Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris

TIMOTHY Q. GENTNER & STEWART H. HULSE
Johns Hopkins University, Department of Psychology

(Received 2 June 1997; initial acceptance 9 September 1997; final acceptance 18 December 1997; MS. number: A7941)

ABSTRACT

The capacity for vocal recognition of individual conspecifics is well documented in many species, but the perceptual mechanisms that underlie this ability in oscines are less well understood. Using operant conditioning, we trained three groups of European starlings on a baseline task to discriminate the songs of one male starling from those of four others. Each subject heard songs from the same five singers, but the to-be-recognized individual varied among birds. We grouped the subjects according to sex and their degree of previous exposure to the songs used as stimuli in this experiment. The first group (N = 5 males) identified their own songs from those of four familiar males. The second group (N = 5 males) was familiar with the song stimuli, but none of the songs was their own. The third group (N = 4 females) was unfamiliar with the songs. After learning the baseline discrimination, the subjects were exposed to new natural and synthetic stimuli. The subjects maintained the ability to identify correctly an individual on the basis of novel song bouts, and showed differential responding on the basis of the sequence of song types in song bouts that were modelled using Markov chains. Based upon patterns of responding to these different stimuli, we conclude that European starlings are capable of individual vocal recognition, and that this process is mediated by mechanisms involving the memorization of individually specific song types, the sequential ordering of song types within different bouts of an individual, and perhaps by individually specific spectral (or voice) characteristics that generalize across song types.
has been suggested for great tits, Parus major (Weary & Krebs 1992), but is clearly not a relevant acoustic feature for song sparrows (Beecher et al. 1994). (4) The sequence in which multiple song types occur (shared or not) may show individual variation. This final strategy remains untested.

In species that sing multiple song types the four mechanisms outlined above may not be mutually exclusive, and there is no a priori reason to suspect that individual vocal recognition in a single species should rely on only one perceptual mechanism. In fact, given the diverse number of contexts in which individual recognition is likely to occur and differences among intended receivers of song, one could easily predict that the opposite would be true. Thus one might expect to see information about an individual singer coded at multiple levels throughout that bird’s song or songs. Support for the hypothesis that the perception and encoding of individual identity information occurs at multiple levels in the song is difficult to gather from a number of separate studies where extrapolation across different dependent measures is usually required. The present set of experiments, however, does allow for direct comparisons to be made between multiple perceptual mechanisms for individual vocal recognition; and once established, it examines the extent to which both males and females rely upon these mechanisms.

This study extends our knowledge of vocal communication in songbirds by first demonstrating the capability for individual vocal recognition in European starlings, Sturnus vulgaris, and then by investigating the role of multiple perceptual mechanisms for individual vocal recognition in this species. We accomplish this goal by using operant conditioning techniques to first train different groups of male and female European starlings in a baseline task to discriminate between multiple exemplars of conspecific male song on the basis of singer identity. Then, using carefully chosen novel exemplars and synthetically constructed song bouts from the same singers as in the baseline task, we examined the specific information that the birds used to solve the baseline task. To generate synthetic song bouts, we took an information theoretic approach based on Markov chain models of transition probabilities between sequential acoustic events (see Chatfield & Lemon 1970). By controlling the subjects’ exposure to both the components of the song bouts (song types) and the sequence of song types within bouts we tested perceptual mechanisms involving memorization, sensitivities to temporal organization, and the recognition of voice characteristics.

European starlings are a sensible choice as subjects for this study. They are semicolonial oscines (Feare 1984) in which both males and females sing long and elaborate song bouts composed of many song types (Adret-Hausberger & Jenkins 1988; Eens et al. 1989). The complexity of male song bouts allows for the possibility that information may be coded at multiple levels, and the nature of starlings’ social structure holds the possibility of extending our knowledge of individual recognition beyond territorial or kin recognition contexts.

**METHODS**

**Subjects**

Ten male and four female European starlings, captured from the wild as adults were used in this experiment. Each subject’s sex was determined by the presence or absence of pale coloured iris pigmentation (‘eye ring’) and the coloration of the base of their beak (Feare 1984). All of the birds were naïve to operant experimental procedures. All subjects were housed in individual cages in mixed-sex avaiaries containing approximately 4–10 other European starlings. Males and females were housed in separate rooms, and additional male starlings not used in this experiment were housed with the females. Timers connected to fluorescent fixtures controlled the light/dark schedule in the aviaries. The females were held on a light/dark schedule synchronized with the natural photoperiod in Baltimore. The males were held on an 11:13 h light:dark cycle. Throughout the course of the experiment all birds were maintained on a diet of Purina Start and Grow (Purina, St Louis, Missouri) at 85% of their ad libitum weights. The birds had access to water at all times. The experiment was run from early October 1995 to late February 1996.

The subjects for this experiment were divided into three groups based upon their sex and the extent of prior exposure to the song stimuli with which they were eventually trained. The five males whose songs were used to generate all of the stimuli for this experiment are referred to as the ‘bird’s-own-song’ group. The five remaining males, referred to as the ‘male-familiar-song’ group, had been housed in group cages with the bird’s-own-song males for several months prior to this experiment and so had extensive experience with the songs of the males in the bird’s-own-song group. The five females were naïve to all of the songs used in this experiment and so are referred to as the ‘female-unfamiliar-song’ group.

