

D.S. Vicario · J.N. Raksin · N.H. Naqvi
N. Thande · H.B. Simpson

The relationship between perception and production in songbird vocal imitation: what learned calls can teach us

Received: 18 February 2002 / Revised: 9 July 2002 / Accepted: 20 July 2002 / Published online: 15 November 2002
© Springer-Verlag 2002

Abstract Songbirds produce calls as well as song. This paper summarizes four studies of the zebra finch long call, used by both sexes in similar behavioral contexts. Female long calls are acoustically simpler than male long calls, which include acoustic features learned during development. Production of these male-typical features requires an intact nucleus robustus archistriatalis, the sexually-dimorphic source of the telencephalic projection to brainstem vocal effectors. In experiments that quantified the long calls produced in response to long call playbacks, intact adult zebra finch males, but not females, show a categorical preference for the long calls of females over those of males. Experiments with synthetic stimuli showed that males classify long call stimuli that they hear by gender, using both spectral and temporal information, but that females use only temporal information. Juvenile males (<45 days) did not show the categorical preference, but it emerged during the same period when the robustus archistriatalis matures anatomically and the first male-typical vocalizations are produced. Adult males with robustus archistriatalis le-

sions lost the categorical preference for female long calls, suggesting that the robustus archistriatalis plays a role in long call discrimination. These results demonstrate that calls complement song as a potent tool for studying the neurobiology of vocal communication.

Keywords Auditory processing · Brain lesions · Sensorimotor integration · Sexual dimorphism · Vocalization

Abbreviations *DUR* duration · *FF* fundamental frequency · *FFM* fast frequency modulation · *HVC* high vocal center · *LC* long call · *RA* robustus archistriatalis · *TR* trial

Introduction

Studies of the songbird vocal control system have yielded fundamental insights into many biological processes, ranging from behavioral to cellular levels (Catchpole and Slater 1995; Konishi 1985; Kroodsma and Miller, 1996; Nottebohm 1991). The phenomenology of song learning has played an important role in expanding our view of learning processes to encompass specialized learning mechanisms – behavioral adaptations – that may operate differently from classical stimulus-response learning (Thorpe 1958; Marler 1991). Song learning can also be described as vocal imitation. Vocal imitation is known to occur only in a few animal taxa – cetaceans, the human primate, parrots, hummingbirds, and songbirds – and in songbirds we have the best understanding of its neural substrate. Song learning and production are served by a specialized set of brain nuclei that exist in songbirds but not other birds (Nottebohm et al. 1976). These areas are larger in males than in females and their development occurs post-natally, in parallel with vocal development (Bottjer et al. 1985; Konishi and Akutagawa 1985). This system was originally described in canaries, but much recent work on brain-behavior correlations has been carried out in zebra

D.S. Vicario (✉)
Psychology Department, Rutgers University,
152 Frelinghuysen Road, Piscataway, NJ 08854, USA
E-mail: vicario@rci.rutgers.edu
Tel.: +1-732-4452907
Fax: +1-732-4452263

J.N. Raksin
Department of Neurobiology & Behavior,
University of California, Irvine,
Irvine, CA 92697-3800, USA

N.H. Naqvi
Graduate Program in Neuroscience and Department of Neurology,
University of Iowa, Iowa City, IA 52242, USA

N. Thande
University of Michigan Health System,
1500 East Medical Center Drive,
Ann Arbor, MI 48109, USA

H.B. Simpson
New York State Psychiatric Institute,
1051 Riverside Drive, New York, NY 10032, USA

finches, a songbird that is more tractable for many laboratory studies.

