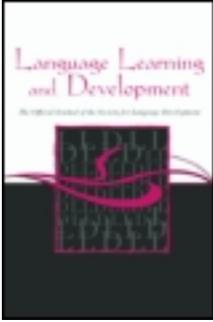


This article was downloaded by: [University of California, San Diego]
On: 28 February 2013, At: 23:27
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



Language Learning and Development

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hlld20>

Mechanisms of Temporal Auditory Pattern Recognition in Songbirds

Timothy Q. Gentner^a

^a University of California, San Diego

Version of record first published: 05 Dec 2007.

To cite this article: Timothy Q. Gentner (2007): Mechanisms of Temporal Auditory Pattern Recognition in Songbirds, *Language Learning and Development*, 3:2, 157-178

To link to this article: <http://dx.doi.org/10.1080/15475440701225477>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Mechanisms of Temporal Auditory Pattern Recognition in Songbirds

Timothy Q. Gentner
University of California, San Diego

Human speech and birdsong share several important features. Both communication systems entail large, acoustically rich repertoires of temporally patterned vocal signals, and both must be learned early in ontogeny (Kuhl, 2003, 2004; Marler, 1970, 1975). Over the last 20 years, these similarities have helped to establish birdsong as an important model system for understanding the neurobiological bases of vocal learning and production (Doupe & Kuhl, 1999; Margoliash, 2002; Zeigeler, 2004). But communication is not a solo endeavor, and the target of this elaborate vocal-motor production system is not the song itself, but rather the behavior of nearby conspecifics (Kroodsma & Miller, 1996). Birds sing to be heard. Thus, birdsong also provides an important model system for the processing of complex acoustic communication signals (Theunissen & Shaevitz, 2006; Zeigeler, 2004).

The prototypical temperate zone songbird sings to modulate aggressive territorial encounters among neighboring male songbirds and to attract prospective female mates. For those on the receiving end of song, these broad (though not exhaustive) evolutionary functions imply a host of auditory perceptual and cognitive processes that must be engaged (Gentner & Margoliash, 2002). Although these processes are interesting in their own regard, recent advances (Gentner, Fenn, Margoliash, & Nusbaum, 2006) suggest that some of the acoustic pattern recognition abilities of songbirds may inform current debates on the evolution of language-relevant representational strategies and syntactic skills. This review describes those recent advances in the context of a larger endeavor to understand the biological basis of temporal pattern recognition in one species of songbird, European starlings.

Correspondence should be addressed to Timothy Q. Gentner, Department of Psychology, University of California at San Diego, 9500 Gilman Drive MC0109, La Jolla, CA 92093-0109. E-mail:tgentner@ucsd.edu

We begin by discussing the behavioral and neural basis for song recognition in starlings. These studies establish that as adult starlings learn to recognize different songs, a rich comprehension of each song's component structure emerges, which is reflected in both their behavior and the responses of single neurons in the auditory forebrain. From this understanding of the functional units of song (or motifs), we then describe the abilities of starlings to attend to temporal sequences of motifs, and to arbitrary rules that describe patterns among classes of motifs. Along the way we note what we consider to be important similarities and differences between birdsong and human speech perception.

BEHAVIORAL MECHANISMS OF INDIVIDUAL VOCAL RECOGNITION

As any birder can attest, vocalizations are a vital diagnostic cue to species identity. Yet songbirds attend to vocal patterns with very fine-scale precision, and their vocal recognition abilities go well beyond simply recognizing members of their own species. Various forms of individual vocal recognition have been observed in nearly every species of songbird studied to date (Stoddard, 1996) and commonly serve as the basis for decisions in more elaborate social behaviors such as female choice (e.g., Lind, Dabelsteen, & McGregor, 1997; O'Loughlen & Beecher, 1997; Wiley, Hatchwell, & Davies, 1991), kin recognition (Beecher, 1991), and male territoriality (Falls & Brooks, 1975; Falls, Krebs, & McGregor, 1982; Godard, 1991; Peek, 1972). Humans are also adept at recognizing familiar talkers by voice alone (Bricker & Pruzansky, 1976). Traditional views of these abilities have separated the acoustic features that carry linguistic content from those that encode talker attributes (e.g., fundamental frequency) and are linguistically irrelevant. More recent data challenge this categorical division of information, however, by suggesting a more integrated perception in which talker recognition can incorporate phonetic aspects of the speech signal as well (Pisoni, 1997; Remez, Fellowes, & Rubin, 1997; Sheffert, Pisoni, Fellowes, & Remez, 2002). As reviewed below, the evidence from songbirds suggests that recognition of individual singers does not rely on idiosyncratic voice cues.

To study individual vocal recognition in starlings, we devised an operational definition that permitted direct laboratory study, whereby the "meaning" of a song can be controlled experimentally. To do this, we use operant conditioning techniques that require subjects to make one response to the songs of a specific bird and a different response to the songs of one or more other birds (Gentner & Hulse, 1998; Hulse, 1995). Typically, the birds are trained to obtain food by pecking buttons on a panel mounted on the side of their cage (Figure 1). They are then reinforced with food for pecking one button, say the left, every time they hear a song from Male A, and for pecking another button, in this case the right, every time they

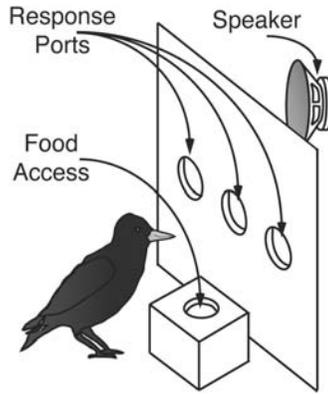


FIGURE 1 Operant apparatus used to demonstrate classification of song stimuli. Subjects start a trial (presentation of a song stimulus selected at random) by pecking the center response port. After the stimulus completes, the subject can either peck at either the left or right response ports depending on the class from which the stimulus was drawn (two-alternative choice tasks), or peck the center port (the “go-nogo” procedure). Correct responses yield food reward. Incorrect responses lead to a short “time out” during which the house light is extinguished and food is inaccessible.

hear a song from Male B. Tasks such as this, in which two sets of stimuli (songs) are associated with similarly reinforced behaviors (peck left/peck right), are called two-alternative choice tasks (2AC). In a close variant, the “go-nogo” (GNG) procedure, behavioral responses to only one set of stimuli are reinforced, leading the subject to cease responding to the nonreinforced stimuli. With both training procedures subjects become proficient at recognizing the songs in each class. Once the basic song recognition is learned we can vary the stimuli in myriad ways to ask questions about the precise acoustic features and associative processes that guide song recognition.