**Apparatus**

All of the experimental sessions were conducted inside a sound attenuating test chamber (IAC Model AC-3, New York). The chamber measured 80 × 60 × 60 cm (width × height × depth). We transported the birds from the aviary to the testing apparatus in a 30 × 20 × 20 cm stainless steel weld-wire cage that was attached to a response panel mounted inside the test chamber. The response panel was suspended from the ceiling of the test chamber, and formed one end of the test cage. The birds gained access to the response panel after a sliding door on the transport cage was removed at the start of each session. The panel contained three horizontally aligned, translucent response buttons (keys). The keys were 2 cm in diameter and were spaced 6 cm apart, centre to centre. Food hoppers (Gerbrands Model G5610, Cambridge, Massachusetts) delivered food (Purina Start and Grow) to a 6.0 × 4.5 cm opening centred 5.5 cm below the middle key. Two 10 W incandescent lamps, located behind a translucent screen mounted on the back wall of the test chamber, provided indirect illumination of the test.
chamber. A speaker (Bose model 101, Framingham, Massachusetts) was located above and behind the response panel, and shielded from the view of the bird by a 5 × 20 cm rectangular aluminium plate, 3 mm thick. A 386 PC equipped with a parallel digital interface board (Keithley Metabyte PIO-12, Tauton, Massachusetts) and a sound card (Creative Labs SB16, Milpitas, California) for D/A signal conversion controlled the stimulus presentation, response contingencies, and data collection using MELODIE version 2.0 (Psychology Software Tools, Pittsburgh, Pennsylvania). Analogue signals from the controlling computer were amplified (Crown model D-75, Elkhart, Indiana) then sent to the speaker in the test chamber. Prior to testing, we set the maximum sound level within the test chamber at 70 ± 2 dB by placing a microphone, connected to a sound level meter (Rion Model NA-20, Tokyo, Japan), at a position inside the test chamber that approximated that of the bird’s head during experimentation.

Male European Starling Song

Male European starling song has been well described (Adret-Hausberger & Jenkins 1988; Eens et al. 1989). Songs are organized into bouts that vary in length, but have a stereotypical gross structure common to nearly all bouts (Eens et al. 1991b). In keeping with Eens et al.’s (1989) nomenclature, a typical song bout is composed of sequentially patterned multiple note clusters referred to as ‘song types’. Song types are generally less than one second long and often repeated several times before the next song type is sung (Adret-Hausberger & Jenkins 1988; Eens et al. 1989). In this way, sequences of song types are strung together in time to produce a single song bout. The length of a song bout correlates positively with age and varies between individuals, with mature males producing longer songs than younger males (Eens et al. 1991b). In keeping with Eens et al.’s conventions of Eens et al. (1989) we used these conventions of Eens et al. (1989) to divide the song bouts for each bird.

Recording

We recorded a total of at least 0.5 h of song from five male starlings. To record the songs of an individual, that bird was placed in a 40 × 30 × 35 cm weld-wire cage and then isolated inside a 1.98 × 1.93 × 1.52 m sound attenuating chamber (Industrial Acoustics, New York) for 24–48 h, after which time a female starling was introduced into the chamber in a separate cage. A directional microphone (Sennheiser ME66 & K6 power module, Wedemark, Germany) was placed between the male and female so that the female was positioned at 180° relative to the microphone’s maximum sensitivity, while the male was at 0°. In this way, we recorded only the vocalizations of the male while he had visual and auditory contact with a female. The same female was used to induce song from all the males. The microphone was connected to a digital tape recorder (Sony TCD-D7 or DTC-690, Tokyo, Japan). Each bird was recorded for at least 2 h following the introduction of the female, and if necessary for longer until a total of 0.5 h of song was recorded. In most cases, the 2 h following introduction of the female was sufficient to record more than enough singing. All of the songs for each bird were transferred directly from DAT to hard disk (Apple Macintosh Quadra 650, Cupertino, California) at the original sampling resolution of 48 kHz × 16 bits, using a D/A board with S/PDIF digital inputs (Digidesign Audiomedia II, Menlo Park, California) and SoundDesignerII version 2.8 software.

Stimuli

Recordings of five male European starlings were used to generate all of the stimuli for this experiment. We used five different stimulus sets. Each stimulus set consisted of the following exemplars sampled from a single bird: eight baseline exemplars, eight novel song bout exemplars, eight novel song type exemplars, six synthetic sequence exemplars and three control exemplars. These exemplars are detailed below. In all, 165 different stimulus exemplars were used in this experiment.

Natural song

A time spectrogram was printed for each song bout (Macromedia, SoundEdit Pro version 1.0.5, San Francisco, California) then divided by human observers into sequences of uniquely labelled song types following the conventions of Eens et al. (1989). We used these sequences of song types to divide the song bouts for each
bird into two subsets. Each subset was made up of samples from the original song bouts with the constraint that they had no song types in common. Each song bout sample was 12–15 s long. These two subsets for each recorded bird served as the basis for the three different types of natural song stimuli used in this experiment. We chose 16 exemplars from one subset, designating eight of these ‘baseline’ exemplars and eight of them ‘novel song bout’ exemplars. We then chose eight exemplars from the opposing subset and designated them ‘novel song type’ exemplars. Thus, the baseline and novel song bout stimuli shared common song types and singer identity, but in most cases were sampled from different song bouts; whereas the novel song type stimuli shared singer identity, but no common song types, with the baseline and novel song bout stimuli. Figure 1 shows time spectrograms for 3 of the 120 natural song stimuli used in this experiment.

