner TQ (2013) Perceptual categories enable pattern generalization in songbirds. Cognition 128:113–118. doi: 10.1016/j.cognition.2013.03.014

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86.

Eliades SJ, Wang X (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453:1102–6. doi: 10.1038/nature06910

Emberson LL, Liu R, Zevin JD (2013) Is statistical learning constrained by lower level perceptual organization? Cognition 128:82–102. doi: 10.1016/j.cognition.2012.12.006

Ganong W (1980) Phonetic categorization in auditory word perception. Journal of Experimental Psychology: Human Perception and Performance 6:110–125.

Gentner T, Hulse S (1998) Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Animal Behaviour 56:579–594.

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440:1204–1207.

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Animal Behaviour 59:443–458.

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674.

Goldstone R (1994) Influences of categorization on perceptual discrimination. Journal of Experimental Psychology- … 123:178–200.

Goldstone RL (1998) Perceptual learning. Annual review of psychology 49:585–612. doi: 10.1146/annurev.psych.49.1.585

Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–79. doi: 10.1126/science.298.5598.1569

Hockett CF (1960) The origin of speech. Scientific American 203:88–111. doi: 10.1126/science.303.5662.1316

Jeanne JM, Sharpee TO, Gentner TQ (2013) Associative Learning Enhances Population Coding by Inverting Interneuronal Correlation Patterns. Neuron 78:352–363. doi: 10.1016/j.neuron.2013.02.023

Jeanne JM, Thompson J V, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. The Journal of Neuroscience 31:2595–606. doi: 10.1523/JNEUROSCI.3930-10.2011

Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a Comparative Neurobiology of Language. Frontiers in Evolutionary Neuroscience 4:1–13. doi: 10.3389/fnevo.2012.00006

Knudsen DP, Gentner TQ (2013) Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. Journal of Neurophysiology 0109:1690–1703. doi: 10.1152/jn.00461.2012

Kuhl PK (2004) Early language acquisition: cracking the speech code. Nature reviews Neuroscience 5:831–43. doi: 10.1038/nrn1533

Margoliash D, Nusbaum HC (2009) Language: the perspective from organismal biology. Trends In Cognitive Sciences 13:505–510.

Maye J, Werker JF, Gerken L (2002) Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82:B101–11.

Meliza CD (2011) Effects of auditory recognition learning on the perception of vocal features in European starlings (Sturnus vulgaris). The Journal of the Acoustical Society of America 130:3115–23. doi: 10.1121/1.3641420

Meliza CD, Chi Z, Margoliash D (2010) Representations of Conspecific Song by Starling Secondary Forebrain Auditory Neurons: Toward a Hierarchical Framework. Journal Of Neurophysiology 103:1195–1208.

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. The Journal of Neuroscience 32:15158–68. doi: 10.1523/JNEUROSCI.0845-12.2012

Miyawaki K, Jenkins JJ, Strange W. (1975) An effect of linguistic experience- the discrimination of [r] and [l] by native speakers of Japanese and English. Attention, Perception, & Psychophysics 18:331–340.

Mountjoy JD, Lemon RE (1995) Extended song learning in wild European starlings. Animal Behaviour 49:357–366. doi: 10.1006/anbe.1995.0048

Pavlova D, Pinxten R, Eens M (2005) Female Song in European Starlings : Sex Differences , Complexity , and Composition. The Condor 107:559–569.

Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. Nature neuroscience 2:1019–25. doi: 10.1038/14819

Samuel a. G (2001) Knowing a Word Affects the Fundamental Perception of The Sounds Within it. Psychological Science 12:348–351. doi: 10.1111/1467-9280.00364

Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. Nature neuroscience 1:513–8. doi: 10.1038/2232

Thompson J V, Gentner TQ (2010) Song Recognition Learning and Stimulus-Specific Weakening of Neural Responses in the Avian Auditory Forebrain. Journal Of Neurophysiology 103:1785–1797.

Werker J, Tees R (1999) Influences on infant speech processing: Toward a new synthesis. Annual review of psychology 50:509–35. doi: 10.1146/annurev.psych.50.1.509

Werker JF, Lalonde CE (1988) Cross-language speech perception: Initial capabilities and developmental change. Developmental Psychology 24:672–683. doi: 10.1037/0012-1649.24.5.672

Werker JF, Pons F, Dietrich C. (2007) Infant-directed speech supports phonetic category learning in English and Japanese. Cognition 103:147–62. doi: 10.1016/j.cognition.2006.03.006

CHAPTER SIX

An avian model of rule-learning:

The role of perceptual and decision-making processes

Introduction

The ability to perceive, organize, and use temporally patterned information is integral to adaptive behavior. However, we commonly take the capacity to extract patterning rules from the environment as a hallmark of higher cognitive function in humans. Indeed, the acquisition of pattern expertise and the recognition of familiar patterns underlie sophisticated behaviors such as analogical reasoning and music (Gentner and Namy 2006; Patel et al. 1998). Central to the power of rule functioning in humans is the ability to learn and generalize patterns. In the auditory domain, this ability contributes to the parsing of richly patterned sound sequences. For instance, 8-month-old infants rapidly and spontaneously learn word boundaries embedded within continuous speech by tracking the statistical relationships between sound patterns (Saffran et al. 1996). In addition, humans are capable of generalizing such pattern knowledge to novel stimuli. Infants as young as 7-months generalize phoneme-sequence patterns to novel strings of phonemes, an ability arguably necessary for acquiring knowledge of language (Marcus et al. 1999). These pattern generalization capabilities remain intact throughout adulthood (Endress et al. 2005; Shanks and Darby 1998). Thus, in humans, pattern knowledge is dissociable from the elements comprising learned patterns (Gomez and Gerken 1999; Marcus et al. 2007; Marcus 2000).

This paper explores the capacity for rule learning in the auditory domain using natural communication sounds in a species of songbird, European starlings (*Sturnus vulgaris*). Previous comparative research has investigated temporal pattern perception in the auditory modality for tones (Hulse et al. 1984; Wright 1998; Wright et al. 1990), and human speech sounds (Hauser et al. 2001; Ohms et al. 2010; Ramus et al. 2000). Still, we know very little about the evolution of pattern recognition capacities, and almost nothing of the biological mechanisms that support pattern recognition and generalization in any species. Improving our understanding of how nonhuman animals amenable to invasive neuroscience techniques acquire and transfer temporal pattern knowledge will assuredly help to advance our knowledge of the biological substrates of temporal pattern learning and generalization (Kiggins et al. 2012; Knudsen and Gentner 2010).

Pattern learning in songbirds

Starlings excel at learning acoustic patterns in both the wild and in controlled-laboratory settings. Acoustically, starling song appears as a patterned sequence of distinct motifs, where each motif comprises multiple shorter notes presented in a stereotyped pattern roughly 200 - 1000ms long (Eens et al. 1988; Gentner 2008). Starlings perennially learn new motifs, which they fold into their already well-patterned songs (Mountjoy and Lemon 1995). In addition, starlings learn the patterning of motifs in other individuals' songs to recognize them. Thus, if starlings hear a familiar conspecifics song with the patterning of motifs artificially disrupted this impairs their ability to accurately identify that individual (Gentner and Hulse 1998). Thus, motif patterning is thought to underlie both successful individual recognition (Gentner and Hulse 1998) and mate selection (Gentner and Hulse 2000a) in the species. Recently, it has been shown that starlings accurately classify artificial motif patterns that take the forms AABB and ABAB (Gentner et al. 2006), where A and B represent sets of ethologically-relevant motif categories (Eens et al. 1988) of starling song. Once starlings acquire knowledge of these patterns through training, pattern classification generalizes to novel configurations of these motifs (Gentner et al. 2006). More recent investigations further shows that, as in humans (Emberson et al. 2013), successful generalize to novel pattern configurations is constrained by the integrity of these motif categorical boundaries when assigning motifs to sets A and B (Comins and Gentner 2013). In addition, as long as motif categorical boundaries are preserved, pattern learning and generalization can occur over highly rich and variable sources of information, including trial-unique instantiations of these patterning forms (Comins and Gentner 2014).

The current set of experiments were designed to accomplish two main objectives: (1) to assess whether training experience can loosen perceptual constraints on pattern generalization such that pattern knowledge may generalize to patterns of wholly novel motifs; and (2) to examine the decision processes involved in pattern generalization. Such understanding informs how general cognitive mechanisms support the generalization of temporally organized, complex auditory patterns.

Methods

Subjects

A total of seven adult European starlings served as subjects in this study (4 in experiment 1, 3 in experiment 2). Subjects were wild caught in southern California in May 2006. All had full adult plumage at the time of capture, and thus were at least one year old. From the time of capture until their use in this study, all subjects were housed in large, mixed sex, conspecific aviaries with ad libitum access to food and water. The photoperiod in the aviary and the testing chambers followed the seasonal variation in local sunrise and sunset times. No significant sex differences have been observed in previous studies of individual vocal recognition (Gentner and Hulse 1998; Gentner 2008; Gentner and Hulse 2000b) and the sex of subjects in this study was not controlled.

Apparatus

Starlings learned to classify the training stimuli using a custom-built operant apparatus, housed in a custom sound attenuation chamber (Acoustic Systems) with internal dimensions 61 x 81 x 56 cm. Inside the chamber, a weld-wire cage (41 x 41 x 35 cm) held the subject and permitted access to a 30 x 30 cm operant panel mounted on one wall. The operant panel contained three circular response ports spaced 6 cm center-to-center, aligned in a row with the center of each port roughly 14 cm off the floor of the cage and with the whole row centered on the width of the panel. Each response port was a PVC housed opening in the panel fitted with an IR receiver and transmitter that detected when the bird broke the plane of the response port with its beak. This 'poke-hole' design allows starlings to probe the apparatus with their beak, in a manner akin to their natural appetitive foraging behavior. Independently controlled light-emitting diodes (LEDs) could illuminate each response port from the rear. Directly below the center port, in the section of cage floor immediately adjacent to the panel, a fourth PVC lined opening provided access to food. A remotely controlled hopper, positioned behind the panel, moved the food into and out of the subject's reach beneath the opening. Acoustic stimuli were delivered through a small full-range audio speaker mounted roughly 30 cm behind the panel and out of the subject's view. The sound pressure level inside all chambers was calibrated to the same standard broadband signal. Custom software monitored the subject's responses, and controlled the LEDs, food hoppers, chamber light and auditory stimulus presentation according to procedural contingencies.

Stimuli

Recordings of three adult male European starlings were used to generate all the stimuli for this experiment. The procedure for obtaining high-fidelity song recordings from male starlings is detailed elsewhere (Gentner and Hulse 1998). Briefly, song was recorded from each male while housed individually in a large double-walled sound-attenuation chamber, but with visual and auditory access to a female starling. Male starlings sing in long continuous episodes called bouts. Song bouts are composed of much shorter acoustic units referred to as motifs (Adret-Hausberger and Jenkins 1988; Eens et al. 1988).

Motifs can be classified by their spectro-temporal characteristics into four broad classes: whistles, warbles, rattles, and high-frequency (Adret-Hausberger and Jenkins 1988; Eens et al. 1988) “Whistle” motifs are continuous, narrow band motifs. “Warble” motifs have a variable acoustic structure and include many heterospecific imitations. “Rattle” motifs are similar to warbles, but characterized by a broadband click train coincident to the more tonal elements (Braaten 2000). Motifs referred to as “high frequency” are characterized by high amplitude notes in the upper frequency spectrum. Of these classes, we extracted eight unique (non-repetitious) warble and rattle motifs from each male. Warbles possess variable acoustic structure with many heterospecific imitations. Rattles are similar to warbles, but are further characterized by a broadband click train coincident to the more tonal elements.

In constructing sets of training stimuli, the motifs extracted from a single male’s repertoire were used. Thus, the initial training set of 32 stimuli was built using warbles and rattles from one male where these elements are denoted as $a_i \in A$ and $b_i \in B$, respectively (i=1 to 8), yielding either AABB or ABAB sequences (Figure 1). The internal dependencies of the motif sequences used during training were the same as those originally used by Gentner et al (2006). Importantly, for these training stimuli each motif appeared exactly once in each possible position, and no motif appeared twice in the same sequence.

The extracted motifs from the songs of another male starling were used to create an additional set of 32 patterned training stimuli. To mark the fact that these stimuli sets were constructed from different motif libraries, we denote these patterns as CCDD and CDCD. From the remaining male starling song repertoire, a 16-exemplar set of patterns was built for generalization testing, denoted EEFF and EFEF.