To understand the nature our stimulus manipulations, it’s helpful to know a bit about how starling song is naturally organized. Male starlings tend to sing in long continuous episodes called *bouts*. Song bouts are composed of much smaller acoustic units referred to as *motifs* (Adret-Hausberger & Jenkins, 1988; Eens, Pinxten, & Verhagen, 1991; Figure 2) that, in turn, are composed of still smaller units called *notes*. Notes can be broadly classified by the presence of continuous energy in their spectrotemporal representations. Although a motif may consist of several notes, the note pattern within a motif is usually stereotyped between successive renditions of that motif. Commonly, each motif is repeated two or more times before the next one is sung. Thus, starling song appears (acoustically) as a sequence of changing motifs, where each motif is an acoustically complex event (Figure 2). Different motifs can vary in duration from roughly 200 to 1000 msec and the number of unique motifs that a mature male starling can sing (i.e., his rep-

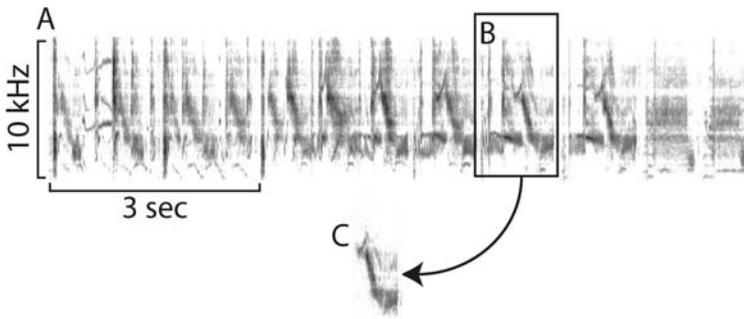


FIGURE 2 Sonogram of starling song segment. Power across the frequency spectrum is shown as a function of time. Darker regions show higher power. Starling song bouts are organized hierarchically. Normal bouts of song can last over 1 min and are comprised of series of repeated motifs. (a) Shows a short sequence of motifs as they might appear in a much longer song bout. A single motif is outlined in (b). Complete song bouts contain many different motifs. Motifs are comprised of stereotyped note patterns. An example of one note is shown in (c).

ertoire size) can exceed 50 or more. Consequently, different song bouts from the same male starling are not necessarily composed of the same set of motifs. Over time, however, the songs of a specific male starling can be characterized by a set of motifs typical of that male starling. Although some sharing of motifs does occur among captive male starlings (Hausberger, 1997; Hausberger & Cousillas, 1995), the motif repertoires of different male starlings living in the wild are generally unique (Adret-Hausberger & Jenkins, 1988; Chaiken, Böhner, & Marler, 1993; Eens, Pinxten, & Verheyen, 1989; Eens et al., 1991; Gentner & Hulse, 1998). Learning which male starlings sing which motifs can therefore provide a diagnostic cue for individual recognition.

The results of our behavioral studies suggest that individual song recognition operates at (or below) the level of the motif. Starlings trained to recognize sets of songs from different conspecific individuals can correctly recognize sets of novel song bouts from the same singers (Gentner & Hulse, 1998; Gentner, Hulse, Bentley, & Ball, 2000). To rule out the use of idiosyncratic source and/or filter properties imparted by individual's vocal apparatus (i.e., "voice" characteristics), we trained birds to recognize isolated motifs shared by two different male starlings (e.g., Bird A singing Motif 1, and Bird B singing Motif 2) and then watched recognition fall to chance for recordings of the same birds singing the opposite motifs (i.e., Bird A singing Motif 2, and Bird B singing Motif 1). These results eliminate a critical role for voice characteristics, suggesting instead the importance of memory for explicit motifs. Consistent with this hypothesis, when we tested the recognition of novel song bouts that have *no* motifs in common with the training songs, performance falls to chance (Figure 3). Data from song sparrows report a similar lack of attention to voice characteristics (Beecher, Campbell, & Burt, 1994). Thus, the in-

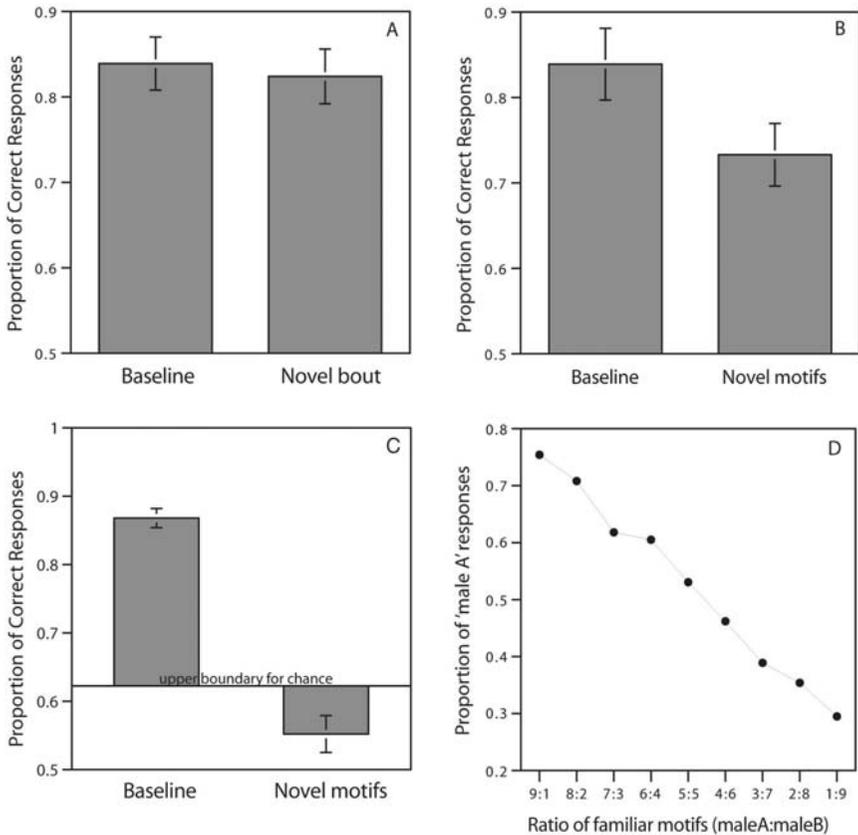


FIGURE 3 Vocal recognition behavior in European starlings. (a) Mean (\pm SEM) proportion of correct responses given during asymptotic performance on an operant recognition task (“Baseline”), and during initial transfer to novel songs containing familiar motifs (“Novel bout”). (b and c) Mean (\pm SEM) proportion of correct responses during transfer from the baseline training to novel songs from the same singers composed of “Novel motifs”; when (b) subjects were exposed to the training and test songs outside of the operant apparatus, or (c) after controlling for all previous song experience where recognition falls below chance. (d) Data showing the close (and approximately linear) relation between the proportions of familiar motifs from two different singers in a song and individual vocal recognition.

corporation of individual identity information into the phonetics of the signal may represent a plesiomorphic state for vocal communication.

If voice characteristics are not used, then it would appear that starlings learn to recognize the songs of other individuals by memorizing sets of motifs associated with individual singers. If true, it should be possible to control song recognition systematically by varying the proportions of motifs in a “target” song that comes

from two vocally familiar male starlings. That is, if subjects memorize a large set of motifs from each singer, recognition behavior should be correlated with the relative proportions of familiar motifs from different male starlings independent of the specific motifs comprising a given song. If they attend to the presence (or absence) of a single motif or a small set of motifs, recognition should not follow relative motif proportions, and should not generalize between songs in which different motifs make up similar proportions. To test these ideas, we again trained starlings to recognize sets of songs from different male starlings and then watched as subjects classified novel song bouts in which motifs from the training songs were combined in different ways (Gentner & Hulse, 2000). Consistent with the motif memorization hypothesis, we observed an approximately linear relation between song classification and the relative proportions of familiar motifs from different singers composing each bout (Figure 3). This suggests that when starlings are compelled to classify conspecific songs, they do so by memorizing large numbers of unique motifs and then organize subsets of these motifs into separate classes. From a human perspective, this might seem a suboptimal strategy for individual vocal recognition. For starlings, however, the solution is functionally parsimonious. Under natural conditions, individual starlings possess unique motif repertoires. Thus, sets of motifs sorted into disjointed perceptual classes, although memory intensive, will correspond to individual identity. Obviously, as the size of the lexicon and sharing between individuals increases, the exclusive use of this strategy for individual vocal recognition becomes increasingly inefficient.