Synthetic song sequences

We generated synthetic song bouts (sequences of song types) for each of the five recorded males using three different types of Markov process models. Markov models provide a method for quantifying the statistical probabilities associated with any sequence of discrete events, and their use in modelling bird song has been well described (Chatfield & Lemon 1970). For any given sequence of events, each single event can be assigned an overall probability of occurrence that is equal to the number of times that event is observed divided by the total number of events in the sequence. Similarly, each ordered pair of events can be assigned a joint probability equal to the number of times that pair of events is observed divided by the total number of pairs of events, as well as a conditional (or transition) probability equal to the probability of a single event given that some other event has just occurred. Similar probabilities can be extended to ordered triplets, quadruplets, etc. Different Markov chain models make use of different probabilities. First-order Markov chains rely only on the probability of occurrence for single events, and thus maintain random transition probabilities between consecutive events. Second-order Markov chains rely on transition probabilities for ordered pairs of events and third-order Markov chains rely on transition probabilities between ordered triplets. We used multiple recordings (mean = 33.0/bird) of natural bouts from our five stimulus males (described as above) to generate a first-, second- and third-order transition matrix for each individual male. These matrices were then used to generate synthetic strings of song types that an individual starling would be likely to sing, and that conformed to either the first-, second-, or third-order model for that singer. Using sound-editing software (SoundEdit Pro) we took a single representative sample of each song type in a bird’s repertoire and then reassembled these song types according to the sequences of song types in the synthetic strings. We made six different ‘synthetic song sequences’ for each of the five recorded males: two sequences based on the first-order model, two second-order sequences and two third-order sequences. As control songs, we sampled three novel sections of natural song bouts from each bird. By the time the subjects were exposed to the control songs none of the song types contained within them was novel, but in most cases the song bouts from which they were sampled were novel. Figure 2 shows time spectrograms for three of the 30 different synthetic sequences used in this experiment.

Information Analysis of Male Song

Uncertainty is inversely related to the amount of information that a given event (or series of events) provides. That is, the more information one has, the more certain one can be about events that are likely to happen. For this reason, information is commonly measured as a reduction in uncertainty. To describe the amount of information captured by each of the different orders of Markov chains that we used to model the syntactical structure of starling song, we calculated maximum, first-, second- and third-order uncertainty values for the natural songs of each individual (see Fig. 3). The maximum (zero order) uncertainty \( U_{\text{max}} \) describes the total amount of information in a system, and for any string of discrete events is given by the equation:

\[
U_{\text{max}} = U_0 = \log_2 k = - \log_2 (1/k)
\]  

where \( k \) is equal to the number of different events, in this case the number of song types in an individual’s repertoire. If the probability of any event \( P_i \) is expressed as a ratio of the frequency of a given event \( F_i \) to the total number of events \( N_i \), then the first-order uncertainty \( U_1 \) is given by the equation:

\[
U_1 = - \sum P_i \log_2 P_i
\]  

The second-order uncertainty \( U_2 \) is given by the equation:

\[
U_2 = - \sum P_{ij} \log_2 P_{ij}
\]  

where \( P_{ij} \) is the joint probability of \( i \) and \( j \), and \( P_{ij} \) is the conditional probability of \( j \) given \( i \). The equation for third-order uncertainty \( U_3 \) is of the same general form as that for equation (3) except that the probabilities for each order pair \( P_{ij} \) and \( P_{ij} \) are replaced by joint and conditional probabilities for ordered triplets \( P_{ijk} \) and \( P_{ijk} \), respectively.

The mean (± SE) uncertainty values for the songs of each of the five male starlings recorded for use in this experiment are shown in Fig. 3. The mean (± SE) repertoire size of the five male starlings recorded for use in this study was 104.80 (± 12.53) song types. The mean (± SE) song bout length was 38.36 (± 6.09) s. The correlation between song bout length and repertoire size was significant (\( r = 0.765, P < 0.05 \)), and conforms with previous reports (Eens et al. 1991b). There was no significant correlation between the repertoire size of a given individual and the number of trials required to reach criterion on the stimulus set in which the songs of that bird
Figure 1. Time spectrograms of three of the 120 different natural song stimulus exemplars: (a) a baseline stimulus exemplar, (b) a novel song bout exemplar and (c) a novel song type exemplar. All three exemplars were sampled from song bouts of the same male starling. The letters above the time spectrograms indicate the codes used for the sequence of song types in each bout. Note that (a) and (b) share common song types (M2, L1 and M3), whereas all of the song types in (c) are unique.
Figure 2. Time spectrograms of three of the 30 different synthetic stimulus exemplars: (a) a sequence with random transition probabilities between song types (first order) in which the frequency of each song type is maintained; (b) a sequence in which the probability of a given song type is conditional on the preceding song type (second order); (c) a sequence in which the probability of a given song type is conditional on the preceding two song types (third order). The three sequences shown were composed from the song types of a single male. Conditional probabilities were obtained through a Markov analysis of all the recorded song bouts from a given individual male.
were free to complete the shaping sequence at their own pace. Most of the animals completed the entire 300 trial sequence in less than four 1.5-h sessions.

Discrimination training

The birds were trained with a two-alternative choice procedure (see Hulse 1995 for a review of the baseline-transfer procedure). A peck to the centre key initiated a trial by starting the playback of a randomly selected stimulus exemplar, after which a single peck to either the left or the right key led to reinforcement or punishment depending on the key with which that stimulus was arbitrarily associated. Correct responses were reinforced with 4-s access to the food hopper. Incorrect responses were punished with an 8-s time-out during which the house lights were extinguished and the food hopper remained inaccessible. Subjects could increase the amount of available feeding time by making correct responses to the playback stimuli. The distribution of responses to the left and right keys was therefore dependent upon the extent to which the subjects were capable of differentiating between the stimulus classes assigned to each of those keys. Thus by learning to peck the left key following certain stimuli and to peck the right key following certain other stimuli, a discrimination was established between the two classes of stimuli associated with the opposing keys.