Apparatus

The operant apparatus where starlings learned to classify the training stimuli is illustrated in Figure 2. Each subject was held in a small weld-wire cage containing an operant panel. On the panel, a centrally located response port was a PVC-housed opening. Inside of this opening was an IR receiver and transmitter used to detect when the bird broke the plane of the response port with its beak. This ‘poke-hole’ design allows starlings to probe the apparatus with their beaks, in a manner akin to their natural appetitive foraging behavior. A remotely controlled hopper, positioned behind the panel, moved the food within and beyond the subject’s reach beneath the opening. Acoustic stimuli were delivered through a small full-range audio speaker mounted behind the panel and out of the subject’s view. The sound pressure level inside all chambers was calibrated to the same standard broadband noise signal. Full details for all the mechanical components of the apparatus, audio interface, digital I/O control hardware, and custom software are available upon request.

Shaping

Upon initially entering the operant chamber, we provided each subject with unrestricted access to the food hopper. Following acclimation to eating from this device, the hopper was lowered beyond the subject's reach. Next, several times per minute, we presented the subject with a blinking LED in the center response port followed by 2-s access to the food hopper. This process recurred until the subject acquired a key-peck response (Brown & Jenkins, 1968). From then, the lowered food hopper would only be engaged if the subject pecked the blinking LED in the center port. Subjects repeated this behavior for 100 trials. After the completion of this phase, the center LED ceased blinking, requiring subjects to peck at the darkened center port to raise the food hopper. Following 100 such trials, pecking the center port initiated the playback of an acoustic stimulus where subjects earned food rewards in accordance with standard go/no-go training procedures described below.

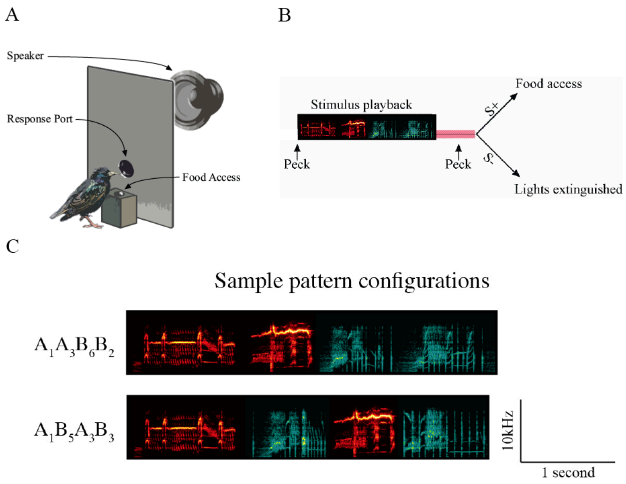

Figure 6.1. (A) Schematic of the operant apparatus. (B) Subjects start a trial by pecking the center response port. After the motif-pattern stimulus ends, the subject either pecks the center response port again or withholds any response depending on the class from which the stimulus was drawn. Correct responses yield a food reward. Incorrect responses lead to the house light being extinguished and food being inaccessible. (C) Spectrogram of stimuli used for training. The top spectrogram reflects a pattern form denoted in the text as AABB, while the lower spectrogram is of the pattern type denoted in the text as ABAB.

Classification training

We trained subjects to classify sets of starling song stimuli abiding by two different patterning rules. To initiate a trial, subjects pecked a darkened center port. Doing so triggered the playback of a training stimulus. Once stimulus presentation concluded, the animal either pecked the darkened center port once more (i.e., produced a 'go' response) or abstained from pecking (i.e., a 'no-go' response). Pecking behavior in response to one set of stimuli was positively reinforced with 2-s access to the food hopper (S+ training set). Alternatively, pecking in response to the other set of training stimuli results were punished in the form of extinguishing the operant box light for 10-50 seconds, thereby prohibiting food hopper access (S- training set). An absence of a response to either the S+ or S- stimuli yielded no operant consequence. Correction trials were available in the experiment during only some training phases and between (not during) blocks containing probe trials. For these subjects, a key-peck response within the response window following an S- stimulus not only extinguished the house lights as usual, but the program would also continue to present the same S- stimulus until the subject abstained from giving a key-peck response. Correction trials were discontinued for this experiment once the animal achieved our criterion for successful classification performance during the initial training phase (criterion: 3 successive d' $\geq$ 1.0). The stimulus exemplar presented on any given non-correction trial was sampled randomly with replacement from the pool of all stimuli the animal was learning to classify. To evaluate classification accuracy, we considered a response to an S+ stimulus and the withholding of a response to an S- stimulus as correct. Withholding a response to an S+

stimulus and responding to an S- stimulus were considered incorrect. Subjects could freely peck at the center response port throughout stimulus presentation, but only the first response within a 2-s response window beginning at stimulus offset triggered reinforcement or punishment. Subjects were on a closed economy during training, with daily sessions lasting from sunrise to sunset, and each subject could run as few or as many trials as they were able. Food intake was monitored daily to ensure each subject's well-being. The inter-trial interval was a minimum of 2-s. Water was always available. All procedures were approved by the UCSD institutional animal care and use committee.

Generalization testing

We applied two procedures to assess the strategies employed by starlings to acquire and generalize knowledge of the patterned sequences presented during training. The first was an abrupt transfer procedure to strictly novel stimuli for 5 consecutive 100-trial blocks. During transfers, we maintained all reinforcement contingencies as in training. The second was a probe procedure. Prior to initiation of the first probe session, the rate of food reinforcement for correct responses to S+ stimuli was lowered from 100% (where it had been during initial training) to 80%, and the rate of punishment (extinguished operant box lights) for incorrect responses to S– stimuli was lowered to 95%. All key peck responses to probe stimuli were non-differentially reinforced. Thus, a response to a probe stimulus had a 40% chance of eliciting 2-s food access, a 40% chance of eliciting punishment (extinguishing operant box lights), and a 20% chance of yielding no consequence at all. Due to the fact that probe stimuli reinforcement is random and non-differential with respect to response outcome, subjects have no opportunity to associate any probe stimulus with a given response. Thus, the classification of probe stimuli is commonly taken as strong evidence that the subject is classifying stimuli based on some set of features common to the probe and training classes rather than learning rote sets of specific exemplars. In the absence of generalization, classification accuracy should reflect chance levels.

Data analysis

We used d-prime (*d′*) to estimate the subjects' sensitivity for classifying stimuli in the two opposing classes (e.g. AABB and ABAB):

$$d' = z(go_1 - \varepsilon) - z(go_2 + \varepsilon),$$

where $z(go_1)$ is the z-score of the proportion of 'go' responses to stimuli in class 1 and $z(go_2)$ is the z-score of the proportion of 'go' responses to stimuli in class 2 adjusted by a factor of $\varepsilon = 0.001$ to set response rate bounds (0, 1), respectively. The measure *d′* is convenient because it eliminates any biases in the response rates (e.g. due to guessing) that may vary across individuals and within individuals over time. To gauge the effect of various song manipulations during the test sessions, we compared *d′* values for different stimulus classes using repeated measures ANOVA, and where appropriate used post-hoc analyses to quantify the significance of specific differences between mean *d′* measures. Estimates for d' are based on a minimum of 25 responses to each relevant class of stimuli under consideration.

Experiment 1: Generalization of artificial motif patterns

Under operant training procedures, we measured subjects' ability to generalize pattern knowledge. Generalization was tested in two ways. First, after learning to classify patterns built with specific transitions between motifs, we presented subjects with sequences containing novel transitions amongst training motifs. This was done separately for two different libraries of training warble and rattle motifs (i.e., training set one: 32 AABB/ABAB; training set two: 32 CCDD/CDCD) and allowed us to discern whether pattern learning was bound specifically to learned transitions. Second, we presented subjects with instantiations of the rule built from a novel set of warbles and rattles (i.e., EEFF/EFEF). This allowed us to explore whether subjects could recognize the learned patterned structures in the absence of any explicit training with the new constituents.

Results

Subjects initially learned to classify 8 AABB and 8 ABAB sequences. Our performance criteria of three consecutive 100-trial blocks with d' ≥ 1.0 was reached in 7750 ± 606 trials (Figure 2a). Shortly after, an additional 8 AABB and 8 ABAB sequences were added to the original training stimuli. Performance remained high during this stage (Figure 2b). Subjects were then transferred to 5 100-trial blocks in which each trial delivered a novel patterned sequence of AABB or ABAB. None of these 500 sequence arrangements were shared between subjects nor did they include any of the original 32 training stimuli. This procedure was used for two reasons. First, previous testing of starlings' ability to discern new patterns built from the training set of warbles and rattles only tested 16 new sequences over 500 trials. Therefore, it could be that subjects generally exhibit rapid learning of those new dependencies. Second, the stimulus combinations presented here were not bound by the same regulations as the training stimuli such that repetitions of single motifs could indeed occur in the instantiation of either pattern. Even though performance declined for some subjects on this transfer condition compared to the responses in the prior 5 training blocks (Subj. 635: $t = 5.02$, $df = 4$, $p < 0.008$; Subj. 637: $t = 0.55$, $df = 4$, $p = 0.609$; Subj. 651: $t = -0.05$, $df = 4$, $p =$

0.961; Subj. 659: t = 2.87, df = 4, $p < 0.05$; all matched pairs t-test), all four subjects' d' performance was significantly above chance levels (Subj. 635: t = 10.69, df = 4, $p < 0.001$; Subj. 637: t = 7.60, df = 4, $p < 0.002$; Subj. 651: t = 7.35, df = 4, $p < 0.002$; Subj. 659: t = 4.87, df = 4, $p < 0.009$; all single sample t-tests where chance d' is 0; Figure 2c). Across this transfer condition, we analyzed subjects' performance for those trials in which a motif item was repeated. Importantly, repeated motifs never occurred in any of the training sets. Consistent with previous work (Comins and Gentner 2014), performance on trials where a motif repetition occurred exceed levels expected by chance (Subj. 635: $\chi^2 = 19.19$, df = 1, $p < 0.001$; Subj. 637: $\chi^2 = 32.18$, df = 1, $p < 0.001$; Subj. 651: $\chi^2 = 30.87$, df = 1, $p < 0.001$; Subj. 659: $\chi^2 = 14.64$, df = 1, $p < 0.001$). Finally, one possible explanation of the strong accuracy for classification of novel A & B motif patterns is that subjects may have learned explicit sub-sequences (or fragments, or n-grams) of motifs during the initial training. If this is true, then the A & B-motif test patterns that share any positionally accurate bi-gram combinations (e.g., where bi-grams occur in their learned positions) with the A & B training patterns should be more easily classified than those that are wholly unique to the test set. Analyzing the response data according to these stimulus differences refutes this possibility. In all subjects, classification accuracy for wholly novel patterns did not differ from that for patterns that contained bi-grams from training stimuli (Subject 635: $\chi^2 = 0.01$, df = 1, $p = 0.91$; Subject 637: $\chi^2 = 0.79$, df = 1, $p = 0.38$; Subject 651: $\chi^2 = 3.46$, df = 1, $p = 0.06$; Subject 659: $\chi^2 = 0.03$, df = 1, $p = 0.87$; all Pearson's chi-squared tests; Figure 3).

