Anim Cogn (2014) 17:1023–1030
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 / Published online: 15 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

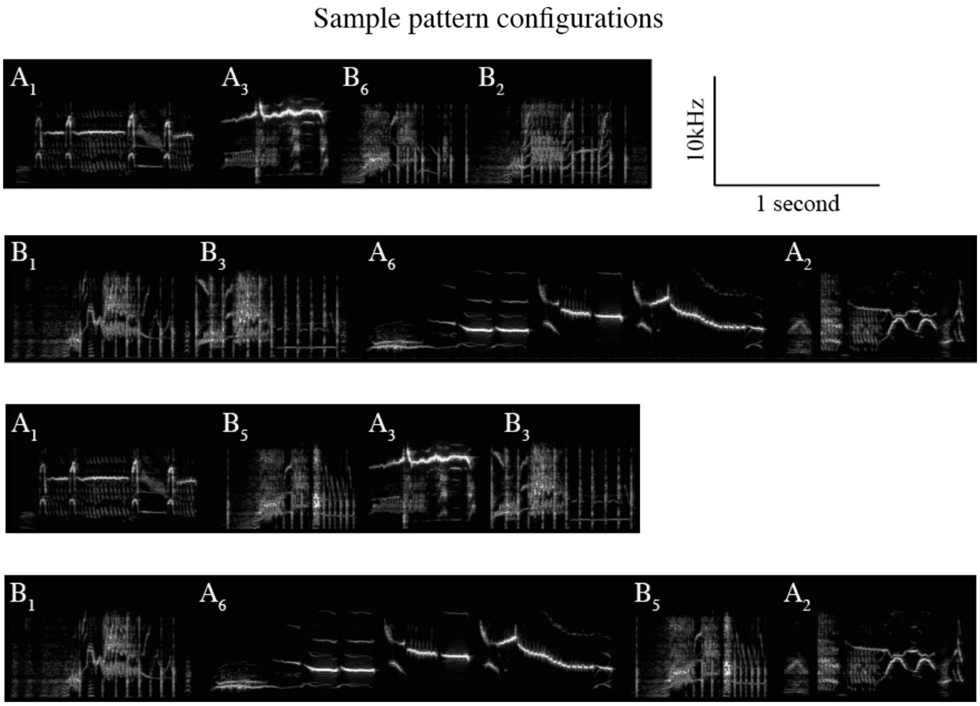


Fig. 1 Example motif sequence stimuli from the AABB and ABAB pattern classes. *Letters* denote motif classes (A: warbles, B: rattles) and *subscript numbers* (1–8) denote unique motifs within each class used to generate patterned sequences. A total of 16,300 (out of 16,384 possible) XXYY and XYXY sequences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

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

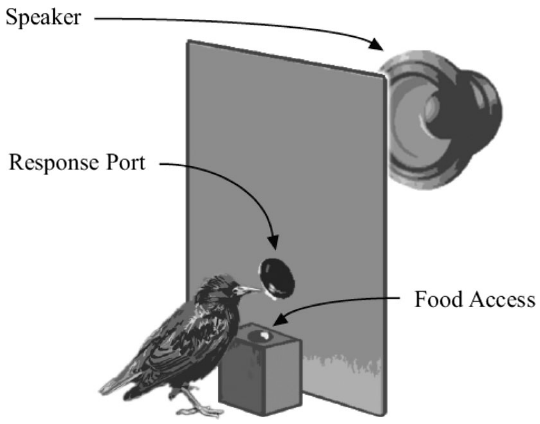


Fig. 2 Schematic of the operant apparatus. Subjects start a trial by pecking the center response port. After the motif pattern stimulus ends, the subject either pecks the center response port again or withholds any response depending on the class from which the stimulus was drawn. Correct responses yield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

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 ($\mu \pm$ 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.

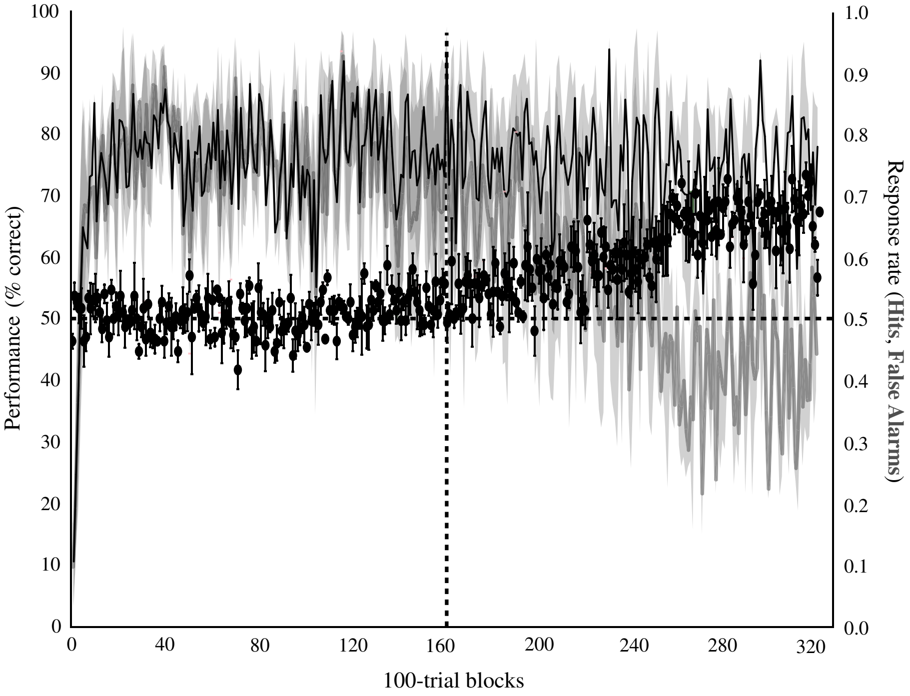
Results

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.

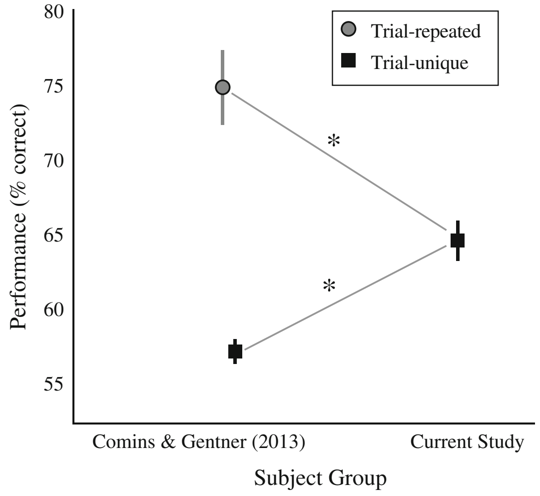


Fig. 4 Mean (±SE) accuracy (% correct) for the classification of the final 500 trial-unique (*black squares*) XXYY/XYXY patterned stimuli for subjects in the current study (*right*) and previous work (*left*) where subjects were trained to a high level of performance with a small set of trail-repeated (*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