The intertrial interval between all trials was 2 s. In the event that a bird failed to respond within 2 s following the completed presentation of a given exemplar, the trial ended and the computer waited for a centre key peck to begin the next trial. In addition to causing an 8-s time-out, an incorrect response to a playback stimulus initiated a correction trial sequence in which the same exemplar was repeated on all subsequent trials until the bird either responded appropriately or not at all. On each trial throughout the course of the experiment there was a 12.0-s ‘observation period’ during which time the stimulus for that particular trial played but responding on the keys had no effect. Immediately following this observation period responses were reinforced as described above. All experimental sessions lasted approximately 1.5 h, and occurred once daily at the same time for each bird. We conducted test sessions from Monday to Saturday. On days that no testing was conducted the birds were fed 15 g of food.

For this specific experiment, the subjects’ task was to discriminate between the songs of an individual male European starling and those of four other male European starlings. For example, one bird was reinforced for pecking the left key each time it heard a song from bird A, and for pecking the right key each time it heard a song from either bird B, C, D, or E. Another bird was reinforced for pecking the left key each time it heard a song from bird B and for pecking the right key each time it heard a song from bird A, C, D, or E. We refer to the key associated with the songs of the single bird as the individual (INDIV) key, and the key associated with the songs from multiple birds as the multiple (MULT) key. Similarly, the songs associated with each key are referred to as INDIV and MULT stimuli, respectively. Given this design and songs from

![Figure 3. Uncertainty plotted as a function of the Markov chain order for all of the song bouts from each of the five male starlings recorded for this experiment. The line connects the means. The large drop between first- and second-order uncertainty values indicates that of the total amount of information present in a particular male starling song \( (U_{\text{max}}) \), most of the information is contained in the second-order transitions between song types.](image-url)
Table 1. Baseline stimulus set configurations

<table>
<thead>
<tr>
<th>Stimulus set</th>
<th>Individual (INDIV)</th>
<th>Multiple (MULT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1, A2, A3, A4</td>
<td>B1, B2, C1, C2</td>
</tr>
<tr>
<td></td>
<td>A5, A6, A7, A8</td>
<td>D1, D2, E1, E2</td>
</tr>
<tr>
<td>2</td>
<td>B1, B2, B3, B4</td>
<td>A1, A2, C3, C4</td>
</tr>
<tr>
<td></td>
<td>B5, B6, B7, B8</td>
<td>D3, D4, E3, E4</td>
</tr>
<tr>
<td>3</td>
<td>C1, C2, C3, C4</td>
<td>A3, A4, B3, B4</td>
</tr>
<tr>
<td></td>
<td>C5, C6, C7, C8</td>
<td>D5, D6, E5, E6</td>
</tr>
<tr>
<td>4</td>
<td>D1, D2, D3, D4</td>
<td>A5, A6, B5, B6</td>
</tr>
<tr>
<td></td>
<td>D5, D6, D7, D8</td>
<td>C5, C6, E7, E8</td>
</tr>
<tr>
<td>5</td>
<td>E1, E2, E3, E4</td>
<td>A7, A8, B7, B8</td>
</tr>
<tr>
<td></td>
<td>E5, E6, E7, E8</td>
<td>C7, C8, D7, D8</td>
</tr>
</tbody>
</table>

*Stimulus exemplars composing each of the five baseline stimulus sets. Letters indicate the identity of the singer and numbers refer to the specific song samples from that individual. We assigned each stimulus set to one subject in each of the three groups. Note that each exemplar was used as a MULT stimulus only once.

five male starlings there are five possible stimulus configurations in which the songs of one individual can be singled out from those of the remaining four (see Table 1).

Baseline

During the baseline sessions each subject learned to discriminate eight INDIV exemplars (all from one bird) from eight MULT exemplars (two from each of the four other birds). The assignment of the INDIV and MULT stimuli to particular keys (either left or right) was counterbalanced across subjects. By design, each of the five different MULT arrangements shared three of its four singers with the other MULT arrangements. However, the specific exemplars that made up each MULT arrangement were unique to that arrangement (see Table 1). One subject from each of the three groups was assigned one of the five stimulus configurations. Subjects in the bird’s own-song group identified their own songs from the songs of the four other starlings. Subjects in the male-familiar-song and female-unfamiliar-song groups were assigned one of the stimulus arrangements at random with the constraint that all five stimulus arrangements were assigned within a group.

Transfer to novel song bout stimuli

We used a complete transfer from the novel song bout stimuli to the novel song bout stimuli to test the arbitrariness of the stimuli classifications in our discrimination. Once an individual bird made correct responses to at least 75% of all the stimuli for three or more consecutive blocks of trials, that bird was transferred to novel song bout stimuli. For novel song bout transfer sessions, the 16 baseline exemplars were simply replaced with 16 novel song bout stimuli. The singer identities for the novel song bout stimuli (and their assignment to either the left or the right key) was exactly the same as that for the baseline stimuli (see Table 1). Recall that the novel song bout stimuli shared many song types with the baseline stimuli, but were in most cases sampled from different song bouts. Comparisons between performance on the first block of transfer trials and the last block of baseline trials were particularly informative. Immediate transfer to the novel song bout stimuli, that is no change in performance between pre- and post-transfer, would indicate that the subjects did not merely learn arbitrary baseline exemplar-key combinations in order to perform the discrimination, arguing instead for a nonarbitrary classification. Non-arbitrary classification is evidence for the formation of natural categories (Herrnstein 1979, 1990). Natural category formation on the basis of an individual’s song bouts is strong evidence that European starlings have a capacity for individual vocal recognition. The subjects were maintained on the novel song bout stimuli for several sessions until performance reached an asymptotic level.