NEURAL CORRELATES TO SONG RECOGNITION

The close relation between starling song motifs and individual vocal recognition suggests the possibility that these functional components may have direct correlates at the neural level. To examine this possibility, we began recording from single neurons in the forebrain of European starlings.

The large-scale architecture and pattern of connectivity within the starling auditory forebrain is shared with other songbird species (Vates, Broome, Mello, & Nottebohm, 1996) and with vertebrates in general (Carr, 1992). The Field L complex (used as a proper noun) is the primary thalamorecipient zone in the auditory telencephalon, and thus is analogous to primary auditory cortex in mammals. Farther along the ascending sensory hierarchy, the caudal-medial nidopallium (NCM) and the lateral and medial regions of the caudal mesopallium (CLM, CMM) are analogous to mammalian secondary auditory cortices. Neurons throughout the starling auditory forebrain show complex patterns of tonotopic organization (Capsius & Leppelsack, 1996; Häusler, 1996; Leppelsack & Schwartzkopff, 1972; Rübsem & Dörrscheidt, 1986), including selectivity to species-specific vocalizations (Bonke, Bonke, & Scheich, 1979; Leppelsack & Vogt, 1976; Müller &

Leppelsack, 1985; Theunissen & Doupe, 1998; Theunissen & Shaevitz, 2006). In other songbirds, the general pattern of increasing response selectivity along the sensory hierarchy (Hsu, Woolley, Fremouw, & Theunissen, 2004; Woolley, Fremouw, Hsu, & Theunissen, 2005) continues into NCM and CMM/CLM (Sen, Theunissen, & Doupe, 2001), suggesting that these regions are involved in the extraction of complex features (Chew, Mello, Nottebohm, Jarvis, & Vicario, 1995; Chew, Vicario, & Nottebohm, 1996; Grace, Amin, Singh, & Theunissen, 2003; Leppelsack, 1983; Sen et al., 2001; Stripling, Volman, & Clayton, 1997). Additional support for the role of NCM and CMM/CLM in the processing of conspecific song comes from studies of stimulus driven expression of the immediate-early-gene (IEG) *zenk*, a putative marker for song-induced experience-dependent plasticity (Jarvis, Mello, & Nottebohm, 1995; Jones et al., 2001; Mello & Clayton, 1995; Mello, Velho, & Pinaud, 2004; Mello, Vicario, & Clayton, 1992; Ribeiro, Cecchi, Magnasco, & Mello, 1998). In starlings, the IEG response in NCM appears tied to stimulus novelty, whereas IEG activity in CMM appears to correlate with the ongoing recognition of familiar songs (Gentner, Hulse, & Ball, 2004).

To examine the role of CMM in the representation of learned conspecific song in adult birds more closely, we trained starlings to recognize two sets of conspecific songs using both the 2AC and GNG operant procedures (Figure 4). Although the bird was anesthetized, we then recorded extra-cellular responses of single neurons in the CMM to an ensemble of acoustic stimuli, including songs used for recognition training (“familiar” songs) and novel conspecific songs (“unfamiliar” songs). To control for biases to any specific song, we varied the stimuli across subjects so that although the stimulus ensemble was similar for each subject, the familiarity or novelty of any given song differed (Gentner & Margoliash, 2003).

As a population, CMM neurons responded selectively (i.e., significantly more strongly) to the class of familiar compared to unfamiliar songs. Moreover, the strong response bias for familiar songs was consistent in animals trained under both the 2AC and GNG procedures (Figure 4).

The apparent plasticity in CMM could be driven solely by exposure to patterns of acoustic variation in the stimulus such that, in the extreme, all the information represented by the cell’s response is present in the acoustic variation of the signal. Alternatively, attention, motivation, and/or reward mechanisms might also shape a cell’s selectivity across time in a manner independent of any particular signal acoustics. Indeed, a significant portion of the response tuning in CMM appears to be under the control of different reinforcement signals present during recognition learning. Whereas the subjects trained using the 2AC procedure showed no reliable difference between response strengths associated with the two sets of training songs, those trained with the S+ and S– classes in the GNG procedure did. For the GNG group, songs associated with positive reinforcement (S+ stimuli) elicited significantly stronger responses than those associated with no reinforcement (S–

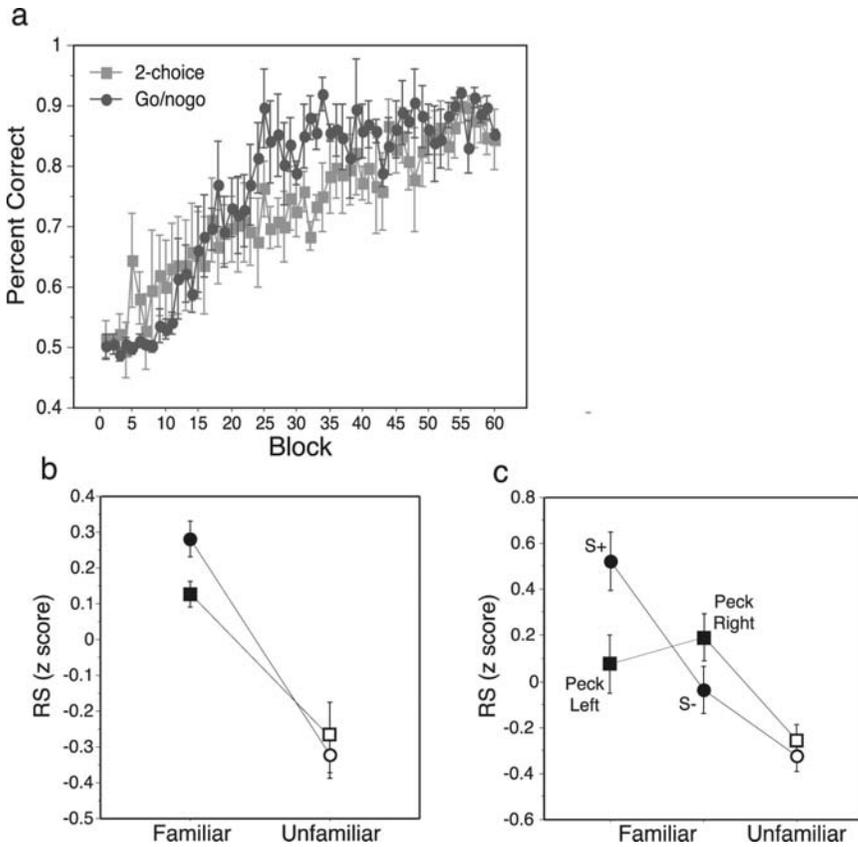


FIGURE 4 Medial regions of the caudal mesopallium (CMM) behavior and physiology. (a) Acquisition curves showing mean performance (as the proportion of correct responses) over the first 60 blocks of recognition training (100 trials/block) with go-nogo (GNG) and two-alternative choice task (2AC) procedures. (b) Mean normalized (z score) response strengths of CCM single units to familiar and unfamiliar songs, split by training regime (2AC: squares; GNG: circles). (c) RS z scores as in (a) but with the responses associated with each of the two sets of training stimuli (2AC: S+ left, S+ right; GNG: S+, S-) shown separately.

stimuli; Figure 5). This suggests that representational plasticity in the system is driven both by “bottom-up” stimulus activation (i.e., exposure) and by “top-down” modulatory mechanisms that map onto variation in behavioral significance.