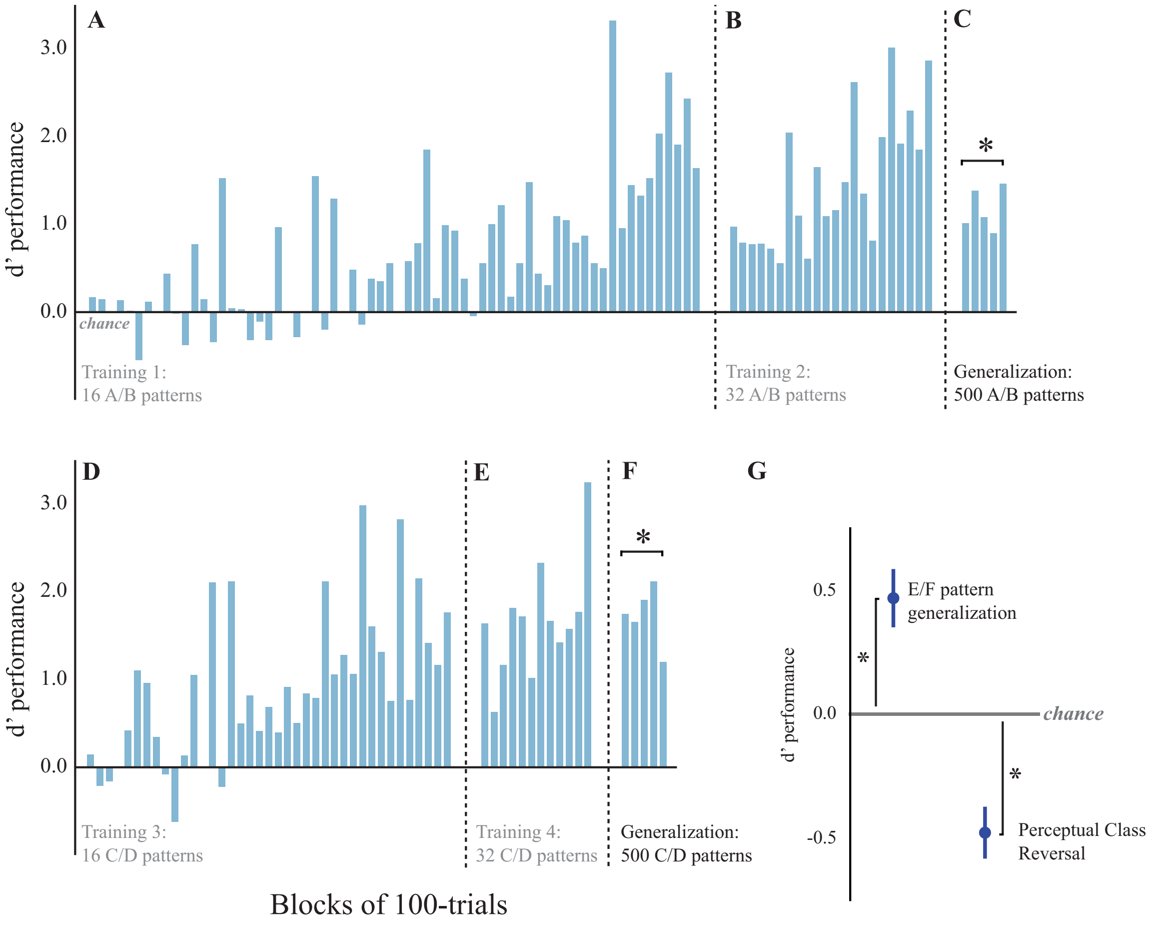

Figure 6.2. Learning results for starlings in experiment 1 (A) Acquisition of baseline 8 AABB and 8 ABAB sequence classification task for example subject 635. Each bar represent the d' value for a block of 100 hundred trials. (B) Classification of 16 AABB versus ABAB patterns. (C) Generalization phase comprised of 500 trials of novel AABB or ABAB patterns. (D) Following training and testing on A/B patterns, birds were transferred to training on 8 CCDD and 8 CDCD patterns, (E) followed by 16 CCDD and 16 CDCD patterns and (F) finally tested on 500 novel C/D patterns. (G) Following this training procedure, subjects as a population demonstrated reliable generalization to familiar patterns built from entirely novel elements (E&F) while, surprisingly, significantly worse than expected performance on novel A&B and C&D patterns when the order of perceptual classes was reversed. This suggests that subjects in experiment 1 might have acquired 'local' rules to generalize pattern knowledge (see text).

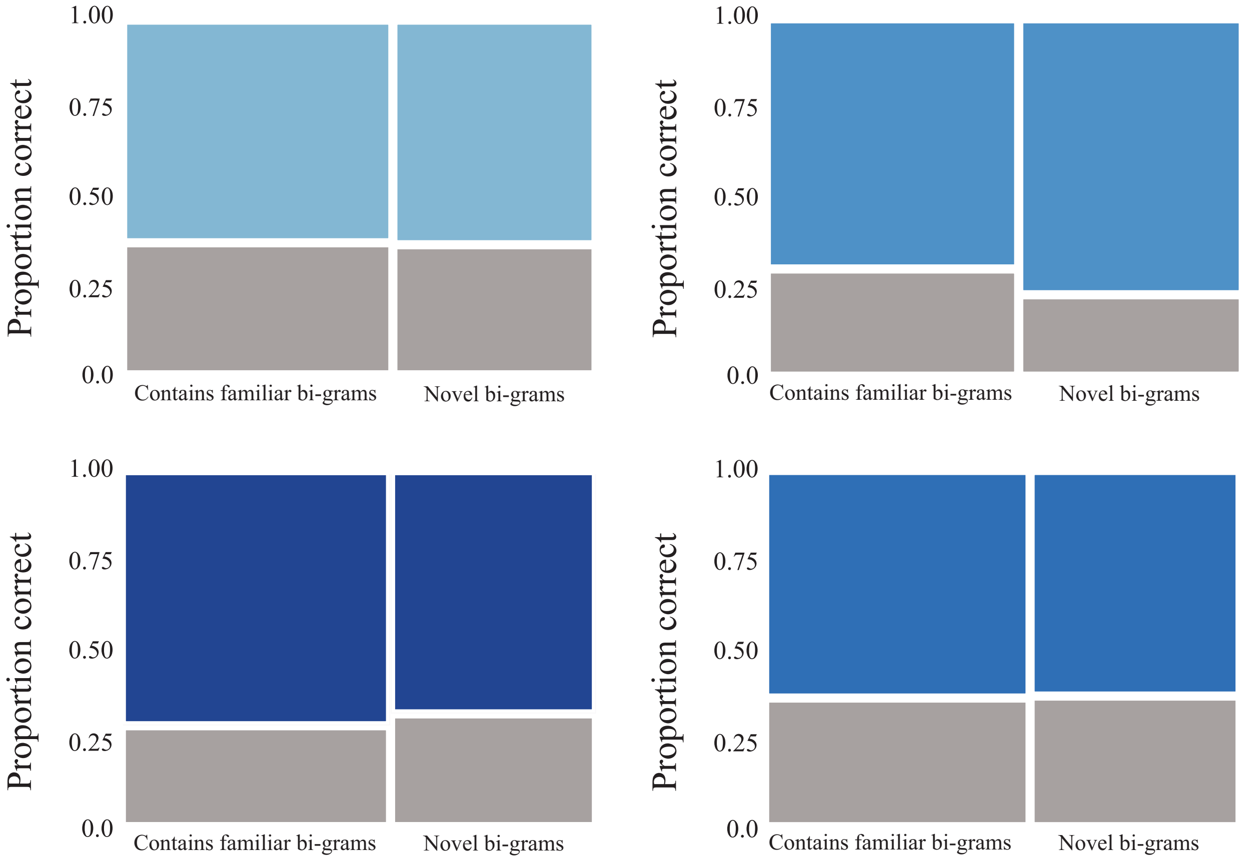

Figure 6.3. Mosaic plots comparing proportion correct of AABB versus ABAB classification during the transfer phase for those pattern containing familiar bi-grams to those without them (see text). Gray represents incorrect responses, while blues represent correct responses. Each subject is represented by a different saturation of blue.

After learning A/B sequences, subjects were trained on patterns composed of a different set of warbles and rattles (C's and D's, respectively). As before, the initial stage of training consisted of 16 explicit patterned sequence, 8 CCDD and 8 CDCD (Figure 2d). Subjects reached our performance criterion of three consecutive 100-trial blocks with a $d' \geq 1.0$ quicker than in the initial AABB/ABAB training phase, requiring only 2650 ± 366 trials ($t = -6.34$, $df = 3$, $p < 0.008$). Soon thereafter, an additional 8 CCDD and 8 CDCD sequences were merged with the current stimulus repertoire, bringing the total set of C/D pattern stimuli to 32 sequences (Figure 2e). Subjects remained in this stage of training until d' scores returned to high levels. At this point, subjects were transferred to 5 100-trial blocks where each trial was a novel patterned sequence of CCDD or CDCD. Overall, d' performance was significantly above chance levels (Subj. 635: t = 11.29, df = 4, $p < 0.001$; Subj. 637: t = 5.35, df = 4, $p < 0.006$; Subj. 651: t = 6.26, df = 4, $p < 0.004$; Subj. 659: t = 4.2, df = 4, $p < 0.02$; single sample t-tests where chance d' is 0; Figure 2f). Additionally, only one subject demonstrated a significant drop in d' performance during this transfer compared to the previous 5 training blocks (Subj. 635: t = -0.625, df = 4, $p = 0.5656$; Subj. 637: t = -0.3234, df = 4, $p = 0.7626$; Subj. 651: t = 0.098, df = 4, $p = 0.926$; Subj. 659: t = -3.38, df = 4, $p < 0.03$; all matched pairs t-test). The results of the two transfer conditions suggest that subjects rely relatively little on knowledge of the transitions between motifs in these patterns.

Pattern generalization to unfamiliar warbles and rattles

Thus far, our tests of pattern generalization extend only to novel arrangements of familiar motifs from the two training sets. In the next phase, we tested subjects' ability to generalize the learned patterned relationships to unique strings of warbles and rattles (E's and F's). Probe trials containing EEFF or EFEF patterns were interspersed with the presentation of stimuli from the aforementioned large baseline classes. Though d' performance dropped precipitously compared to baseline (t = 9.32, df = 3, $p < 0.003$ matched pairs t-test), subjects classified EEFF and EFEF sequences above chance levels on both the individual level (3 out of 4 cases: chi-square tests, $p < 0.05$) and as a population (t = 3.98, df = 3, $p < 0.03$ single sample t-test where chance d' is 0; $\mu \pm$ SE d' scores of 0.47 ± 0.12; Figure 2g).

Perceptual class reversals

Here we sought to explore the rule used by starlings to classify these patterns. Given that these subjects always encountered a warble motif in the second position of AABB/CCDD/EEFF patterns and a rattle in the second position of ABAB/CDCD/EFEF patterns, subjects might learn highly specific rules about the relationship between perceptual classes and their temporal positions in the sequence. If this were the case, we would anticipate that subject would misclassify sequences that preserve the overall motif patterning structure but are contrarian to specific relations between motif categories and their temporal positions. For example, because subjects were required to differentiate patterns that unfold as AABB from patterns that proceed as ABAB, they might consider a pattern organized in the reverse direction, such as BBAA, as more similar to the ABAB forms due to the presence of a rattle motif in the second position and a warble motif in the third position. To test for this possibility, we manipulated stimuli so that the global patterning structure would be preserved but the perceptual classes in all positions would be switched.

When subjects encounter the patterns BBAA/DDCC and BABA/DCDC, the proportion of probe false alarms drove a significant change in performance compared to baseline (t = -18.71, df = 3, $p < 0.001$; matched pairs t-test). Further, these negative d’ scores from subjects are significantly different from expected chance levels (t = -3.78, df = 3, $p < 0.04$; Figure 2g). Thus, it does appear that subjects rely on the information present in the second and third sequence positions more so than a more global integration of information across all four constituents of the pattern for this task.

Discussion

Overall, these results support the conclusion that, for starlings, acquired knowledge about motif patterns is tied initially to the representation of specific motifs and to the presentation of those motifs in sequences. Learned pattern recognition does not automatically generalize to other sets of motifs. With appropriate experience, however, acquired patterning knowledge can generalize, albeit modestly, to sets of novel motifs. Thus, European starlings can learn complex temporal patterns, and can learn to apply these patterns in multiple contexts. Based on these results, we conclude that knowledge of a pattern in this species is, in principle, dissociable from the explicit elements that evince a given pattern.

Experiment 2: Do starlings learn rules 'rationally'?

The propensity for rule learning over temporal patterns seen during experiment 1 raises an important question: what factors influence the use of a particular rule governing pattern constructions when other candidate rules are present? Insights from cognitive psychology suggest that human rule-inductions arise from rational decision models that favor recognizing pattern structures according to the narrowest set of distinguishable features (Frank and Tenenbaum 2010). As an example, consider the following: what do 5, 15, 25, 35 and 1005 have in common? The most apparent answer is that these numbers are all divisible by 5. How about 5, 15, 25, 35, 1005 and 1? In this case, divisible by 5 is no longer accurate, but concluding that all of these numbers are divisible by 1 is accurate. Interestingly, divisibility by 1 was equally accurate in the former list of numbers as well – so why does the answer *divisible by 5* appear to leap from the page for this set? The proposed solution is that *a priori*, grabbing a random handful of numbers that are all divisible by 5 is much less likely than a handful of numbers all divisible by 1. A rational rule-induction process considers whether a candidate rule accurately describes the data and whether this candidate rule is *a priori* more unlikely than other competing candidate rules.

Whether such rational decision-making processes for rule learning are unique to humans or shared across taxa remains an important open question. If starlings are to serve as a useful model organism for human rule learning, it is critical to know whether their rule induction process is similar. Based on results from the "perceptual class reversals" condition in experiment 1, subjects appear to acquire a specific patterning rule akin to *warble-warble-rattle-rattle* as opposed to learning a more flexible and broader rule like *category 1 - category 1 - category 2 - category 2*. Unlike the former rule, the latter would be able to recognize the global similarity between patterns that proceed as *warble-warble-rattle-rattle* and those that proceed as *rattle-rattle-warble-warble.* Thus, in the current experiment, we sought to explore whether subjects could, indeed, acquire this broader patterning rule to classify such sequences using only a subtle modification of the prior training procedure.