Transfer to novel song type stimuli

We used a complete transfer from the novel song bout stimuli to the novel song type stimuli to test for perceptual mechanisms of individual vocal recognition independent of specific song types. Recall that the novel song type stimuli shared no song types with any of the other stimuli used in either the baseline or novel song bout transfer portions of the experiment. Therefore, the transfer to novel song type stimuli is a strong test of any vocal recognition mechanisms independent of specific song types. Immediate transfer to the novel song type stimuli can only occur if a mechanism for recognizing an individual’s song is independent of the specific song types in that individual’s repertoire. After an individual bird responded at or above the 75% performance criterion for three or more blocks of trials with the novel song bout stimuli, that bird was transferred to the novel song type stimuli. For novel song type transfer sessions, the 16 novel song bout stimuli were replaced with 16 novel song type stimuli. The singer identities for the novel song type stimuli (and their assignment to either the left or the right key) was exactly the same as that for the baseline and novel song bout stimuli (see Table 1). Subjects were maintained on the novel song type stimuli for several sessions until performance again reached an asymptotic level.

Partial transfer to synthetic song sequences

We used partial transfers to the synthetic song sequences and control songs to examine the role of syntax as a mechanism for vocal recognition. Recall that the synthetic song sequences were composed using first-, second- and third-order Markov models derived from each individual’s natural songs, and as a result encoded systematically the natural predictability of song type sequences for each male. The synthetic song bout sequences were not only informative in testing the role of syntax for vocal recognition, but in making predictions about the level of syntactic information that we expected to be most important for recognition. Because the majority of uncertainty is accounted for by the second-order transition probabilities in song (see Fig. 3), we predicted that disrupting the second-order transitions should be most detrimental to individual recognition.
After an individual’s performance on the novel song type stimuli was at or above the 75% criterion for three or more consecutive blocks or trials, the novel song bout and baseline stimuli were reintroduced to the discrimination. This combined stimulus set contained all of the exemplars from the baseline and both transfer sets for that specific stimulus configuration. Now, each bird was exposed to 48 different natural song exemplars (24 INDIV stimuli and 24 MULT stimuli) during a single session. After performance on the combined stimulus set reached an asymptotic level, the rate of reinforcement for correct responses was lowered from 100 to 80%. Once performance again reached asymptote, we began partial transfer sessions with the synthetic song sequences.

For partial transfer sessions, each subject continued to classify the 48 natural song stimuli but on 10% of the trials a synthetic song sequence or one of the control songs was presented. Responses to the synthetic sequences were never reinforced, and therefore subjects did not directly learn to associate them with either class of stimuli (INDIV or MULT). Presumably then, if subjects tended to associate the synthetic song stimuli with one or the other class of the natural song stimuli, they did so on the basis of strategies used to classify the natural song bout stimuli. Recall that we made six different synthetic song sequences for each of the five singers that we recorded, and that the task for each subject was to discriminate one of those singers from the other four. The synthetic song sequences presented to a given subject during the partial transfer sessions were those sequences derived from the singer that that subject was ‘identifying’. Differences in responding between the random and ordered synthetic sequences can be used to infer the relative importance of song type order within a bout for vocal recognition in European starlings. If the temporal sequence of song types within a bout is important for individual recognition, then one would expect to see significant differences in the way that the different sequences of synthetic stimuli are classified. If the temporal sequence of song types within a bout is not important, then all of the synthetic stimuli should be classified in the same way. The subjects were maintained on the partial transfer sessions until each bird responded at least 20 times to each of the six different synthetic sequences and two control songs. The overall rate of reinforcement was 80% throughout all the partial transfer sessions.

Analysis

For each subject, we recorded a single response (or lack thereof) and the stimulus presented on every trial, then analysed these data in blocks of trials. The response following each stimulus presentation could be coded as either ‘correct’ or ‘incorrect’ depending upon the particular stimulus event presented on that trial. Alternatively, responses could be considered independently from the stimulus and coded as either INDIV or MULT depending upon the key that was pecked. Performance could thus be expressed as the probability of responding correctly to any stimulus event, or as the probability of making either an INDIV or MULT response to any stimulus event. We examined each subject’s performance during baseline, novel song bout and novel song type transfer sessions in blocks of 64 trials. We analysed each subject’s performance during the partial transfer sessions as a single block of trials that contained all of that subject’s responses to the synthetic sequences and control stimuli. We used chi-square to assess performance within a given block with respect to chance. We used a factorial analysis of variance (ANOVA) to test for differences between groups in their acquisition rates and asymptotic levels of performance on the baseline task. To assess the transfer to novel stimuli, we compared mean performance over the five blocks of trials prior to transfer with performance on the first block of trials after transfer using a repeated measures ANOVA. The repeated measures design allowed us to account for individual differences in performance. Responses to the different synthetic sequences and control songs were normalized by dividing the probability of making an INDIV response to a given test stimulus by the probability of making an INDIV response to the INDIV baseline stimuli, and then pooled according to the type of stimulus exemplar (first order, second order, third order, INDIV control, or MULT control). All percentage scores were arcsine transformed to correct for any deviations from normality (Zar 1984). The F values reported in the text reflect this arcsine transformation. Identical ANOVAs were done on raw percentage scores and yielded similar results. For ease of interpretation the figures show raw performance data. Data from correction trials and trials for which the subject made no response were not included in any of the analyses.

RESULTS

Baseline Discrimination

All the subjects learned the baseline task by accurately discriminating the songs of an individual starling from the songs of four other starlings. Figures 4 and 5 display the asymptotic performance and acquisition data for the baseline stimulus set. At asymptote (Fig. 4), the probability of making a correct response to any baseline stimulus exemplar was significantly greater than that expected by chance ($\chi^2_{2,11}=2447$, $P<0.001$), and the subjects made reliably different responses to the two (INDIV and MULT) classes of baseline stimuli ($F_{1,11}=159.8$, $P<0.0001$). There were no significant differences between the three groups in their mean asymptotic performance on either class of baseline stimuli ($F_{2,11}=0.34$, NS for INDIV exemplars; $F_{2,11}=2.38$, NS for MULT exemplars; Fig. 4). Once performance reached asymptotic levels, there were no significant differences among any of the five different baseline stimulus configurations ($F_{4,9}=2.24$, NS).