Response Selectivity

The strong population response bias for familiar songs indicates that CMM neurons do not respond equally to all songs. In fact, roughly 64% of the cells in our

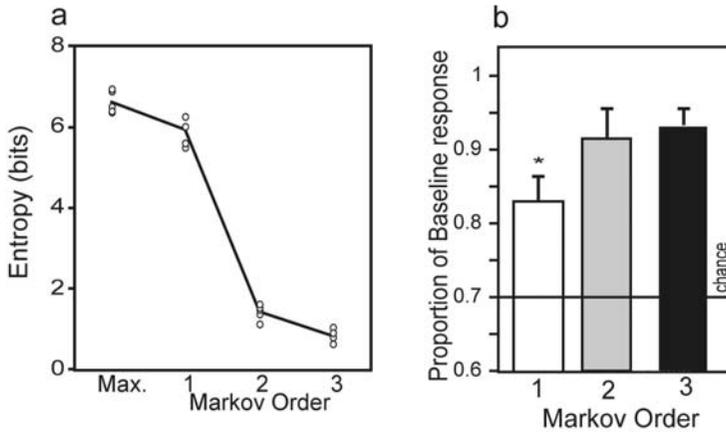


FIGURE 5 Sensitivity to motif sequence. (a) Mean uncertainty plotted as a function of the Markov chain order (motif sequence length) for all the song bouts from each of the five male starlings (open circles). The large drop between first and second order values indicates most of the sequence information in a particular male starling's songs is contained in transitions between pairs of motifs. (b) Performance during partial transfer sessions with synthetic songs that varied sequence reliability, normalized for each individual according to their baseline performance.

sample of CMM gave a selective response to one of the test stimuli (Gentner & Margoliash, 2003); and of these cells, almost all (93%) preferred one of the training songs (Figure 5). For nonselective cells, the song that elicited the strongest response was equally likely to be familiar or unfamiliar.

Motif Selectivity

For many of the song-selective cells, responses were restricted to one or a small number of repeated motifs within one or a few songs, typically with suppression of background activity for all other motifs. These responses appear to be driven by acoustic variation at the level of the motif, responding on average to about 8 motifs, whereas nonselective cells respond to significantly more motifs on average (roughly 20). Additional analyses confirm that the "song" selectivity described in this population of CMM neurons is derived from selective tuning for spectro-temporally complex features centered at the level of the motif (Gentner & Margoliash, 2003). The precise form of these features is not yet characterized. Studies mapping the receptive fields of CMM neurons as a function of complex acoustic features are underway.

On a broader scale, we now recognize that experience-dependent plasticity guides many forms of learning (e.g., Bakin & Weinberger, 1990; Gilbert, Sigman,

& Crist, 2001; Kay & Laurent, 1999; Kilgard, 2003) across a variety of vertebrate sensory systems and brain regions, including the primary sensory cortices once thought to provide static stimulus representations (Nagel & Doupe, 2006; Schoups, Vogels, Qian, & Orban, 2001). The data from starlings suggest that the plasticity mechanisms engaged during song recognition learning give rise, through hierarchical convergence (Sen et al., 2001), to selective neural representations for acoustic features diagnostic of individual (or small sets of) motifs. More simply, songbirds build specific representations for the functionally relevant units of their communication signals. These observations suggest a system that is simultaneously constrained in its immediate representational capacities by its idiosyncratic history, and is tremendously adaptive such that it is able to acquire a broad (currently undetermined) range of representations. Based on these observations, it is tempting to speculate that similar, experience-dependent, hierarchical processes may give rise to the phonetic, phonemic, syllabic, and word-level representations that must be correlated with language experience in humans. Unfortunately, a cellular-level understanding of speech representation in the human auditory system is not yet possible. It is our hope that by the time such experiments do become empirically tractable, the birdsong system will be sufficiently advanced so as to provide a set of testable hypotheses that can be used to examine the similarities and differences (both qualitative and quantitative) across these representational systems.

SENSITIVITY TO SUPRAMOTIF TEMPORAL PATTERNS

The prior sections on individual song recognition in starlings are informative in establishing the close correspondence between neural representation and behaviorally relevant variation in song. This correspondence highlights the functional importance of motif-level song organization. Under normal conditions, however, single motifs are almost never produced in isolation but typically occur as part of long and elaborate song bouts where 25 to 30 different motifs may be strung together in close succession. Building on the success of the song recognition studies just described, we next discuss the perception of higher order patterns constructed from motifs. What do starlings know about the temporal patterning of motifs within song bouts (see Hauser, Barner, & O'Donnell, this issue), and how do they know it?

MOTIF SEQUENCES

Sensitivity to the temporal patterning of motifs could arise in its simplest form through attention to the explicit sequences in which familiar motifs are produced. As an example, a bird might learn to recognize motifs A, B, C, and D (where letters

denote different motifs) *and* the sequence ABCD. If so, then presenting the same motifs in a different order may affect song recognition.

To test the role of motif sequencing in vocal recognition, we trained starlings to recognize normally patterned songs from different singers using a 2AC operant procedure. We then had the birds classify synthetic motif sequences in which the transition probabilities between now familiar motifs were varied systematically using Markov models of each singer's natural song (Gentner, Duffy, Kaloudis, Ellis, & Ball, 1998). Most of the variability in normal motif transitions is accounted for by the second-order transition probabilities (Figure 5). That is, the motif likely to occur at any position in a song bout is most strongly predicted by the motif at the immediately preceding position.

As expected, and consistent with a primary role for motif acoustics, all of the synthetic song sequences were correctly recognized (Figure 5). However, the randomly ordered song bouts were significantly more difficult to recognize than those that followed the natural transition probabilities. In other words, starlings do attend to motif sequencing. Moreover, the responses to the synthetic song bout sequences suggest a close correspondence between production variability in temporal sequencing and perceptual sensitivity. Indeed, anecdotal observations during extracellular electrophysiological recordings from CMM, as described above, point to the presence of cells tuned to temporal combinations of motifs. It is unclear at present whether the close correspondence between behavioral performance and the concentration of information in the second-order motif transitions of natural songs represents a perceptual constraint in terms of neuronal temporal integration windows or an artifact of the motif sequences in natural song bouts. The upper bound on sensitivity to significant temporal sequence information encoded beyond the second-order transitions has not been explored.

MOTIF PATTERNS

Recently we have begun a second approach to studying temporal pattern perception that moves beyond simple transition probabilities between adjacent motifs to ask if starlings can acquire an abstract rule that describes the patterning of familiar motifs (Gentner et al., 2006). That is, instead of learning the explicit sequence of motifs ABCD as mentioned above, can a starling learn the pattern *ABCD*, where the letters now denote sets of motifs that can occur at each position. Rules that describe sequences of patterned strings have a rich history in the theory of formal grammars.