As in experiment 1, subjects were initially trained and tested on the classification of AABB/ABAB patterns. However, during the second phase, subjects were trained and tested on DDCC/DCCD patterns (i.e., patterns that begin with a rattle motif), which could serve as "counter-factual" evidence to the rule that patterns proceed as *warble-warble-rattle-rattle* instead of *category 1 - category 1 - category 2 - category 2.* Finally, we used a probe procedure to test whether these starlings could generalize pattern knowledge to novel motifs (EEFF/EFEF) and to patterns comprised of familiar items where the temporal positions of perceptual classes are reversed (in this experiment, BBAA/BABA and CCDD/CDCD).

Results

Subjects were initially trained to classify 8 AABB and 8 ABAB sequences. Performance criteria of three consecutive 100-trial blocks with $d' \geq 1.0$ was reached in 7000 ± 2494 trials (Figure 4a; acquisition time not statistically different compared to subjects' performance in experiment 1; $t = 029$; $p = 0.795$; unmatched t-test). Next, a different set of 8 AABB and 8 ABAB sequences were added to the original training stimuli and subjects continued to perform well at classifying training stimuli (Figure 4b). During the transfer phase of 500 novel A/B patterns, all subjects maintained better than chance performance (Subj. 692: $t = 6.736$, $df = 4$, $p < 0.003$; Subj. 725: $t = 6.312$, $df = 4$, $p < 0.004$; Subj. 830: $t = 4.046$, $df = 4$, $p < 0.02$; Subj. 659: $t = 4.87$, $df = 4$, $p < 0.009$; all single sample t-tests where chance d' is 0; Figure 4c).

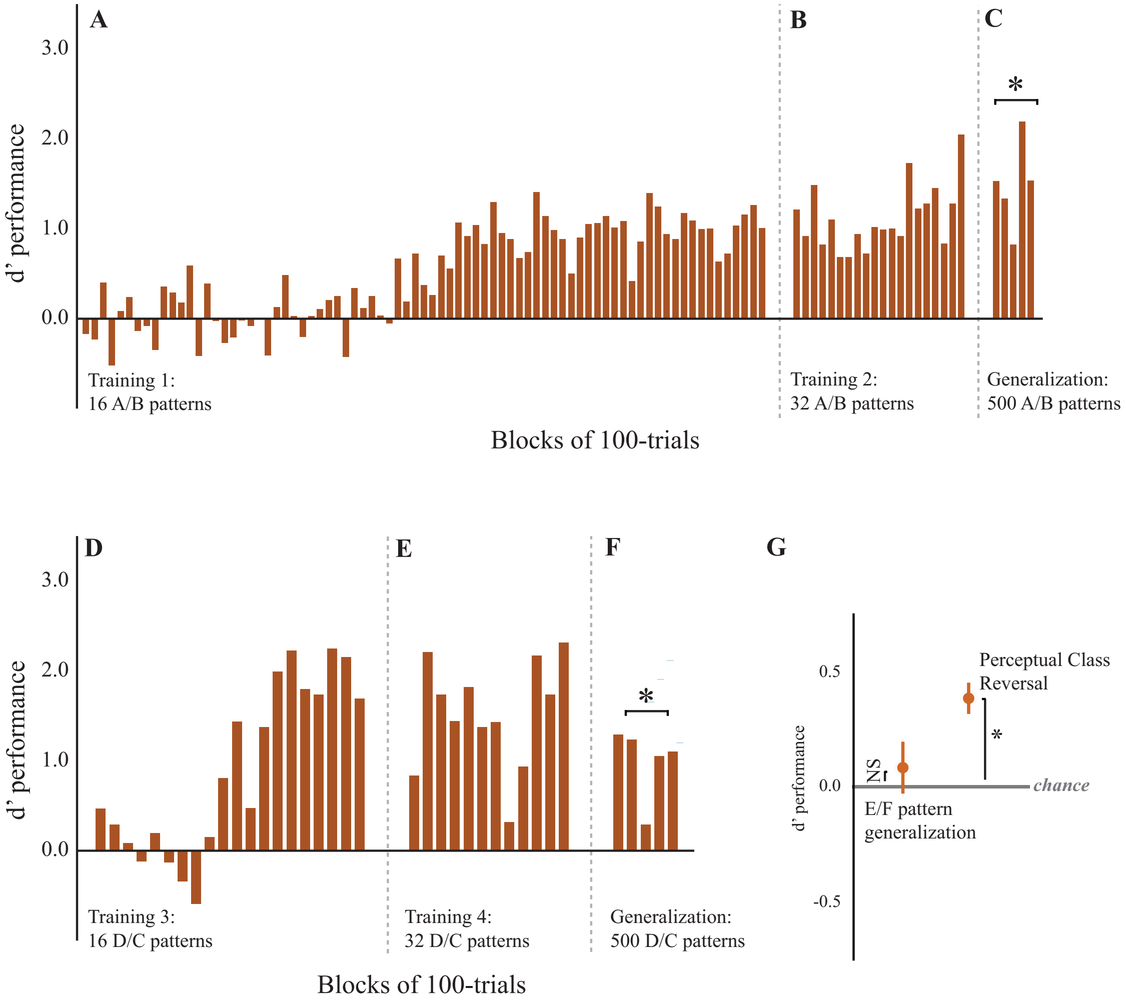

Figure 6.4. Learning results for starlings in experiment 2 (A) Acquisition of baseline 8 AABB and 8 ABAB sequence classification task for example subject 692. Each bar represent the d' value for a block of 100 hundred trials. (B) Classification of 16 AABB versus ABAB patterns. (C) Generalization phase comprised of 500 trials of novel AABB or ABAB patterns. (D) Following training and testing on A/B patterns, birds were transferred to training on 8 DDCC and 8 DCDC patterns, (E) followed by 16 DDCC and 16 DCDC patterns and (F) finally tested on 500 novel D/C patterns. (G) Following this training procedure, subjects as a population did not show any reliable generalization to familiar patterns built from entirely novel elements (E&F). Intriguingly, though starlings in experiment 2 did not generalize to strings of wholly novel motifs, they did show reliable classification performance on novel A&B and C&D patterns where the order of perceptual classes was reversed. Unlike subjects in experiment 1, this suggests that birds in experiment 2 might have acquired 'broader' rules to classify patterns (see text).

After learning AABB/ABAB sequences, subjects were trained on patterns made from a new set of warbles and rattles (C's and D's, respectively). Crucially, these motifs were now arranged in reverse order so that the prior learned relationship between motif categories and temporal positions was inverted. Thus, while patterns made from As and Bs always proceeded as *warble-warble-rattle-rattle* and *warble-rattle-warble-rattle*, the patterns made from Cs and Ds for this group of birds would proceed as *rattle-rattle-warble-warble* and *rattle-warble-rattle-warble*. The goal was to provide subjects with evidence supporting a broader rule, such as *category 1 - category 1 - category 2 - category 2*. Once again, subjects learned to classify 16 explicit patterned sequences, 8 DDCC and 8 DCDC (Figure 4d), though interestingly not statistically quicker than they acquired knowledge of AABB versus ABAB patterns. After reaching criterion, novel configuration of 8 DDCC and 8 DCDC sequences was added to the stimulus set such that subjects were classifying 32 patterned sequences of D/C motifs (Figure 3e). Finally, subjects were transferred to 5 100-trial blocks of DDCC or DCDC, where each trial presented a novel D/C patterned sequence. As before, their d' performance was significantly above chance levels on these trials (Subj. 692: t = 5.55, df = 4, $p < 0.006$; Subj. 725: t = 6.06, df = 4, $p < 0.004$; Subj. 830: t = 7.92, df = 4, $p < 0.002$; single sample t-tests where chance d' is 0; Figure 4f). Overall, the foregoing results of the two transfer conditions show that subjects in this experiment accurately learn and generalize temporal patterns comprised of familiar elements.

Pattern generalization to unfamiliar warbles and rattles

Following training with A/B and D/C patterns, starlings were tested using probe trials on patterns comprised of novel warbles and rattle (EEFF and EFEF). Interestingly, neither analyses of individual performance (in all 3 cases, chi-square tests, $p > 0.05$) nor group performance indicate any meaningful generalization to patterns built from novel motifs (t = .073, df = 2, $p = 0.55$; single sample t-test where chance d' is 0; $\mu \pm$ SE d' scores of 0.08 ± 0.11; Figure 4g).

Perceptual class reversals

We hypothesized that if the starlings here indeed acquired a different rule compared to subjects from experiment 1, they should continue to perform well when classifying patterns of familiar elements when the temporal positions of warbles and rattles were reversed. In other words, given that these starlings were trained on patterns of the form AABB (*warble-warble-rattle-rattle*) and ABAB followed by DDCC (*rattle-rattle-warble-warble*) and DCDC, we anticipated subjects would be able to accurately classify novel arrangements of these familiar elements that present the motifs in reverse order. Thus, we probed starlings with BBAA/BABA and CCDD/CDCD patterns. As predicted, starlings classified probe trials significantly better than chance ($t = 5.63$, $df = 2$, $p < 0.03$ single sample t-test where chance d' is 0; Figure 4g).

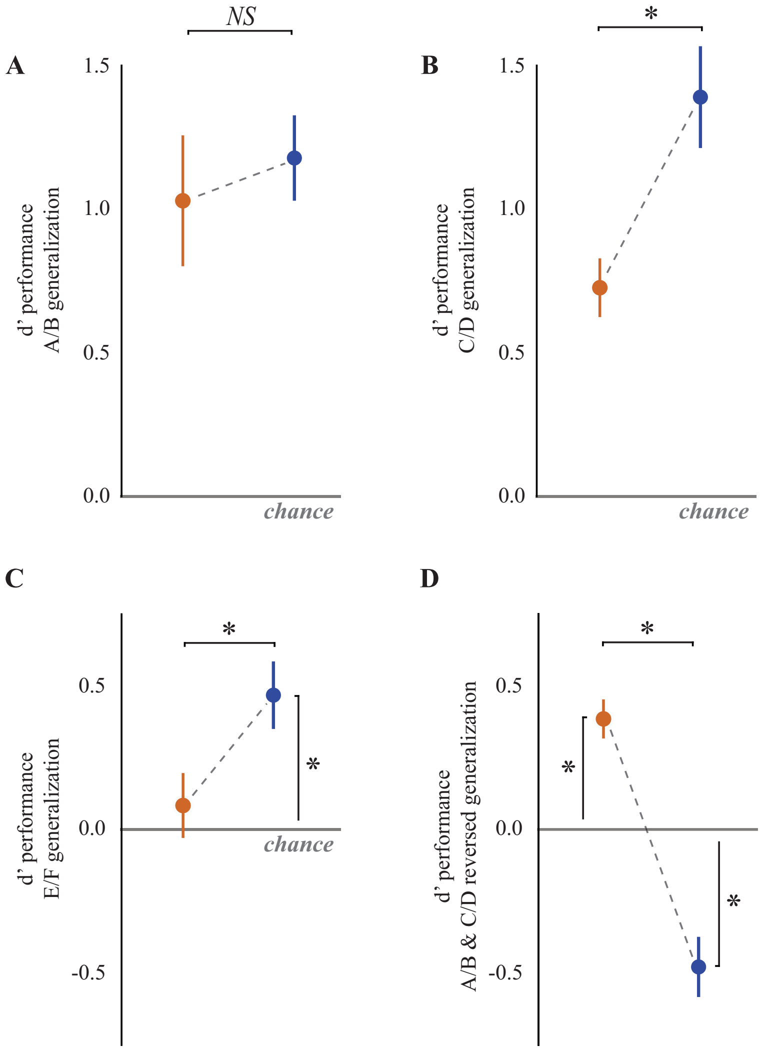

Figure 6.5. (A) Starlings from experiment 1 and 2 performed comparably on generalization trials for A/B pattern, with both group performing significantly above chance. (B) Starlings from experiment 1 significantly outperformed starlings from experiment 2 when generalizing to patterns built from Cs and Ds – though starlings from experiment 2 still performed better than expected due to chance. (C) When tested on patterns built from unfamiliar motifs (Es and Fs), only subjects from experiment 1 demonstrated reliable performance. (D) Intriguingly, this difference might be related to starlings in experiment 1 learning a considerably “simpler” rule than starlings in experiment 2 – thus, when starlings in experiment 1 were presented with patterns that reversed the traditional order of warbles and rattles, they performed significantly worse than expected by chance while subjects in experiment 2 still accurately classified these probe trials.