The mean acquisition rates did not differ significantly between the three groups ($F_{2,11}=0.66$, NS; Fig. 5a); but acquisition was more rapid for some of the stimulus configurations than for others ($F_{4,9}=6.72$, $P<0.01$). This large variation in acquisition rates between the five stimulus configurations probably precluded the identification of significant differences in acquisition between
groups. However, the bird's-own-song subject tended to be the first within each stimulus configuration to reach the criterion (Fig. 5b).

**Transfer to Novel Song Bout Stimuli**

All of the subjects showed immediate transfer to the novel song bout stimuli (Fig. 6). The probability of making a correct response during the first block of novel song bout stimuli was significantly above that expected by chance for all three groups ($\chi^2_{p,4}=154.12, P<0.001$ for bird's-own-song; $\chi^2_{p,4}=113.93, P<0.001$ for male-familiar-song; $\chi^2_{p,3}=116.04, P<0.001$ for female-unfamiliar-song); and the probability of making a correct response did not change between pre- and posttransfer trials ($F_{1,11}=1.10$, NS). Performance on the novel song bout transfer stimuli did not differ significantly among the three groups of subjects ($F_{2,11}=1.56, NS$, see Fig. 6), or among the five stimulus configurations ($F_{4,9}=1.64, NS$, data not graphed).

Transfer performance was equally robust for both the INDIV and MULT novel song bout stimuli. Transfer from the baseline INDIV stimuli to the novel song bout INDIV stimuli produced no significant changes in performance for any of the three groups of subjects ($F_{1,4}=0.01, NS$ for bird's-own-song; $F_{1,4}=0.01, NS$ for male-familiar-song; $F_{1,3}=0.001, NS$ for female-unfamiliar-song). Similarly, transfer from the baseline MULT stimuli to the novel song bout MULT stimuli produced no significant changes in performance for any of the three groups of subjects ($F_{1,4}=4.24, NS$ for bird's-own-song; $F_{1,4}=1.66, NS$ for male-familiar-song; $F_{1,3}=3.56, NS$ for female-unfamiliar-song).

**Transfer to Novel Song Type Stimuli**

The transfer to the novel song type stimuli was not as robust as the earlier transfer to novel song bout stimuli. The subjects were able to classify correctly the novel song type stimuli, but post-transfer performance dropped significantly from pretransfer levels ($F_{1,11}=23.86, P>0.001$; Fig. 6). Despite the drop in performance between the pre- and post-transfer trials, the probability of making a correct response during the first block of novel song type stimuli remained above that expected by chance for all three groups ($\chi^2_{p,4}=18.87, P<0.001$ for bird's-own-song; $\chi^2_{p,4}=61.77, P<0.001$ for male-familiar-song; $\chi^2_{p,3}=73.08, P<0.001$ for female-unfamiliar-song). The drop in post-transfer performance was significant for the bird's-own-song group ($F_{1,4}=15.83, P<0.05$), but not for the other two groups ($F_{1,4}=4.54, NS$ for male-familiar-song; $F_{1,3}=9.71, NS$ for female-unfamiliar-song) unless combined ($F_{1,8}=9.28, P<0.05$). Performance on the novel song type transfer stimuli did not differ significantly among the five different stimulus configurations ($F_{4,9}=2.47, NS$, data not graphed).

Patterns of responding to both the INDIV and MULT novel song type transfer stimuli were similar. Transfer from the novel song bout INDIV stimuli to the novel song type INDIV stimuli produced a significant drop in overall performance ($F_{1,11}=20.96, P<0.001$) that reflected a
significant drop in performance for the birds own-song group \((F_{1,4}=10.19, P<0.05)\), but not for the other two groups \((F_{1,4}=6.17, \text{NS} \text{ for male-familiar-song}; F_{1,3}=6.18, \text{NS} \text{ for female-unfamiliar-song})\) unless combined \((F_{1,8}=13.45, P<0.01)\). Similarly, transfer from the novel song bout MULT stimuli to the novel song type MULT stimuli produced a significant drop in performance overall \((F_{1,11}=16.74, P<0.01)\) that reflected a significant drop in performance for the bird’s-own-song group \((F_{1,4}=10.19, P<0.05)\), but not for the other two groups \((F_{1,4}=6.17, \text{NS} \text{ for male-familiar-song}; F_{1,3}=6.18, \text{NS} \text{ for female-unfamiliar-song})\) even when combined \((F_{1,8}=4.70, \text{NS})\).