A formal grammar is a collection of rules that operate on a set (or sets) of elements (a "vocabulary") to produce patterned strings (the "language"). Conversely, any grammar, once known, can be used to decide whether a given string is acceptable (i.e., grammatical) within a given language (Chomsky, 1957). Formal gram-

mars can be classified hierarchically according to the complexity of the patterns they can produce or recognize (Hopcroft & Ullman, 1979). Finite-state grammars (FSG; Figure 6) are the most limited type of formal grammar and have been thought to describe all animal communication systems (but see Suzuki, Buck, & Tyack, 2006). Human languages minimally require a grammar more complex than finite-state, called a context-free grammar (CFG; Figure 6), in part to support the

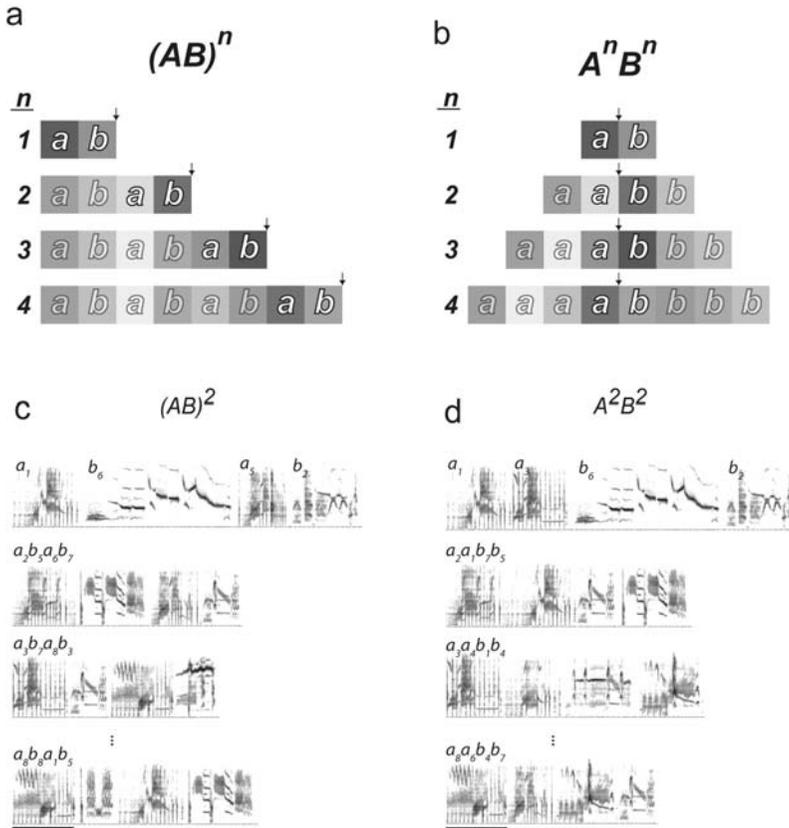


FIGURE 6 Grammatically generated motif patterns. Both the (a) finite-state grammar (FSG) and (b) context-free grammar (CFG) generate relatively simple temporally patterned sequences of elements (lowercase letters) of the sets denoted by “A” and “B”. Increasingly longer strings of the form $(AB)^n$, where n gives the number of AB iterations, are produced by rules that append elements to the end of a shorter, $n - 1$ order, string. An equivalent set of rules could append elements to the start of the $n - 1$ string. In contrast, increasingly longer strings with the form $A^n B^n$ are produced by rules that embed elements into the center of an $n - 1$ sequence. Sonograms (frequency range: .2 – 10.0 kHz; scale bars = 1 sec) showing four of the eight sequences constructed from (c) the FSG, $(AB)^n$, and (d) the CFG, $A^n B^n$, used in the initial FSG versus CFG pattern classification training with $n = 2$.

recursive hierarchical embedding common in many syntactic structures (Chomsky, 1957; Hopcroft & Ullman, 1979). In principle, there is no reason why grammatical rules such as these cannot be used to produce patterned sequences from any set of discrete elements (e.g., Fitch & Hauser, 2004; Fitch, Hauser, & Chomsky, 2005; Newport, Hauser, Spaepen, & Aslin, 2004; Ramus, Hauser, Miller, Morris, & Mehler, 2000) including birdsongs.

Recently, we've shown that European starlings (*Sturnus vulgaris*) can learn to classify temporal patterns of song motifs generated by both FSG and CFG. We used eight "rattle" and eight "warble" motifs (two acoustically distinct starling song motif classes) to create all 4,096 possible sequences (two complete "languages") for each of two distinct grammars (Figure 6). Our CFG generated strings of the form A^2B^2 , whereas the FSG generated strings of the form $(AB)^2$, where A and B refer to the rattles and warbles. We trained 11 European starlings, using a GNG operant conditioning procedure (Figure 1), to classify subsets of sequences from each language (Gentner et al., 2006).

Although many more have since been trained, 9 of the original 11 starlings eventually learned to classify the FSG and CFG sequences accurately, and compared to recognition of natural songs (Gentner & Hulse, 1998), the task was clearly difficult (Figure 7). As an initial test of learning generalization, and to rule out rote memorization of the CFG and FSG patterns, we transferred the four subjects that learned quickest from the 16 baseline training stimuli to 16 novel sequences from the same two grammars, A^2B^2 and $(AB)^2$, maintaining the same operant contingencies to the novel CFG and FSG stimuli as used during baseline training. Subjects correctly classified the novel CFG and FSG sequences during the first transfer session (Figure 7). These results demonstrate that subjects did not simply memorize the training stimuli, but instead acquired some more general knowledge about features diagnostic of the two grammars and applied this knowledge to classify the novel stimuli correctly. Given that the same elements (motifs) composed the sequences in each class, this knowledge must be related to the patterning of elements in each class. Additional generalization tests using "probe" procedures that control for the direct effects of differential reinforcement learning on the test stimuli show very similar results (Gentner et al., 2006). To understand what about the patterns the starlings learned, we conducted several additional tests.

Classification of Agrammatical Sequences

One solution strategy is to learn only the FSG and treat the CFG as the complement set, labeling anything a CFG that is not a legal FSG sequence. Learning only the FSG would be consistent with interpretations of comparable experiments conducted with cotton-top tamarin monkeys (Fitch & Hauser, 2004; Hauser, Newport, & Aslin, 2001; Hauser, Weiss, & Marcus, 2002). To test this possibility, we constructed 16 motif sequences based on four different agrammatical patterns

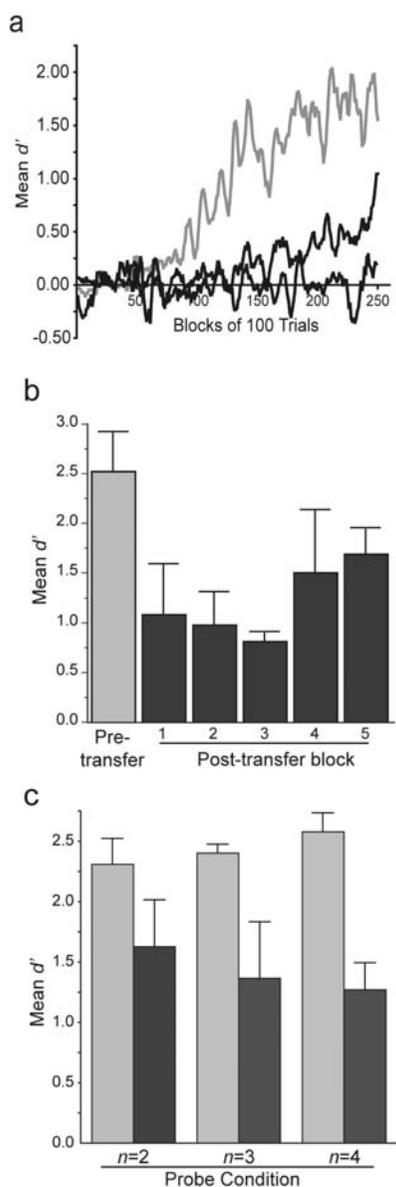


FIGURE 7 Pattern recognition. (a) Acquisition curves for the baseline finite-state grammar/context-free grammar (FSG/CFG) classification, showing mean d' over the first 250 blocks (100 trials/block) for birds that learned quickly and were subjected to further testing (grey, $N = 4$) and birds that learned slowly (black, $N = 5$) or not at all (black, $N = 2$). (b) Mean d' just prior to (light bar) and over the first 100 trials after transfer to novel motif sequences in the FSG and CFG training patterns (dark bars), showing significant discrimination. (c) Subjects correctly classified novel $A^n B^n$ and $(AB)^n$ sequences when $n = 2, 3$, and 4 (dark bars) during probe sessions while maintaining accurate recognition of the baseline patterns (light bars).