Comparing generalization performance between experiments 1 and 2

The foregoing results are consistent with starlings from experiment 1 being able to generalize a patterning rule to sequences of unfamiliar motifs, albeit by use of a local rule. Contrarily, starlings from experiment 2 appear to have acquired a broader rule thereby allowing them to more flexibly generalize novel patterns built from familiar motifs, but at the expense of extrapolating this knowledge to sequence of unfamiliar motifs. In this section, we compare the generalization performance across subjects from experiments 1 and 2 to test these assessments. In comparing performance on A/B generalization trials we find no marked difference between starlings from experiment 1 and 2 ($t = 0.54$, $p = 0.62$; unmatched pairs t-test; Figure 5a). Given that training and testing conditions were identical up until this point for both experiment 1 and 2, this demonstrates that these two groups of birds were equivalent in their generalization abilities. Following the next training phase with C/D (or D/C) patterns, starlings in experiment 2 were markedly impaired compared to the performance seen by experiment 1 subjects ($t = 3.24$, $p < 0.03$; unmatched pairs t-test; Figure 5b). This supports the notion that the starlings in experiment 2 were employing a different rule than subjects in experiment 1 during this generalization test. During the generalization test for patterns built from unfamiliar motifs, subjects from experiment 1 significantly outperformed starlings from experiment 2 ($t = 2.3$, $p = 0.06$; unmatched pairs t-test; Figure 5c), which is consistent with our failure to find any reliable generalization on E/F probe patterns for experiment 2 subjects. Finally, we find a significant enhancement in experiment 2 performance over experiment 1 performance in classifying patterns of familiar elements where the order of perceptual classes was reversed ($t = -6.9$, $p < 0.002$; unmatched pairs t-test; Figure 5d). This lends credence to the idea that, unlike starlings from experiment 1, subjects in experiment 2 acquired a different, broader rule for classifying novel patterns created from familiar elements. As a result, their performance is robust to pattern variations that invert the order of perceptual classes across time.

General Discussion

Understanding the behavioral expression of temporal pattern knowledge and its relationship to the underlying pattern elements is central to both psychological and neurophysiological theories of how brains process complex real-world auditory signals and generalize knowledge to new situations. This is particularly true for signals involved in conspecific communication. Previous work demonstrates that starlings can learn to classify sequences of conspecific song motifs that follow patterns defined by two different patterning forms and can generalize this acquired knowledge to novel sequences drawn from the same elements, i.e. those constructed using the same patterning rules and the same motif vocabulary heard in training (Comins and Gentner 2014; Comins and Gentner 2013; Gentner et al. 2006). The present study examined the extent to which pattern recognition acquired in the context of one motif vocabulary might generalize to another motif vocabulary. In so doing, we are interested in the broader question of whether acquired knowledge of a pattern is separable from the elements that explicitly instantiate a pattern, and the general character of the knowledge that underlies learned recognition of patterned motif sequences.

We find that generalization is closely constrained by the animal’s prior experience. After learning to recognize patterns instantiated by one set of motifs, starlings failed to recognize the same patterns built from a new set of motifs even when these motifs were drawn from the same acoustic categories as the training motifs. If, however, starlings were given explicit experience in applying the same learned patterns to multiple sets of motifs, pattern knowledge generalized to sets of novel motifs. In addition, we demonstrate how variations in pattern experience can alter decision-making processes that are highly reminiscent of "rational" rule-learning effects found in humans. Taken together, this evidence suggests that starlings might serve as a powerful model for studying the neurobiological basis of rule learning.

Learned Generalization

We first trained starlings to recognize patterns built from two sets of song motifs (A&B), where each set was defined by the spectro-temporal characteristics common to two natural classes of starling learned vocalizations (i.e. rattles and warbles; (Braaten 2000). Next, we transferred the subjects classify patterns built from two unfamiliar sets of motifs (C&D). Performance on this transfer was initially quite poor. Over time, however, subjects once more reached a stable accurate asymptotic level of performance and we tested their generalization to patterns built from two sets of unfamiliar motifs once more (E&F). The E and F motifs were drawn from the same natural rattle and warble classes as the A/C and B/D motifs, respectively. To our surprise, the subjects were able to identify the well-formed sequences composed of all three sets of motifs (A/B, C/D, and E/F; Figure 2g and 4c) at levels significantly above chance. This result suggests that appropriate experience can alter the default strategies for motif-pattern generalization, and broaden the context over which acquired patterning knowledge can be used. Thus it appears that generalization itself is learned.

The inability of subjects to immediately generalize pattern knowledge when initially transferred to patterns built from unfamiliar warbles and rattles is consistent with the possibility that starlings simply cannot apply the patterning rules acquired in one context to larger sets of motifs. However, experience can temper these constraints so that pattern knowledge can generalize to sequences built from wholly unfamiliar motifs. It is noteworthy that the level of performance observed in the second generalization test for E & F-motif patterns is close to that seen for the generalization of acquired pattering rules in other animals including rats (Murphy, 2008), pigeons (Herbranson, 2003; Herbranson, 2008), and the original AGL descriptions in humans (Reber, 1967; Reber, 1969).

Our results entreat the question, do constraints on the pattern generalization seen here constitute a fundamental difference between the pattern recognition abilities of humans and songbirds? Beginning in infancy, humans are prodigious pattern generalizers. For example, when children are exposed to inconsistent use of grammatical forms during language acquisition, they tend to regularize those forms as they learn the language (Hudson-Kam, 2005). However, human pattern generalization does not exist without its own limitations (Conway and Christiansen 2006). Indeed, while infants readily generalize patterning rules to speech sounds (Gomez, 1999), the ability to generalize rules to non-speech signals requires that they first hear the rule instantiated over speech sounds (Marcus et al, 2007). Given results showing that starlings require familiarity with the perceptual organization of motifs for pattern generalization (Comins and Gentner 2013), this suggests human infants might share similar constraints initially – though, they eventually overcome such constraints to understand patterns operating over more functional rather than perceptual categories. We have demonstrated that at least in principle the ‘rules’ describing patterned sequences and the classes of elements that instantiate those patterns are dissociable in non-humans.

Which Rule?

Developmental evidence has demonstrated that human infants deploy a suite of statistical learning mechanisms to identify patterning rules across multiple timescales in language, from phonemes to words (Aslin and Newport 2012). Even more, infants can recognize familiar patterns of speech elements over time, even when the speech elements themselves are unfamiliar. But how do newborns identify the veridical rules governing pattern constructions in speech and language despite the fact that the number of candidate-rules giving rise to these structures is potentially infinite (Chomsky 1965)? In the classic experiment by Marcus et al (1999), for example, it was shown that infants can recognize the similarity between phoneme patterns such as ABA, where A and B represent any two distinct speech sounds. But what strategy are infants using to recognize the familiar structure across renditions of the pattern made from different speech sounds? Recent findings show that infant rule-inductions arise from rational decision models that favor recognizing pattern structures according to the narrowest set of distinguishable features consistent with their experience (Dawson and Gerken 2011; Gerken 2010). For instance, sequences of speech syllables like */ba/-/di/-/ba/* or */do/-/di/-/do/* can both be described by "broad" rules, like ABA, and "narrow" rules like A-*/di/*-A (where any syllable A occurs before and after the specific syllable */di/*). Only when presented counterfactual evidence to the narrow rule, such as *te-la-te*, will infants generalize pattern knowledge according to a broader rule, ABA.

Results from our second experiment address how experience shapes the kinds of rules starlings acquire to classify patterns. Consistent with work in human infants (Dawson and Gerken 2011), songbirds use more locally defined rules to classify patterns in the absence of counterfactual evidence that those rules are obsolete. However, when starlings are presented such counterfactuals, they do generalize this local rule to other contexts (evinced by starlings from experiment 2 accurately classifying patterns where perceptual classes were reversed).

Strikingly, despite being matched for experience with motifs patterns, experiment 2 starlings never demonstrated reliable pattern generalization when those patterns were built from unfamiliar motifs. So why is the broad rule group successfully generalizing patterns of familiar elements and failing to apply this knowledge to a novel situation? This failure might be linked to how these seemingly rational rule learning processes interact with other cognitive systems, such as perception, memory and attention (Endress 2013). For instance, one requirement for success on this task is the identification and categorization of motifs into warbles and rattles and then assigning these categories to temporal positions in memory. Previous and ongoing work has shown that starlings in fact improve their perceptual discrimination of warbles and rattles as a result of classifying patterns of these motifs (Comins & Gentner, *Chapter Five*). Thus, given that experiment 1 birds know where to focus their attention during the task (second and third positions of the pattern), they might develop more perceptual expertise along those dimensions that divide rattles and warbles, and can employ this knowledge in the context of EEFF and EFEF pattern generalization.

Songs, speech and language

Our results demonstrate a close, although not exclusive, relationship between the acoustic features of underlying pattern elements and the representation for patterns. Statistical learning among humans also shows effects that may be tied ultimately to acoustic properties of speech stimuli, but with curious constraints. For instance, adult humans can learn non-adjacent dependencies between speech sounds so long as the patterned sequences are made up of segments (e.g. consonant-vowel-consonant, C-V-C; or V-C-V) rather than syllables (CV-CV-CV) (Newport 2004). A species of non-human primate, the cotton-top Tamarin, however, can learn non-adjacent statistical regularities among CV and vowel sounds, but not consonants. Other results indicate a significant advantage for audition in statistical learning (Conway 2005) further supporting a strong tie between pattern and sensory representations.

These results combined with the present study, indicate that prior experience and the resulting expectancies constrain processing of patterned acoustic sequences. It will be important to future work to examine these constraints comparatively. Moreover, as these processes appear to operate on classes of underlying elements, understanding the emergence of such classes is of critical importance.

Finally, although language is uniquely human, several contemporary theories of language comprehension suggest that syntactic ambiguity, and its resolution, derives in part from the lexical representation of words (Samuel 2001). These theories imply that if the representation of meaning and syntax are separable at all, they are nonetheless closely tied to one another. Similar ideas are being applied to understanding the relationship between the perceptual, or phonological, structure of language and its linguistic structure. For instance, shared phonological (i.e., perceptual) features in words correlate with grammatical category identity, such as case-endings or gender-markers (Frigo and McDonald 1998; Kelly 1992; Monaghan et al. 2007; Shi et al. 1998). Thus, recent theories suggest that humans infants can learn the phonological structure of a language to serve as the scaffolding for later knowledge over more abstract features of language (Frigo and McDonald 1998; Kelly 1992; Monaghan et al. 2007; Shi et al. 1998).

Acknowledgements

This chapter, in part, is currently being prepared for submission for publication of the material. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

References

Adret-Hausberger M, Jenkins P (1988) Complex Organization of the Warbling Song in the European Starling Sturnus vulgaris. Behaviour 107:138–156.

Aslin RN, Newport EL (2012) Statistical Learning: From Acquiring Specific Items to Forming General Rules. Current Directions in Psychological Science 21:170–176. doi: 10.1177/0963721412436806

Braaten RF (2000) Multiple levels of representation of song by European starlings (Sturnus vulgaris): Open-ended categorization of starling sing types and differential forgetting of song categories and exemplars. Journal of Comparative Psychology 114:61–72. doi: 10.1037//0735-7036.114.1.61

Chomsky N (1965) Aspects of the theory of syntax. The MIT Press, Cambridge, MA

Comins J a, Gentner TQ (2014) Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. Animal cognition. doi: 10.1007/s10071-014-0732-5

Comins JA, Gentner TQ (2013) Perceptual categories enable pattern generalization in songbirds. Cognition 128:113–118. doi: 10.1016/j.cognition.2013.03.014

Conway CM, Christiansen MH (2006) Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. Psychological science 17:905–12. doi: 10.1111/j.1467-9280.2006.01801.x

Dawson C, Gerken L (2011) When global structure “Explains Away” local grammar: a Bayesian account of rule-induction in tone sequences. Cognition 120:350–9. doi: 10.1016/j.cognition.2010.11.016

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86.

Emberson LL, Liu R, Zevin JD (2013) Is statistical learning constrained by lower level perceptual organization? Cognition 128:82–102. doi: 10.1016/j.cognition.2012.12.006

Endress AD (2013) Bayesian learning and the psychology of rule induction. Cognition 127:159–76. doi: 10.1016/j.cognition.2012.11.014

Endress AD, Scholl BJ, Mehler J (2005) The role of salience in the extraction of algebraic rules. Journal of experimental psychology General 134:406–19. doi: 10.1037/0096-3445.134.3.406

Frank MC, Tenenbaum JB (2010) Three ideal observer models for rule learning in simple languages. Cognition

Frigo L, McDonald JL (1998) Properties of Phonological Markers That Affect the Acquisition of Gender-Like Subclasses. Journal of Memory and Language 39:218–245. doi: 10.1006/jmla.1998.2569

Gentner D, Namy LL (2006) Analogical Processes in Language Learning. Current Directions in Psychological Science 15:297–301. doi: 10.1111/j.1467-8721.2006.00456.x

Gentner T, Hulse S (1998) Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Animal Behaviour 56:579–594.