Partial Transfer to Synthetic Stimuli

All of the synthetic song bout sequences were correctly recognized, but the randomly ordered bouts were significantly more difficult to recognize than the other forms of synthetic bouts. The probability of making an INDIV response to any of the synthetic song bouts was well above that expected by chance for all three groups of subjects \((\chi^2_{6}=197.48, P<0.001, \text{ for bird’s-own-song}; \chi^2_{3}=350.16, P<0.001, \text{ for male-familiar-song}; \chi^2_{3}=154.17, P<0.001, \text{ for female-unfamiliar-song})\), but not all the synthetic stimuli were responded to in the same way \((F_{2,20}=6.18, P<0.01; \text{ Fig. 7})\). The normalized level of responding to the randomly ordered (first-order) synthetic song bouts was significantly lower than that for the second-order synthetic song bouts \((F_{1,10}=8.06, P<0.05)\) and for the third-order synthetic song bouts \((F_{1,10}=10.91, P<0.01)\). We observed no significant difference between the probability of making an INDIV response to the second-order synthetic song bouts and that for the third-order synthetic song bouts \((F_{1,10}=0.56, \text{ NS}); \text{ nor was there a significant difference between groups in the way that each responded to the synthetic stimuli} (F_{2,20}=1.24, \text{ NS})\). All subjects maintained above chance
Individual Vocal Recognition

In field playback experiments, discrimination between two song stimuli is often sufficient to demonstrate individual vocal recognition because the behaviours that one measures are agreed to be functional and as such non-arbitrary. However, the lack of a differential response to field playback stimuli does not necessarily indicate that the receiver failed to perceive a difference between those stimuli. Therefore, the assessment of perception in the field can only be indirect. In an operant experiment, on the other hand, one relies on differences in an arbitrary behaviour to assess perception directly. However, the presence of differential behaviour in an operant context (such as the ability to discriminate between two stimuli) may be the result of an arbitrary difference between the stimuli that is functionally insignificant. This logical constraint is easily overcome by using operant tasks to demonstrate not only simple discrimination between stimuli, but categorization as well, where the latter more strongly implies functional significance.

Given the above considerations, two prerequisites must be met for a demonstration of individual vocal recognition in the laboratory. First, the group of all vocalizations from which an individual’s vocalizations are to be recognized must possess discriminable differences; and second, at least one of those differences must vary in a manner that is nonarbitrary with respect to individual identity. That is, not only must different singers produce vocalizations that are discriminable for the receiver, but those different vocalizations must be perceived (or categorized) as having come from different singers. Reinforcement in this experiment was contingent upon the ability to discriminate between individual vocalizations. Therefore, the fact that all of the subjects were capable of learning to discriminate multiple song bouts sung by a single male starling from song bouts sung by four other male starlings indicates that there must be discriminative differences among the baseline stimuli. Similarly, because the subjects were able to classify correctly the novel song bout stimuli on the basis of classification strategies learned for the baseline stimuli, ostensibly that the songs of an individual are associated with a single key, these strategies must make use of acoustic features that are non-arbitrary with respect to singer identity. Thus, we have evidence for both discrimination and categorization. If the subjects had simply memorized a set of baseline stimulus-key associations that led to high reinforcement, then the introduction of the novel song bout stimuli would require that new associations be learned and initial performance on the novel stimuli would be at chance. This was not the case.

Most conventional field tests of individual vocal recognition have involved the playback of songs from neighbours and strangers to territorial males, and as such rely on aggressive territorial responses to a small set of songs as an assay of perception (Stoddard 1996). As a result, demonstrations of individual vocal recognition in female oscines, nonterritorial oscines outside the parent–offspring context, or in songbirds with large repertoires are relatively rare. Our results present an example of all three. Although the primary function of male song in European starlings appears to be mate attraction (Eens et al. 1991b, 1993; Mountjoy & Lemon 1996), a number of findings suggest that this is not the only function. Male starling song also appears to function in nest site defence (see Eens et al. 1993) and male–male nest site competition (Mountjoy & Lemon 1991), as well as in male dominance hierarchy establishment and maintenance (Hens & Pinxten 1996), as has been observed in female great tits (Lind et al. 1997), dunnocks, Prunella modularis (Wiley et al. 1991) and song sparrows (O’Loghlen & Beecher 1997).

In the absence of territorial aggression, it is difficult to find behaviours that are clearly functional, relatively easy to measure, and closely correlated with song exchanges between individuals. The convergence between the results of the present study and the field work presented above strongly suggest that individual vocal recognition, and thus acoustic communication, is occurring in a number of behavioural contexts that have yet to be well documented by field biologists. Indeed, in the few cases where female responses to playbacks have been examined (as above), individual recognition has been observed. Our results suggest that similar findings should be possible in even more subtle social contexts such as in large roosts, but this awaits future field work. One note regarding this last point is that we observed no sex differences in any of our measures of performance. So, although individual vocal recognition is probably occurring in different social contexts for male and female starlings, no evidence suggests that the underlying psychological mechanisms differ between sexes.
Mechanisms for Individual Vocal Recognition

Individual vocal recognition of conspecific male songs by European starlings appears to rely on at least three perceptual mechanisms: (1) memory for the specific song types that an individual sings; (2) the sequence in which those song types are presented within a bout; and possibly (3) ‘voice’ characteristics imparted to at least some of the song types in an individual’s repertoire.

Memorization mechanisms can be accounted for by individual variation in male song at the level of the song type. That starlings use the memorization of specific song types to recognize individual males can be directly inferred from the results of the two novel song transfer sessions (see Fig. 6). Because performance dropped significantly following the transfer to the novel song type stimuli, but not following the transfer to the novel song bout stimuli, some efficiency in the solution of the recognition task must have been gained through a memory for the specific song types that an individual sings. The composition of song types in an individual starling’s repertoire is a reliable cue for singer identity. In general, the repertoires of male starlings are composed of a unique set of song types (Eens et al. 1989; Chaiken et al. 1993), but song sharing can occur between males engaged in close social interaction (Hausberger et al. 1995). Song sharing between the males we recorded was of the order of 0–3 song types per pair of birds (unpublished data), although all of the birds we recorded for this task lived in very close proximity for several months prior to this experiment.

The notion that individual vocal recognition proceeds through the memorization of specific song types is intuitively parsimonious, but has an implied physiological constraint linked to individual memory capacities. This idea has been articulated as one possible constraint on repertoire size (Falls 1982). However, a well controlled experiment.