(AAAA, BBBB, ABBA, BAAB; four exemplars/pattern using the same A and B motif vocabularies as the two grammars) and presented them as additional probe stimuli during the novel grammatical probe sessions noted above. If the subjects had indeed learned the CFG and FSG rules, then they should treat the agrammatical stimuli differently than the novel grammatical stimuli. Indeed, the response patterns for the agrammatical probe stimuli differed from the response patterns for FSG stimuli for all four birds, and differed from the response patterns for CFG stimuli for three of the four birds. These results suggest that subjects learned both the CFG and the FSG patterns.

Another strategy that could account for the accurate classification of CFG and FSG patterns, and the generalization to novel instances of each grammar, is attention only to “low-level” features that differ between the two patterns. For example, the task could reduce to the discrimination between “AB” and “AA”, or “AB” and “BB”, if only the initial (primacy) or terminal (recency) motif pairs are attended to, respectively. Comparing the responses to different classes of agrammatical stimuli, however, demonstrated that for three of the four birds, neither of these simple primacy or recency effects could account for the classification behavior. A similar argument rules out the putative strategy that the subjects counted the number of transitions between “A” and “B” motifs in each sequence to classify the CFG and FSG stimuli, and a host of other potential solution strategies (see Gentner et al., 2006). What remains is the conclusion that starlings learned the general motif patterns defined by each grammar.

Classification of Higher Order Sequences

Although time and memory capacity both constrain the functional length of any grammatical string, part of the power of a generative grammar is its capacity to describe strings of arbitrary length. To test whether our subjects generalized from A^2B^2 to higher orders of grammatical structure, we probed subjects with $n = 3$; that is, A^3B^3 and $(AB)^3$, and $n = 4$ motif sequences although subjects maintained baseline $n = 2$ discrimination. All subjects accurately classified the $n = 3$ CFG and FSG sequences and the $n = 4$ CFG and FSG sequences, presumably by applying the patterning rules acquired during $n = 2$ training (Figure 7).

Testing Finite-State Approximations

One nontrivial solution strategy that could explain all of our results is to learn a “simpler” grammar that approximates (but does not explicitly define) the patterns in the A^nB^n sequences. Sequences that follow the pattern A^nB^n constitute a subset of those that follow the more general pattern A^*B^* , where the number of *as* and *bs* can vary independently. Although a CFG is required to produce sequences in which the number of *as* and *bs* are matched, as in A^nB^n , the whole of A^*B^* can be

generated by a FSG. To test the possibility that subjects learned an A^*B^* finite-state approximation to the CFG rather than A^nB^n explicitly, we examined their responses to the following A^*B^* patterns: A^1B^3 , A^3B^1 , A^2B^3 , A^3B^2 (four randomly chosen sequences for each pattern, same A/B motif vocabularies as with all the other stimuli). These sequences were delivered as additional probe stimuli during sessions when novel $n = 2-4$ grammatical sequences were also presented. If subjects learned the FSG A^*B^* , then the pattern of response to these stimuli should match the response to A^nB^n probes presented in the same session. The data refute this hypothesis. All subjects showed a strong bias to treat the A^*B^* patterns differently than the A^nB^n reference stimuli, while maintaining accurate classification of the A^nB^n and $(AB)^n$ reference and training stimuli. These results suggest that subjects did not solve the recursive classification tasks by learning a finite-state approximation to the CFG. Rather, it appears that subjects learned A^nB^n , or a functionally equivalent rule (Gentner et al., 2006).

CONCLUSIONS

The foregoing data suggest that starlings are surprisingly sensitive to the temporal patterning of motifs in a song. These sensitivities extend to the recognition of patterns described by the rules of some formal grammars, including at least one set of rules that entails recursive embedding. It has been hypothesized that recursion forms the computational core of a uniquely human narrow faculty for language (Fitch et al., 2005; Hauser, Chomsky, & Fitch, 2002). The A^nB^n language that starlings can learn is only one of several possible context-free languages that entail embedding and recursion (Chomsky, 1957), and there may well be more complex context-free languages that starlings cannot learn to recognize (see Gentner et al., 2006). In any case, at least a simple level of recursive syntactic pattern processing is shared between humans and other animals.

In comparison to natural language, the FSG and CFG patterns learned by the starlings are quite simple. The recursion entailed by phrase structure grammars such as A^nB^n is captured by the formal construct of a stack that defines a pushdown automaton (Hopcroft & Ullman, 1979), in which the stack allows the grammar to express long-range dependencies not found in simpler grammars such as FSGs. The present results demonstrate that starlings, suitably trained, can learn to identify patterns in both of these simple classes, thus showing that they have access to the computational equivalent of a stack (at least of limited depth). To be clear, these results *do not*, as some popular accounts have suggested, show that starlings can represent or learn grammars that are anything like human languages or anything like a context-free approximation to human language, nor do they suggest that starlings can learn *any* CFG. Indeed the development of algorithms with the capacities

to learn CFGs that can approximate natural-language syntax is still very much an open research area in computational linguistics.

In contrast, the computational requirements for learning the simple CFGs described here are actually quite modest (Rodríguez, 2001), requiring at a minimum a push-down automaton to remember or count the number of embeddings. There is ample evidence that animals have memories for acoustic objects and can achieve simple counting (Brannon & Terrace, 1998), and single neurons with acoustic response properties that are sensitive to the number of distinct elements in a sequence have been reported (Edwards, Alder, & Rose, 2002). Elaboration of recognition, classification, and memory capacities in species such as vocal learners that produce long temporal strings of component signals may help explain why starlings were adept at the syntactic pattern classification tasks reported here. Said differently, the ability to process recursive structure revealed through operant conditioning may reflect an opportunistic assembly of available cognitive mechanisms that serve other functions in starlings (see Hauser et al., this issue). Although there has been some discussion of the role of syntactic rules in structuring bird song in chickadees and wrens (Clucas, Freeberg, & Lucas, 2004; Hailman & Ficken, 1986; Holland, Dabelsteen, & Paris, 2000), the existence of complex patterning rules in songbird vocalizations has not received adequate attention. Nonetheless, there is no strong evidence to support the notion that songbirds (or any nonhuman species) use syntax to vary the *semantic* content of vocal signals in the combinatorial manner common among humans. By itself, the ability to process simple recursive syntactic structures may be of little functional significance or evolutionary value, and may represent a necessary but insufficient precondition to the use of unbounded signal sets. It is interesting to consider that these sorts of linguistically primitive pattern recognition abilities might need to be paired with more sophisticated gestural systems as a permissive step to subsequent evolution of language in its modern forms (see Tomesello, this issue).