Gentner TQ (2008) Temporal scales of auditory objects underlying birdsong vocal recognition. The Journal of the Acoustical Society of America 124:1350–1359.

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440:1204–1207.

Gentner TQ, Hulse SH (2000a) Female European starling preference and choice for variation in conspecific male song. Animal Behaviour 59:443–458.

Gentner TQ, Hulse SH (2000b) Perceptual classification based on the component structure of song in European starlings. The Journal of the Acoustical Society of America 107:3369–3381.

Gerken L (2010) Infants use rational decision criteria for choosing among models of their input. Cognition 115:362–6. doi: 10.1016/j.cognition.2010.01.006

Gomez RL, Gerken L (1999) Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70:109–35.

Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. Cognition 78:B53–64.

Hulse S, Humpal J, Cynx J (1984) Discrimination and Generalization of Rhythmic and Arrhythmic Sound Patterns by European Starlings (Sturnus vulgaris). Music Perception 1:442–464.

Kelly MH (1992) Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. Psychological Review 99:349–64.

Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a Comparative Neurobiology of Language. Frontiers in Evolutionary Neuroscience 4:1–13. doi: 10.3389/fnevo.2012.00006

Knudsen DP, Gentner TQ (2010) Mechanisms of song perception in oscine birds. Brain and language 115:59–68.

Marcus G, Vijayan S, Brandi Rao S, Vishton PM (1999) Rule Learning by Seven-Month-Old Infants. Science 283:77–80.

Marcus GF (2000) Pabiku and Ga Ti Ga: Two Mechanisms Infants Use to Learn About the World. Current Directions in Psychological Science 9:145–147. doi: 10.1111/1467-8721.00080

Marcus GF, Fernandes KJ, Johnson SP (2007) Infant rule learning facilitated by speech. Psychological science 18:387–91. doi: 10.1111/j.1467-9280.2007.01910.x

Monaghan P, Christiansen MH, Chater N (2007) The phonological-distributional coherence hypothesis: cross-linguistic evidence in language acquisition. Cognitive psychology 55:259–305. doi: 10.1016/j.cogpsych.2006.12.001

Mountjoy JD, Lemon RE (1995) Extended song learning in wild European starlings. Animal Behaviour 49:357–366. doi: 10.1006/anbe.1995.0048

Ohms VR, Gill A, Van Heijningen C. (2010) Zebra finches exhibit speaker-independent phonetic perception of human speech. Proceedings Biological sciences / The Royal Society 277:1003–9. doi: 10.1098/rspb.2009.1788

Patel AD, Gibson E, Ratner J, Besson M (1998) Processing syntactic relations in language and music: An event-related potential study. Journal of Cognitive …

Ramus F, Hauser MD, Cory M. (2000) Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys. Science 288:349–351. doi: 10.1126/science.288.5464.349

Saffran J, Aslin R, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274:1926–1928.

Samuel a. G (2001) Knowing a Word Affects the Fundamental Perception of The Sounds Within it. Psychological Science 12:348–351. doi: 10.1111/1467-9280.00364

Shanks D, Darby R (1998) Feature-and rule-based generalization in human associative learning. Journal of Experimental Psychology: … 24:405–415.

Shi R, Morgan JL, Allopenna P (1998) Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. Journal of child language 25:169–201.

Wright AA (1998) Auditory List Memory in Rhesus Monkeys. Psychological Science 9:91–98. doi: 10.1111/1467-9280.00017

CHAPTER SEVEN

Temporal and syntactic processing in songbirds

Introduction

Language is full of patterns. Yet, we know little about how the brain perceives, organizes and uses temporally patterned information. We know even less about how these basic neurobiological processes relate to speech acquisition, recognition, production and ultimately to language (Mauk and Buonomano 2004). While much can be learned about the neurobiology of language using non-invasive neuroscience techniques in humans, such as functional magnetic resonance imaging and electroencephalography (Poeppel et al. 2012), these techniques cannot answer how individual neurons and neural circuits contribute to language-relevant temporal pattern processing. Development of biologically plausible, neuron-level computations for behaviors supporting language, such as pattern learning and generalization, requires animal models that can master these aspects of temporal processing (Hockett 1960; Kiggins et al. 2012).

Despite the need for language-relevant model systems, progress has stumbled over an enigmatic question of *which* patterns provide insights into the neurobiology of language? In recent years, under the strong influence of computational linguistics (O'Donnell et al. 2005), the search for fruitful model systems has become preoccupied with whether or not other species are capable of learning equivalents of highly elaborate patterns found in language (e.g., *mildly context-sensitive language* constructions, such as "John Mary Peter Jane lets help teach swim" or *context-free language* constructions like "the starling the cats want was tired" see (Berwick et al. 2011)). As a result, many comparative psychologists have focused on pattern complexity, using formal mathematical grammars, as the key to organizing language-relevant cognitive processes across species (O'Donnell et al. 2005; Berwick et al. 2011; Beckers et al. 2012). This focus on mathematically formal grammars is at odds, however, with limitations imposed by basic organismal biology through, for example, working memory or lung capacity (Christiansen and Chater 2008; Chater et al. 2009; Frank et al. 2012). Remembering that “the map is not the territory” (Korzybski 1933), we suggest that formal models of syntactic complexity, however useful in some domains, provide inappropriate targets for the neurobiological implementation of language from either an evolutionary (Margoliash and Nusbaum 2009) or basic neuroscience (Bloomfield et al. 2011; Frank et al. 2012) perspective.

We wish to move the study of temporal pattern processing in animals, and its connection to language, beyond questions tied to idealized formal grammars. We suggest instead that comparative efforts seek to understand how well-defined pattern knowledge persists across multiple levels of abstraction, an approach that resonates more closely with modern cognitive psychology and neuroscience (Penn et al. 2008; Penn 2009). Our review highlights recent results supporting two central ideas: (1) songbirds are able to learn and generalize pattern knowledge and (2) the level of sensory abstraction at which the learned pattern is implemented is as crucial to generalization as the form of the pattern. These constraints on the generalization of pattern knowledge provide important insights for how multiple cognitive systems contribute to supporting temporal pattern recognition and, therefore, components of language.

Songbirds can learn patterning rules

Learning patterned sequences across time appears to be an important aspect of songbird biology. Certain species of songbird, such as the European starling (*Sturnus vulgaris*), produce songs with a clear acoustic hierarchical organization. The most basic level of starling song is referred to as a note, and separate notes are identified based on the contiguous presence of power in their spectrotemporal representations (Meliza et al. 2010; Meliza 2011). The arrangement of notes in starling song is structured into repeatable patterns called motifs, that usually last between 200 and 1000ms (Gentner and Hulse 2000). In addition, while motif repertoires between starlings are largely unique (Eens and Pinxten 1992), motifs can nonetheless be broadly classified into four classes based on spectrotemporal features (whistles, warbles, rattles and high frequency events) shared across all starlings (Adret-Hausberger and Jenkins 1988). The songs of adult starlings unfold in time as non-random sequences of changing motifs (Eens et al. 1988; Gentner and Hulse 1998), where each motif represents a dynamic, stereotyped, auditory event (see Figure 1). The many layers of temporal structure found in starling song (Eens et al. 1988; Gentner 2008), has led researchers to postulate that the temporal patterning of motifs within song plays a crucial role in many important social behaviors, including individual recognition (Gentner 2007) and mate selection (Knudsen and Gentner 2010).

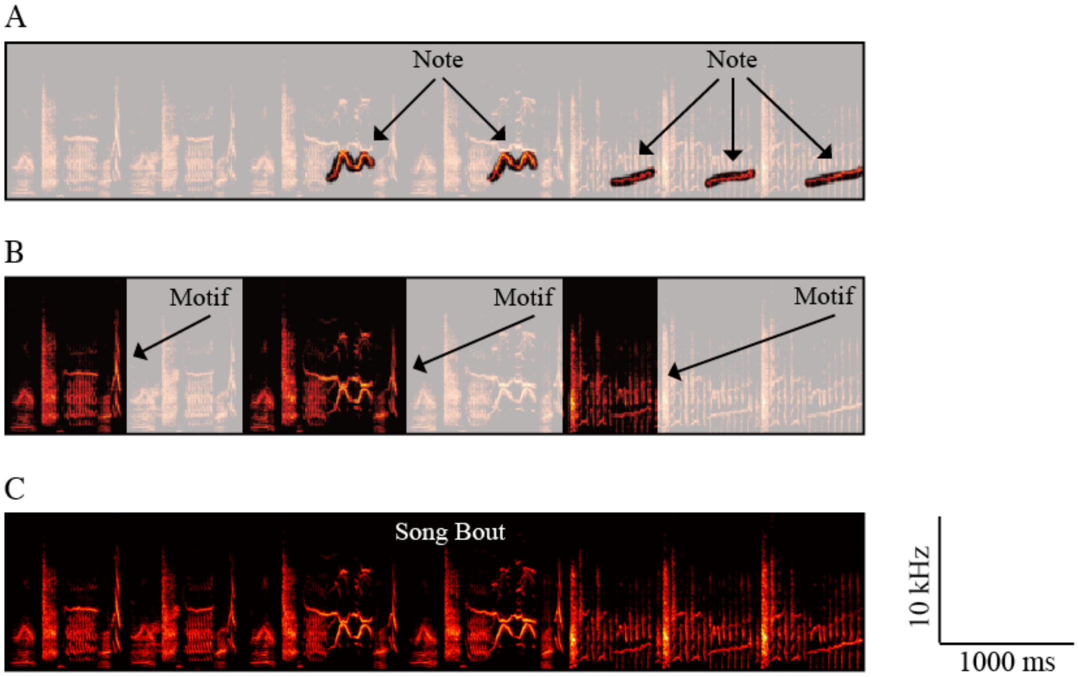

Figure 7.1. Segment of male starling song used to demonstrate the hierarchical temporal scales of acoustic organization. (A) The most fundamental component of songs are notes – bursts of continuous energy throughout their spectrotemporal representations. (B) These notes are clustered into stereotyped packages to form motifs. (C) These motifs occur in a specific temporal pattern across multiple renditions of a given starling's song.

Laboratory studies confirm that starlings can use the natural patterning of motifs in song to successfully recognize individuals. In one such experiment, starlings were trained to recognize songs from different conspecific males, and successfully generalize this learned individual vocal recognition to novel songs from the familiar training singers (Gentner and Hulse 1998). When presented with novel songs from familiar singers had randomly ordered motifs, however, recognition performance was significantly impaired compared to songs in which the natural motif patterning was preserved, suggesting that starlings are sensitive to the patterning of motifs in song (Gentner 2007). More recent studies have directly examined the sensitivity of starlings to temporal patterning of motifs in the absence of additional acoustic cues. This is important because the motif repertoires of different singers are largely unique (Eens et al. 1988), and thus the songs produced by each bird differ in both the ordering of motifs and the spectrotemporal features comprising each motif (Gentner 2007). Acoustic differences between motifs are not required, however, as starlings can readily learn to dissociate motif sequences comprised of *exactly* the same motifs based solely on their patterning (Comins and Gentner 2010).

The foregoing behaviors provide an interesting context within which to investigate neurobiological mechanisms of temporal pattern processing, particularly statistical learning (Saffran et al. 1996; Gómez and Gerken 2000). Importantly, the patterns used in these studies were differentiable by the transition probabilities between specific motifs. This explicit tie between pattern and elements means that any patterning knowledge acquired from these sequences cannot generalize beyond the training stimuli. The ability to generalize pattern knowledge, however, is central to human development and language acquisition (Marcus et al. 1999; Marcus 2000; Marcus et al. 2007). Insofar as animals show similar pattern learning and generalization, comparative neurobiological mechanisms may be profitably studied. To differentiate this important point, we refer to generalizable pattern knowledge as a *rule* (Pothos 2005; Pothos 2007; Bregman and Gentner 2010; Knudsen and Gentner 2010). In keeping with the definition put forth by Marcus (Marcus 2005), rules provide a metric of similarity between two patterns. Importantly, constraints on the generalizability (or measures of similarity) of any rule reflect the level of abstraction at which the rule is operating, from explicit elements to perceptual or functional categories. Understanding the neurobiological implementation of patterning rules requires careful articulation of both the rule and the corresponding level of sensory abstraction.