...individual variation in male song at the level of the song type. That starlings use the memorization of specific song types to recognize individual males can be directly inferred from the results of the two novel song transfer sessions (see Fig. 6). Because performance dropped significantly following the transfer to the novel song type stimuli, but not following the transfer to the novel song bout stimuli, some efficiency in the solution of the recognition task must have been gained through a memory for the specific song types that an individual sings. The composition of song types in an individual starling’s repertoire is a reliable cue for singer identity. In general, the repertoires of male starlings are composed of a unique set of song types (Eens et al. 1989; Chaiken et al. 1993), but song sharing can occur between males engaged in close social interaction (Hausberger et al. 1995). Song sharing between the males we recorded was of the order of 0–3 song types per pair of birds (unpublished data), although all of the birds we recorded for this task lived in very close proximity for several months prior to this experiment.

The notion that individual vocal recognition proceeds through the memorization of specific song types is intuitively parsimonious, but has an implied physiological constraint linked to individual memory capacities. This idea has been articulated as one possible constraint on repertoire size (Falls 1982). However, a well controlled experiment.
randomly ordered bouts composed of novel song types in order to control for the temporal ‘nonvoice’ cues.

The effect of temporal sequence on the ability to recognize the songs of individual males is demonstrated by the fact that the first-order sequences were significantly more difficult to classify correctly than either the second- or third-order sequences (see Fig. 7). Although the first-order sequences maintained the overall frequency of each song type, the transitions between song types were random. The second- and third-order sequences, on the other hand, accounted for transition probabilities between ordered pairs of song types and ordered triplets of song types, respectively. Our behavioural results fit closely with the information theoretic measures of uncertainty made on the natural song bouts from each of the five recorded males (see Fig. 3), which showed a large drop in uncertainty between the first- and second-order sequences and a much smaller drop between the second- and third-order sequences.

Information functions to decrease uncertainty in the prediction of future events. The more information one has, the less uncertain one is about future events. Thus, the drop in uncertainty (Fig. 3) indicates that a relatively large amount of information can be gained by attending to the transition probabilities between ordered pairs of song types, and that relatively little additional information is to be gained by attending to the transition probabilities among ordered triplets of song types. If one assumes even a monotonic increase in the neural processing load required to monitor transition probabilities in progressively longer strings, then the benefit from the relatively small amount of information gained by attending to anything longer than a second-order string might not outweigh the neurobiological costs. Selection pressures would therefore favour perceptual mechanisms that could attend to second-order transition probabilities and, also, constrain the evolution of higher Markov order sensitivity. This is precisely the pattern we observed (Fig. 7) where subjects appeared to be sensitive to the difference between first- and second-order sequences, but not to the differences between second- and third-order sequences. Thus, perceptual sensitivity of the receiver covaries with the information content of the signal.

The results of the partial transfer sessions are the first reported indication that sequential transition probabilities play an important role in the perception of song. Markov sequence models have been applied to song production in several different species of North American thrushes (Dobson & Lemon 1978), cardinals, Cardinalis cardinalis (Lemon & Chatfield 1971), rose-breasted grosbeaks, Pheucticus ludovicianus (Lemon & Chatfield 1973) and American redstarts (Lemon et al. 1993). In these species, as in European starlings, most song sequences are sequentially patterned communication signals.

The differences among groups in performance on the novel song type transfer stimuli (Fig. 6) are also informative. The bird’s-own-song group performed particularly poorly with the novel song type stimuli, but given their relation to the stimuli it seems unreasonable to propose that these subjects did not recognize the novel song type stimuli as being their own songs. Had they not recognized any of the stimuli as being their own, then their performance should have been similar to the male-familiar-song group and it was not. What is more likely, is that the subjects in the bird’s-own-song group used a different strategy to solve the baseline task than did the subjects in the other groups. The specific nature of the strategy used by the bird’s-own-song group is unknown, but one possibility is that the subjects in the bird’s-own-song group were performing a finer grained discrimination that was more closely correlated with the specific song types in their own repertoire than with singer identity. The particular salience of a bird’s own song has been suggested by the results of perceptual studies (Cynx & Nottebohm 1992), field playback experiments (Falls 1985), and in the electrophysiological responses of cells in song system nuclei (Margoliash 1983). The rapid acquisition of the baseline discrimination for the bird’s-own-song group fits well with this earlier literature.

The fact that we observed no differences between the male-familiar-song group and the female-unfamiliar-song group in any of the transfers (Fig. 6) suggests two further points. First, familiarization with the songs of an
individual starling does not necessarily facilitate later recognition of that individual; and second, both sexes appear equally adept at the task of recognizing individual males. These two points clearly require further experimentation in which the degree of familiarization is controlled between the sexes, but the second point regarding a lack of any sex differences is supported by all of the results in the present study. It is conceivable, given the present design, that female starlings are much more adept at recognizing individual males on the basis of song than are males, but that these differences were obscured by the males’ previous exposure to the songs recorded for use in this experiment. Experience can effect the perceptual salience of specific contact call features in budgerigars (Brown et al.1988), and so it would be difficult to dismiss similar possibilities for the present set of results. Further experimentation regarding sex differences in the individual vocal recognition capabilities of European starlings is currently underway.

Acknowledgments

The research reported in this article was supported by National Science Foundation Research Grant IBN 9317868. We thank Gregory Ball, Scott MacDougall-Shackleton and Michael Beecher for their comments and suggestions on the manuscript. Nate Greenslit lent invaluable assistance in scoring time spectrograms. We are grateful to Aaron Bevington, Evan Ellis, David Fessler, Matt Geschke, Jamie Klimey, Shawna MacDonald and Brendan Reeves for their help in running the experiments.

References