In establishing our claims of starling temporal pattern recognition, we have attempted to rule out the most plausible finite-state solution strategies that could account for accurate classification of A^nB^n patterns. In practice, however, the stimulus sets used to test such claims must be finite. Thus, the theoretical possibility remains that a FSG, however heavily contrived, may account for the observed behavior. Of course, theoretical issues relating to indeterminacy in proving the use of a context-free rather than FSG extend to studies of grammatical competence in humans as well. In this regard, one should be cautious of relying exclusively on ideal mathematical descriptions of formal grammars and automata as models for the biological basis of complex temporal pattern recognition. Although theoretical descriptions of syntax are well developed, we know almost nothing of the underlying biological brain mechanisms that support pattern rule representation, learning, and use in *any* organism, including humans. In the pursuit of this understanding, it may prove more useful to consider many species differences as quantitative rather

than qualitative distinctions in cognitive mechanisms. There may be no single property, processing capacity, or brain mechanism that marks the many ways in which human language differs from nonhuman communication systems (Pinker & Jackendoff, 2005).

The comparative studies of pattern recognition presented here can clarify with biological detail the commonalities and differences between components of human language and communication in other species. It will be important for future studies to gauge the extent of the syntactic abilities reviewed here, by examining the processing of more sophisticated recursive structures common in human languages, including those that entail dependencies among nested items. In this context, we note that the CFG/FSG classification task posed to the starlings had no semantic demands; it was purely about pattern recognition. As such, it differs markedly from the functional demands of comprehension that humans bring to the perception of speech sounds. Current work in the laboratory is examining the generalization of acquired syntactic knowledge in starlings, with the hope of exploring the interface between syntactic and semantic knowledge in a more simple communication system than human language. Together with the recent observation of neurons selective for learned sequences of motifs in starlings (Gentner & Margoliash, 2003), these results open the way to a comparative exploration of the physiological and cellular-level brain mechanisms for some complex forms of syntactic processing.

REFERENCES

- Adret-Hausberger, M., & Jenkins, P. F. (1988). Complex organization of the warbling song in starlings. *Behaviour*, *107*, 138–156.
- Bakin, J. S., & Weinberger, N. M. (1990). Classical conditioning induces cs-specific receptive field plasticity in the auditory cortex of the guinea pig. *Brain Research*, *536*(1–2), 271–286.
- Beecher, M., Campbell, S. E., & Burt, J. (1994). Song perception in the song sparrow: Birds classify by song type but not by singer. *Animal Behaviour*, *47*, 1343–1351.
- Beecher, M. D. (1991). Success and failures of parent–offspring recognition in animals. In P. G. Hepper (Ed.), *Kin recognition* (94–124). Cambridge, England: Cambridge University Press.
- Bonke, D. A., Bonke, D., & Scheich, H. (1979). Connectivity of the auditory forebrain nuclei in the guinea fowl (*numida meleagris*). *Cell and Tissue Research*, *200*, 101–121.
- Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*(5389), 746–749.
- Bricker, P. D., & Pruzansky, S. (1976). Speaker recognition. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 295–326). New York: Academic.
- Capsius, B., & Leppelsack, H. J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hearing Research*, *96*(1–2), 59–70.
- Carr, C. E. (1992). Evolution of the central auditory system in reptiles and birds. In D. B. Webster, R. R. Fay, & A. N. Popper (Eds.), *The evolutionary biology of hearing* (pp. 511–544). New York: Springer-Verlag.

- Chaiken, M., Böhner, J., & Marler, P. (1993). Song acquisition in European starlings, *sturnus vulgaris*: A comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. *Animal Behaviour*, *46*, 1079–1090.
- Chew, S. J., Mello, C., Nottebohm, F., Jarvis, E., & Vicario, D. S. (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 Science*, *92*, 3406–3410.
- Chew, S. J., Vicario, D. S., & Nottebohm, F. (1996). A large-capacity memory system that recognizes the calls and songs of conspecifics. *Proceedings of the National Academy of Science*, *93*, 1950–1955.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Clucas, B. A., Freeberg, T. M., & Lucas, J. R. (2004). Chick-a-dee call syntax, social context, and season affect vocal responses of Carolina chickadees (*poecile carolinensis*). *Behavioral Ecology & Sociobiology*, *57*, 187–196.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631.
- Edwards, C. J., Alder, T. B., & Rose, G. J. (2002). Auditory midbrain neurons that count. *Nature Neuroscience*, *5*, 934–936.
- Eens, M., Pinxten, M., & Verheyen, R. F. (1989). Temporal and sequential organization of song bouts in the european starling. *Ardea*, *77*, 75–86.
- Eens, M., Pinxten, R., & Verheyen, R. F. (1991). Organization of song in the european starling—species-specificity and individual-differences. *Belgian Journal of Zoology*, *121*(2), 257–278.
- Falls, J. B., & Brooks, R. J. (1975). Individual recognition by song in white-throated sparrows. II: Effects of location. *Canadian Journal of Zoology*, *53*, 1412–1420.
- Falls, J. B., Krebs, J. R., & McGregor, P. K. (1982). Song matching in the great tit (*parus major*): The effect of similarity and familiarity. *Animal Behaviour*, *30*, 997–1009.
- Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*(5656), 377–380.
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*(2), 179–210.
- Gentner, T. Q., Duffy, D. L., Kaloudis, K., Ellis, E., & Ball, G. F. (1998). Behaviorally relevant variation in male song induces differential expression of the *ieg* zenk in a sub-region of ncm in female starlings. *Society for Neuroscience Abstracts*, *24*, 700.
- 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.
- Gentner, T. Q., & Hulse, S. H. (2000). Perceptual classification based on the component structure of song in european starlings. *Journal of the Acoustical Society of America*, *107*, 3369–3381.
- Gentner, T. Q., Hulse, S. H., & Ball, G. F. (2004). Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *190*, 1001–1010.
- Gentner, T. Q., Hulse, S. H., Bentley, G. E., & Ball, G. F. (2000). Individual vocal recognition and the effect of partial lesions to hvc on discrimination, learning, and categorization of conspecific song in adult songbirds. *Journal of Neurobiology*, *42*(1), 117–133.
- Gentner, T. Q., & Margoliash, D. (2002). *The neuroethology of vocal communication: Perception and cognition* (Vol. 16). Springer-Verlag.
- Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature*, *424*(6949), 669–674.
- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, *31*, 681–697.