Several lines of evidence support of the notion that songbirds (starlings) can generalize learned pattern knowledge to sets of familiar auditory elements (motifs). We trained starlings to classify sequences of motifs that subscribe to forms $(AB)^n$ and A^nB^n (where $2 \leq n \leq 4$; (Gentner et al. 2006)). Here, A and B each represent a set of 8 motifs from perceptually distinct classes of starling song motif known as warbles and rattles, respectively. This experiment demonstrated that starlings were able to classify patterns without relying on numerous alternative solutions (e.g., A^1B^3, A^1B^3, A^2B^3, A^3B^2, AAAA, BBBB, ABBA, BAAB). In addition, once starlings have acquired knowledge of these patterns based on training with 16 sequences, classification can be transferred to other small sets of novel sequences of the same warbles and rattles that follow the learned patterns (Gentner et al. 2006). Thus, starlings can generalize learned patterns over sets of familiar elements.

This study of pattern learning in starlings (Gentner et al. 2006), as well as similar results in Bengalese finches (Abe and Watanabe 2011), have drawn criticism, based on the concern that subjects might have employed some *simple* strategy (e.g., primacy, recency, “phonetic generalization”) to solve the task (Corballis 2007; van Heijningen et al. 2009; Beckers et al. 2012; Ten Cate and Okanoya 2012). To address these concerns, a recent study trained a small set of starlings on a considerably more difficult task. In this task, starlings were trained to recognize the patterning forms XXYY and XYXY, where on any given trial an X could represent a motif from *either* set A or set B (and vice versa for Y) and therefore subjects needed to distinguish AABB and BBAA patterns from ABAB and BABA. This stimulus design removes the possibility that subjects solve the task by anticipating the presence of certain acoustic cues in specific locations of the pattern, because all motifs are equally likely to occur in every position of the sequence. In addition, instead of training subjects on just a subset of all possible patterned sequences (as in (Gentner et al. 2006), where only 16 out of 8,092 possible sequences were used in training), nearly all possible sequences (at least 16,300 out of 16,384) were used during training, and no single sequence was presented more than twice. This procedure provided two distinct advantages for assessing songbird pattern learning abilities: (1) it renders potentially trivial solution strategies useless (e.g., no specific motif or category of motif occurs more frequently in any position), and (2) it precludes identification of serendipitous acoustic cues that might be present in the smaller sets of training and testing sequences (Ten Cate and Okanoya 2012) used in early work (Gentner et al. 2006). Learning to classify XXYY and XYXY patterns correctly is only possible by recognizing differences in the temporal structures governing the organization of motifs (Comins and Gentner 2014).

In addition to demonstrating that starlings are able to learn motif patterning rules from trial-unique exemplars, our recent work (Comins and Gentner 2014) adds further insight into the nature of the patterning rules themselves. More detailed analyses of the starlings' behavior in response to select subsets of the training stimuli indicate that performance did not rely on simply comparing which motif or pairs of motifs occurred in given locations of the pattern. Rather, the starlings appear to be accumulating evidence from at least 3 (if not all 4) of the motifs in the sequence before making a decision about which pattern a sequence follows. In the future, more work will need to be done in order to identify the patterning rules acquired in this study; regardless, the evidence firmly shows that songbirds can generalize learned temporal relationships between multiple items in a sequence.

Constraints on pattern generalization

Although considerable debate exists (and will likely continue) regarding the complexity of various rules that birds and other non-human animals can learn (Berwick et al. 2013), it is nonetheless clear that at least some rules are learnable. But all rules are defined over some context, and because the rules we are interested in understanding are defined over sensory events, it is crucial to understand both the nature of the relevant sensory events and their interactions (if any) with the rules in question. To explore these issues we again trained starlings to recognize differently patterned strings of motifs, but instead of focusing on the overtly reinforced patterns, we asked how the perceptual organization of the pattern elements contributes to or constrains learning and generalization (Comins and Gentner 2013).

We trained two groups of starlings to distinguish patterns of the form XXYY from XYXY, where on each trial X and Y represented a motif from either set A or B. Thus, the subject needed to distinguish AABB and BBAA patterns from ABAB and BABA patterns; A and B were defined sets of motifs. Following training on a small subset of all possible XXYY and XYXY sequences, we transferred subjects to 500 novel sequences built from the same sets of A and B motifs. Crucially, for one group of starlings, the motif sets (A and B) preserved natural, perceptual category boundaries (warbles and rattles; see (Eens et al. 1988; Braaten 2000)). For the other group, however, the same motifs were pseudo-randomly assigned membership into sets A and B (Comins and Gentner 2013).

This seemingly subtle alteration of the way that pattern elements were organized had profound effects. Both groups of starlings were able to learn the XXYY and XYXY patterns regardless of whether they were implemented over sets of perceptually organized motifs or pseudo-randomly organized motifs. Surprisingly, however, only the birds trained with motif sets that followed the natural boundaries were able to generalize pattern knowledge to novel motif sequences. These results suggest that pattern learning may be agnostic to the perceptual structure of underlying elements, but these same perceptual differences tightly constrain the generalization of pattern knowledge.

It is likely that similar constraints regulate pattern generalization at multiple levels of sensory abstraction. In the foregoing examples, it is clear that pattern knowledge is implemented over perceptually well-defined sets of motifs. Definition of these sets likely reflects a combination of our operant training and the animals' prior experiences with different motifs in the context of other song-driven behaviors. One question deserving of future attention is the extent to which acquired pattern knowledge can generalize to sequences of novel elements. Our results suggest that such generalization is possible, but that it will be tightly constrained by how these elements fall into the same open-ended perceptual or functional categories that are used to carry the patterning information during training. This behavioral interaction between pattern knowledge and element abstraction (i.e. categorization) has important implications for understanding the neurobiological basis of temporal pattern knowledge and thus language.

Neurophysiology

We assert that the same close ties to sensory encoding of elements evidenced in the foregoing behavioral results, will also hold for a neurobiological understanding of temporal pattern processing, and ultimately the neurobiological basis of language. This does not imply that language relevant processes must necessarily be restricted to specific sensory modalities, but only that patterns which convey information in language are necessarily carried (and constrained biologically) by *some* sensory modality. The sensory regions that encode pattern elements, therefore, provide the best available targets to study the neurobiology of temporal pattern processing with cellular and circuit-level precision.

The temporal pattern elements of starling songs are motifs, and motifs are preferentially encoded in the higher-level regions of the auditory telencephalon. The songbird auditory system (Figure 2a) follows the general vertebrate plan {Carr, 1992 #751}. Field L2a is the primary telencephalic target of the auditory thalamus, nucleus ovoidalis (Ov){Karten, 1968, p05432}, and is the input layer for a columnar circuit, homologous to mammalian primary auditory cortex {Wang, 2010 #623}{Dugas-Ford, 2012 #1435}, that includes L1, L3, and caudal mesopallium, CM (Fig. 2a). Field L subregions also project to the caudomedial nidopallium, NCM. The NCM and lateral CM (CLM), in turn, share reciprocal connections with the medial CM (CMM).

Encoding of songs in the songbird forebrain mirrors the coarse, hierarchical increase in selectivity observed for complex signals in mammalian visual (Maunsell and Newsome, 1987; Rust and DiCarlo, 2010) and auditory (Kikuchi et al., 2010) cortices. Neurons throughout the songbird auditory forebrain show selectivity to species-specific vocalizations {Bonke, 1979 #477;Leppelsack, 1976 #632;Muller, 1985 #474}, that generally increases from Field L2, to L1 and L3{Theunissen, 2004 #789;Theunissen, 1998 #790;Theunissen, 2000 #740}, and then NCM and CM{Muller, 1985 #474;Grace, 2003 #456;Bonke, 1979 #477;Leppelsack, 1976 #632;Muller, 1985 #474;Gentner, 2003, p02120;Gentner, 2004 #475;Thompson, 2010 #793;Jeanne, 2011 #631}. Although the detailed circuitry within and between field L, NCM and CM remains to be determined, the increasing selectivity between regions is consistent with a functional hierarchy that is tuned throughout to conspecific song{Hsu, 2004 #472;Woolley, 2005 #473}, and further refined by experience{Gentner, 2003, p02120;Sockman, 2002 #633;Sockman, 2005 #1439;Phan, 2006 #678;Thompson, 2010 #793;Jeanne, 2011 #631}.

Neural encoding in NCM and CM in particular is shaped strongly by both the stimulus acoustics and the learned behavioral relevance tied to different motifs {Gentner, 2003, p02120; Thompson, 2010 #793;Jeanne, 2011 #631}. In NCM, response modulation to learned stimuli appears as a rapid and long-lasting stimulus-specific response weakening to more behaviorally-relevant (i.e., familiar vs unfamiliar) stimuli (Chew et al. 1995; Thompson and Gentner 2010), and effect that emerges in part from stimulus specific response modulation through local inhibition (Thompson et al. 2012). Neurons in both CMM and CLM appear encode the acoustic features that differentiate motifs. That is, the spiking patterns of neurons in both regions convey significant amounts of information about motif identity, and this stimulus specific information is greater in medial CM than in lateral CM {Jeanne, 2011 #631}. The stimulus specific responses in both CMM and CLM are modulated by learning, such that in both regions familiar motifs drive stronger average spike rates than unfamiliar (or novel) motifs {Gentner, 2003, p02120; Jeanne, 2011 #631}.

To understand the relationship between neural encoding of stimulus features and familiarity in more detail, Jeanne et al. (2013) recorded from CM in birds that had been trained behavioral so that a subset of familiar motifs carried task-relevant information, while another subset of familiar motifs carried task-irrelevant information. This design ruled out the possibility that reward alone modulates CM responses, and instead requires that any differences between responses evoked by task-relevant and task-irrelevant stimuli are tied to information about the specific behaviors conveyed (or not) by each motif. Consistent with prior studies, Jeanne et al (2013) observed a strong effect of stimulus familiarity in CMM. That is, both the task-relevant and task-irrelevant motifs drove strong responses in CMM, compared to novel motifs. In contrast, however, only the task-relevant motifs evoked responses significantly greater that novel motifs in CLM.

To investigate how motif acoustics might carry information about task-relevant behaviors, Jeanne et al. (2013) also examined population activity in the broad spiking (putative efferent projection) neurons in CLM. Typically, when neurons are similarly tuned (i.e. they have a positive signal correlation), their responses also tend to covary trial-by-trial (i.e. they have a positive noise correlation) (Cohen and Maunsell, 2009; Cohen and Newsome, 2008; Gu et al., 2011; Gutnisky and Dragoi, 2008; Hofer et al., 2011; Kohn and Smith, 2005). This is thought to result because the common inputs that convey similar tunings also share noise. Surprisingly, however, Jeanne et al (2013) found that learning can invert this canonical relationship, leading to a more orthogonal relationship between the signal and noise correlations but only for task-relevant motifs (Figure 2b). In theory (Oram et al., 1998), negative relationships can improve population coding because common noise among dissimilarly tuned neurons can be subtracted, which strengthens the signal while dissipating the noise. Indeed, the task-relevant motifs are encoded in the CLM projection neurons with higher fidelity (Jeanne et al 2013). The improved encoding of task-relevant motifs in CLM efferents suggest that this region may provide a convergence zone for the contextual modulation of high-level auditory representations in which auditory signals may be “packaged” into population responses that can efficiently drive behaviors. One simple circuit whereby signal and noise correlations could be uncoupled in a stimulus specific way is shown in Figure 2c. In this model, shared excitatory inputs driven by stimulus A and B, are combined with stimulus-specific, independent inhibition, that serves to dampen the average response of a neuron while preserving the overall noise correlation between trial-to-trial spiking activity. Understanding whether this kind of inhibition is actually present, and if so what its source (or sources) might be, is a topic of ongoing research. In any case, this basic circuit may serve other processes such as attention as well.

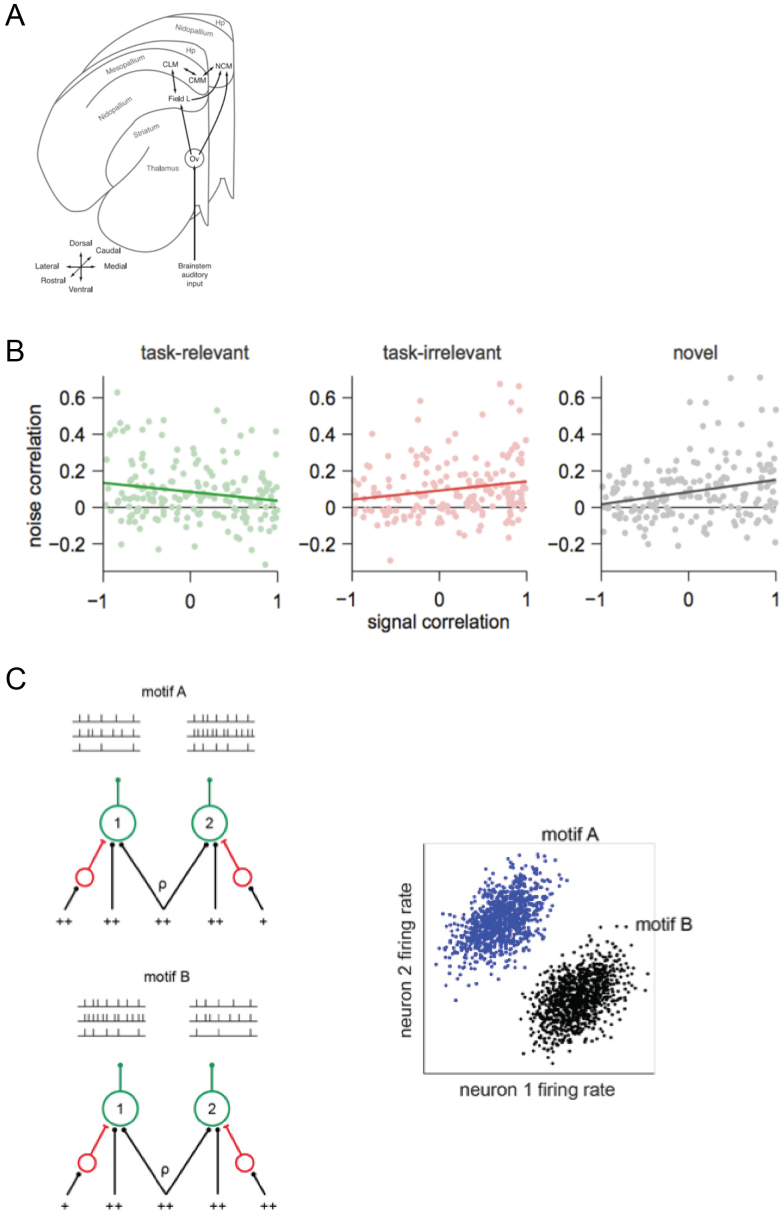

Figure 7.2. (A) Schematic of the avian auditory system. Reprinted from Jeanne et al (2011). (B) Plots showing the signal and noise correlations between pairs of neurons in response to task-relevant, task-irrelevant and novel motifs. Regression lines are indicated by color lines. Reprinted from Jeanne et al (2013).(C) Circuit model of how strong positive noise correlations and negative signal correlations might arise between two neurons that respond to two different stimuli, motif A and B.

Considerable debate on the evolution and neurobiology of language has focused on identifying *which* complex patterns are learnable across taxa. Here, we propose that comparative approaches to biolinguistics move beyond mapping animal pattern expertise for mathematical models of grammatical structures in language (Berwick et al. 2011; Berwick et al. 2012; Beckers et al. 2012). Setting aside the notion of pattern 'complexity', we can begin to ask how the brain encodes *any* patterning rule. Characterizing the nature of the (probably many) patterning rules that animals can use is important, and we argue that the space of possible patterning rules is likely to be much more rich that currently appreciated. We therefore caution against the kinds of dismissals proffered by some researchers that the nature of patterning rule representation in animals is "stimulus-bound" and therefore falls short of the kinds of abstract rule learning found in humans (ten Cate, *this issue*). While we agree that humans are capable of reasoning about temporal relations and patterns in more abstract way than non-human animals (Penn et al. 2008; Kiggins et al. 2012), there is probably nothing "simple" about the use of patterns in non-humans or their neurobiological implementation. Indeed, despite the fact that humans are capable of abstract reasoning (Gentner and Namy 2006), there is considerable evidence that sound-based (i.e., perceptual) features of words correlate with more abstract features of language, such as case-endings and gender-markers (Kelly 1992; Frigo and McDonald 1998; Gómez and Gerken 2000; Monaghan et al. 2007)). This underlying perceptual structure might provide important scaffolding for acquiring more abstract knowledge of language during development. The kind of category/rule interactions we describe above and their neural coding may provide an attractive model for these processes. A complete neurobiology of language cannot rest entirely on non-human models, but until we understand why non-human models are insufficient they have much to offer.

Acknowledgements

This chapter, in full, is currently material under review. Comins, Jordan A.; Gentner, Timothy Q. The dissertation author was the primary investigator and author of this paper.

References

Abe K, Watanabe D (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. Nature Neuroscience 14:1067–1074. doi: 10.1038/nn.2869

Adret-Hausberger M, Jenkins P (1988) Complex Organization of the Warbling Song in the European Starling Sturnus vulgaris. Behaviour 107:138–156.

Beckers GJL, Bolhuis JJ, Okanoya K, Berwick RC (2012) Birdsong neurolinguistics: songbird context-free grammar claim is premature. Neuroreport 23:139–45. doi: 10.1097/WNR.0b013e32834f1765

Berwick R, Okanoya K, Beckers G, Bolhuis J (2011) Songs to syntax: the linguistics of birdsong. Trends in Cognitive Sciences 15:113–121.

Berwick RC, Beckers GJL, Okanoya K, Bolhuis JJ (2012) A Bird's Eye View of Human Language Evolution. Frontiers in evolutionary neuroscience 4:5. doi: 10.3389/fnevo.2012.00005

Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ (2013) Evolution, brain, and the nature of language. Trends in cognitive sciences 17:89–98. doi: 10.1016/j.tics.2012.12.002

Bloomfield TC, Gentner TQ, Margoliash D (2011) What birds have to say about language. Nature neuroscience 14:947–8. doi: 10.1038/nn.2884

Braaten RF (2000) Multiple levels of representation of song by European starlings (Sturnus vulgaris): Open-ended categorization of starling sing types and differential forgetting of song categories and exemplars. Journal of Comparative Psychology 114:61–72. doi: 10.1037//0735-7036.114.1.61

Bregman M, Gentner TQ (2010) Syntactically Complex Vocal Systems. In: MD B, J M (eds) Encyclopedia of Animal Behavior. Oxford Academic Press, pp 368–374

Chater N, Reali F, Christiansen MH (2009) Restrictions on biological adaptation in language evolution. Proceedings of the National Academy of Sciences of the United States of America 106:1015–20. doi: 10.1073/pnas.0807191106

Chew SJ, Mello C, Nottebohm F. (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. Proceedings of the National Academy of Sciences of the United States of America 92:3406–10.

Christiansen MH, Chater N (2008) Language as shaped by the brain. The Behavioral and brain sciences 31:489–508; discussion 509–58. doi: 10.1017/S0140525X08004998

Comins JA, Gentner TQ (2014) Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. Animal cognition. doi: 10.1007/s10071-014-0732-5

Comins JA, Gentner TQ (2010) Working memory for patterned sequences of auditory objects in a songbird. Cognition 117:38–53.

Comins JA, Gentner TQ (2013) Perceptual categories enable pattern generalization in songbirds. Cognition 128:113–118. doi: 10.1016/j.cognition.2013.03.014

Corballis MC (2007) Recursion, Language, and Starlings. Cognitive science 31:697–704.

Eens M, Pinxten R (1992) No overlap in song repertoire size between yearling and older starlings Sturnus vulgaris. Ibis 134:72–76.

Eens M, Pinxten R, Verheyen RF (1988) Temporal and sequential organisation of song bouts in the starling. Ardea 77:75–86.

Frank SL, Bod R, Christiansen MH (2012) How hierarchical is language use? Proceedings Biological sciences / The Royal Society 279:4522–31. doi: 10.1098/rspb.2012.1741

Frigo L, McDonald JL (1998) Properties of Phonological Markers That Affect the Acquisition of Gender-Like Subclasses. Journal of Memory and Language 39:218–245. doi: 10.1006/jmla.1998.2569

Gentner D, Namy LL (2006) Analogical Processes in Language Learning. Current Directions in Psychological Science 15:297–301. doi: 10.1111/j.1467-8721.2006.00456.x

Gentner T (2007) Mechanisms of Temporal Auditory Pattern Recognition in Songbirds. Language learning and development 3:157–178.

Gentner T, Hulse S (1998) Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Animal Behaviour 56:579–594.

Gentner TQ (2008) Temporal scales of auditory objects underlying birdsong vocal recognition. The Journal of the Acoustical Society of America 124:1350–1359.

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440:1204–1207.

Gentner TQ, Hulse SH (2000) Perceptual classification based on the component structure of song in European starlings. The Journal of the Acoustical Society of America 107:3369–3381.

Gómez R, Gerken L (2000) Infant artificial language learning and language acquisition. Trends in Cognitive Sciences 4:178–186.

van Heijningen CAA, de Visser J, Zuidema W, ten Cate C (2009) Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proceedings Of The National Academy Of Sciences Of The United States Of America 106:20538–20543.

Hockett CF (1960) The origin of speech. Scientific American 203:88–111. doi: 10.1126/science.303.5662.1316

Kelly MH (1992) Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. Psychological Review 99:349–64.

Kiggins JT, Comins JA, Gentner TQ (2012) Targets for a Comparative Neurobiology of Language. Frontiers in Evolutionary Neuroscience 4:1–13. doi: 10.3389/fnevo.2012.00006

Knudsen DP, Gentner TQ (2010) Mechanisms of song perception in oscine birds. Brain and language 115:59–68.

Korzybski A (1933) A non-Aristotelian system and its necessity for rigour in mathematics and physics. Science and Sanity 1:747–761.

Marcus G (2005) Opposites detract: Why rules and similarity should not be viewed as opposite ends of a continuum. Behavioral and Brain Sciences 28:28–29. doi: 10.1017/S0140525X05380010

Marcus G, Vijayan S, Brandi Rao S, Vishton PM (1999) Rule Learning by Seven-Month-Old Infants. Science 283:77–80.

Marcus GF (2000) Pabiku and Ga Ti Ga: Two Mechanisms Infants Use to Learn About the World. Current Directions in Psychological Science 9:145–147. doi: 10.1111/1467-8721.00080

Marcus GF, Fernandes KJ, Johnson SP (2007) Infant rule learning facilitated by speech. Psychological science 18:387–91. doi: 10.1111/j.1467-9280.2007.01910.x

Margoliash D, Nusbaum HC (2009) Language: the perspective from organismal biology. Trends In Cognitive Sciences 13:505–510.

Mauk MD, Buonomano DV (2004) The neural basis of temporal processing. Annual review of neuroscience 27:307–40. doi: 10.1146/annurev.neuro.27.070203.144247

Meliza CD (2011) Effects of auditory recognition learning on the perception of vocal features in European starlings (Sturnus vulgaris). The Journal of the Acoustical Society of America 130:3115–23. doi: 10.1121/1.3641420

Meliza CD, Chi Z, Margoliash D (2010) Representations of Conspecific Song by Starling Secondary Forebrain Auditory Neurons: Toward a Hierarchical Framework. Journal Of Neurophysiology 103:1195–1208.

Monaghan P, Christiansen MH, Chater N (2007) The phonological-distributional coherence hypothesis: cross-linguistic evidence in language acquisition. Cognitive psychology 55:259–305. doi: 10.1016/j.cogpsych.2006.12.001

O'Donnell TJ, Hauser MD, Fitch WT (2005) Using mathematical models of language experimentally. Trends in cognitive sciences 9:284–9. doi: 10.1016/j.tics.2005.04.011

Penn D (2009) On becoming approximately rational: The relational reinterpretation hypothesis. Rational Animals, Irrational …

Penn DC, Holyoak KJ, Povinelli DJ (2008) Darwin's mistake: explaining the discontinuity between human and nonhuman minds. Behavioral and Brain Sciences 31:109–30. doi: 10.1017/S0140525X08003543

Poeppel D, Emmorey K, Hickok G, Pylkkänen L (2012) Towards a new neurobiology of language. The Journal of Neuroscience 32:14125–31. doi: 10.1523/JNEUROSCI.3244-12.2012

Pothos E (2005) Opposites detract: Why rules and similarity should not be viewed as opposite ends of a continuum. Behavioral And Brain Sciences

Pothos EM (2007) Theories of artificial grammar learning. Psychological bulletin 133:227–44. doi: 10.1037/0033-2909.133.2.227

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. Science (New York, NY) 274:1926–8.

Ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Philosophical transactions of the Royal Society of London Series B, Biological sciences 367:1984–94. doi: 10.1098/rstb.2012.0055

Thompson JV, Gentner TQ (2010) Song Recognition Learning and Stimulus-Specific Weakening of Neural Responses in the Avian Auditory Forebrain. Journal Of Neurophysiology 103:1785–1797.

Thompson JV, Jeanne JM, Gentner TQ (2012) Local inhibition modulates learning-dependent song encoding in the songbird auditory cortex. Journal of neurophysiology 109:721–733. doi: 10.1152/jn.00262.2012