- Godard, R. (1991). Long-term memory for individual neighbors in a migratory songbird. *Nature*, *350*, 228–229.
- Grace, J. A., Amin, N., Singh, N. C., & Theunissen, F. E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. *Journal of Neurophysiology*, *89*(1), 472–487.
- Hailman, J. P., & Ficken, M. S. (1986). Combinational animal communication with computable syntax: Chick-a-dee calling qualifies as “language” by structural linguistics. *Animal Behaviour*, *34*, 1899–1901.
- Hausberger, M. (1997). Social influences on song acquisition and sharing in the European starling (*sturnus vulgaris*). In C. Snowden & M. Hausberger (Eds.), *Social influences on vocal development* (128–156). Cambridge, England: Cambridge University Press.
- Hausberger, M., & Cousillas, H. (1995). Categorization in birdsong: From behavioural to neuronal responses. *Behavioural Processes*, *35*(1–3), 83–91.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–1579.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*(3), B53–64.
- Hauser, M. D., Weiss, D., & Marcus, G. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*(1), B15–22.
- Häusler, U. (Ed.). (1996). *Measurement of short-time spatial activity patterns during auditory stimulation in the starling*. New York: Plenum.
- Holland, J., Dabelsteen, T., & Paris, A. L. (2000). Coding in the song of the wren: Importance of rhythmicity, syntax and element structure. *Animal Behaviour*, *60*, 463–470.
- Hopcroft, J., & Ullman, J. (1979). *Introduction to automata theory, languages, and computation*. Reading, MA: Addison Wesley.
- Hsu, A., Woolley, S. M., Fremouw, T. E., & Theunissen, F. E. (2004). Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *Journal of Neuroscience*, *24*, 9201–9211.
- Hulse, S. H. (1995). The discrimination-transfer procedure for studying auditory perception and perceptual invariance in animals. In G. M. Klump, R. J. Dooling, R. R. Fay, & W. C. Stebbins (Eds.), *Methods in comparative psychoacoustics* (Vol. 10, pp. 319–330). Basel, Switzerland: Birkhäuser Verlag.
- Jarvis, E. D., Mello, C. V., & Nottebohm, F. (1995). Associative learning and stimulus novelty influence the song-induced expression of an immediate early gene in the canary forebrain. *Learning and Memory*, *2*, 62–80.
- Jones, M. W., Errington, M. L., French, P. J., Fine, A., Bliss, T. V., Garel, S., et al. (2001). A requirement for the immediate early gene *zif268* in the expression of late *Itp* and long-term memories. *Nature Neuroscience*, *4*(3), 289–296.
- Kay, L. M., & Laurent, G. (1999). Odor- and context-dependent modulation of mitral cell activity in behaving rats. *Nature Neuroscience*, *2*, 1003–1009.
- Kilgard, M. (2003). Cholinergic modulation of skill learning and plasticity. *Neuron*, *38*, 678–680.
- Kroodsma, D. E., & Miller, E. H. (Eds.). (1996). *Ecology and evolution of acoustic communication in birds*. Ithaca, NY: Comstock/Cornell.
- Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. Proceedings of the National Academy of Science, USA, *100*, 9645–9646.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews: Neuroscience*, *5*, 831–843.
- Leppelsack, H.-J. (1983). Analysis of song in the auditory pathway of song birds. In J. P. Evert, B. R. Capranica, & D. J. Ingle (Eds.), *Advances in vertebrate neuroethology* (pp. 783–799). New York: Plenum.
- Leppelsack, H.-J., & Schwartzkopff, J. (1972). Properties of acoustic neurons in the caudal neostriatum of birds. *The Journal of Comparative Physiology*, *80*, 137–140.

- Leppelsack, H. J., & Vogt, M. (1976). Response to auditory neurons in the forebrain of a song bird to stimulation with species-specific sounds. *Journal of Comparative Physiology*, *107*, 263–274.
- Lind, H., Dabelsteen, T., & McGregor, P. K. (1997). Female great tits can identify mates by song. *Animal Behaviour*, *52*, 667–671.
- Margoliash, D. (2002). Evaluating theories of bird song learning: Implications for future directions. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *188*, 851–866.
- Marler, P. (1970). Birdsong and speech development: Could there be parallels? *American Scientist*, *58*, 669–673.
- Marler, P. (1975). On the origin of speech from animal sounds. In J. F. Kavanaugh & J. E. Cutting (Eds.), *The role of speech in language* (11–37). Cambridge, MA: The MIT Press.
- Mello, C. V., & Clayton, D. F. (1995). Differential induction of the zenk gene in the avian forebrain and song control circuit after metrazole-induced depolarization. *Journal of Neurobiology*, *26*(1), 145–161.
- Mello, C. V., Velho, T. A., & Pinaud, R. (2004). Song-induced gene expression: A window on song auditory processing and perception. *Annals of the New York Academy of Science*, *1016*, 263–281.
- Mello, C. V., Vicario, D. S., & Clayton, D. F. (1992). Song presentation induces gene-expression in the songbird forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, *89*, 6818–6822.
- Müller, C. M., & Leppelsack, H. J. (1985). Feature extraction and tonotopic organization in the avian forebrain. *Experimental Brain Research*, *59*, 587–599.
- Nagel, K. I., & Doupe, A. J. (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron*, *51*, 845–859.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance ii. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*(2), 85–117.
- O’Loughlen, A. L., & Beecher, M. D. (1997). Sexual preferences for mate song types in female song sparrows. *Animal Behaviour*, *53*, 835–841.
- Peek, F. W. (1972). An experimental study of the territorial function of vocal and visual display in the male red-winged blackbird (*agelaius phoeniceus*). *Animal Behaviour*, *20*, 112–118.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What’s special about it? *Cognition*, *95*(2), 201–236.
- Pisoni, D. B. (1997). Some thoughts on “normalization” in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego, CA: Academic.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349–351.
- Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology—Human Perception and Performance*, *23*, 651–666.
- Ribeiro, S., Cecchi, G. A., Magnasco, M. O., & Mello, C. V. (1998). Toward a song code: Evidence for a syllabic representation in the canary brain. *Neuron*, *21*, 359–371.
- Rodriguez, P. (2001). Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Computation*, *13*, 2093–2118.
- Rübsamen, R., & Dörrscheidt, G. J. (1986). Tonotopic organization of the auditory forebrain in a songbird, the European starling. *Journal of Comparative Physiology*, *158*, 639–646.
- Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in v1 neurons. *Nature*, *412*(6846), 549–553.
- Sen, K., Theunissen, F. E., & Doupe, A. J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. *Journal of Neuropsychology*, *86*, 1445–1458.
- Sheffert, S. M., Pisoni, D. B., Fellowes, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1447–1469.

- Stoddard, P. K. (1996). Vocal recognition of neighbors by territorial passerines. In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 356–374). Ithaca, NY: Comstock/Cornell.
- Stripling, R., Volman, S. F., & Clayton, D. F. (1997). Response modulation in the zebra finch neostriatum: Relationship to nuclear gene regulation. *Journal of Neuroscience*, *17*, 3883–3893.
- Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Information entropy of humpback whale songs. *Journal of the Acoustical Society of America*, *119*, 1849–1866.
- Theunissen, F. E., & Doupe, A. J. (1998). Temporal and spectral sensitivity of complex auditory neurons in the nucleus hvc of male zebra finches. *Journal of Neuroscience*, *18*, 3786–3802.
- Theunissen, F. E., & Shaevitz, S. S. (2006). Auditory processing of vocal sounds in birds. *Current Opinions in Neurobiology*, *16*, 400–407.
- Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*taenopygia guttata*). *Journal of Comparative Neurology*, *366*, 613–642.
- Wiley, R. H., Hatchwell, B. J., & Davies, N. B. (1991). Recognition of individual males' songs by female dunnocks: A mechanism increasing the number of copulatory partners and reproductive success. *Ethology*, *88*, 145–153.
- Woolley, S. M., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience*, *8*, 1371–1379.
- Zeigler, H., & Marler, P. (2004). Behavioral neurobiology of birdsong (Conference proceedings, New York, December 12–14, 2002). *Annals of the New York Academy of Science*, *xiii–xvii*, 1–788.