Vocal imitation occurs in humans but has not been documented in non-human primates; thus the study of the brain mechanisms of vocal imitation learning in birds is additionally motivated by the possibility of gaining real insights into basic mechanisms that underlie human speech learning (Doupe and Kuhl 1999). Although the intense focus on one particular behavior, song, with a known neural substrate, has been fruitful, we may achieve additional insights from related but complementary approaches. This paper will describe the initial results obtained with one such approach: the study of communication calls. This work builds on our knowledge of the mechanisms of song behavior, but provides a different window on these mechanisms. In songbirds, song is a specialized behavior, produced primarily by males in most species, that serves reproductive and territorial functions. Songbirds, like other birds, also produce non-song vocalizations that serve a wider range of communicative functions. We call these other vocalizations "calls" and they are essential and varied communication signals. In songbirds, some of these calls are learned through a process of vocal imitation similar to that observed for song. However, calls have been much less studied than song in laboratory experiments.

Calls differ from song in several ways that can be exploited experimentally. Calls are produced by both males and females and, in some species, can be shown to be learned through a process of vocal imitation, just like song, and to depend on the same forebrain vocal pathway. Calls can be used to probe vocal behavior in ways that are difficult or impossible to do with song. First, both males and females often produce calls in similar behavioral contexts. So a new range of behaviors can be studied, including the behavior of females, which in many species do not sing. Second, calls are used for a range of communicative purposes, not just courtship and territorial defense. Third, calls generally do not depend on reproductive or seasonal cycles and thus are readily elicited in the laboratory. Fourth, although some calls have learned acoustic features, they are acoustically simpler than most songs, and thus easier to analyze and synthesize. Finally, calls provide an opportunity to study the perceptual side of vocal communication.

Calls are vocal signals that are often exchanged between individuals, and the behavioral response to a call is often another call, not a more covert behavioral or hormonal response. Thus a call "dialog" at least metaphorically resembles a speech "dialog". This is significant because it enables us to study not only vocal production mechanisms, but also the sensory processing associated with call discrimination. Further, the tendency of some calls to be used in vocal exchanges that depend on perceptual as well as motor processing enables the auditory side of vocal communication to be studied in a naturalistic way. A better understanding of

perceptual processing is useful not only for studying communication, but also for studying vocal imitation, because birds must hear, discriminate and remember the vocal sounds that they will subsequently imitate. In addition, the vocal learning that accompanies practice in young birds also depends on auditory processing of their own vocal attempts.

Work in our laboratory has focused on one such call produced in zebra finches, the "long" call, originally named by Price (1979). This is the longest, loudest call in this species and is common in laboratory and field (Zann 1996). The long call is frequently given when a zebra finch hears the vocalizations of another bird, but cannot see it, although it is also used in other behavioral situations (Zann 1996). Exchanges occur between neighbors, strangers, and mates and may influence social affiliation and pair bonding. In the present study, we have explored these vocal interactions by using natural call response behavior (rather than an operant) to test the role of gender-typical features of the long call (LC) in call discrimination. Zann originally documented a clear sex difference in the LC and showed that this depended on learning in males (Zann 1984, 1985). As a result of the learning process, LCs, like song, are species-typical, but are also virtually unique to each individual male. The LCs of females show idiosyncratic differences, but these have not been shown to reflect learning. Production of the LC in males depends on the same specialized telencephalic pathway (Simpson and Vicario 1990) originally shown to be necessary for song production (Nottebohm et al. 1976).

This paper will summarize basic results from four types of experiments in zebra finches. First, lesion studies that assessed the contribution of the telencephalic song control pathway to LC production will be described. Second, studies using a behavioral paradigm that quantifies the LC responses made by isolated male and female zebra finches to playbacks of the LCs of other birds will be presented. The results reveal many of the basic acoustic parameters that birds use to distinguish calls, and also show that male and female birds process the same signals very differently. Third, recent studies of the call responses of young birds during the course of vocal development show that the discrimination behavior of normal adult males emerges with a developmental time course that parallels vocal motor development. Finally, experiments in which brain lesions were combined with behavioral assessment will be presented. The results show that male-typical discrimination of LCs by adult male birds depends on an intact vocal control pathway. This unexpected observation suggests that vocal control nuclei may contribute to sensory as well as motor aspects of vocal behavior. Together, these results document the salient features of the LC, and relate them to brain mechanisms of auditory discrimination, vocal production and processes of sensory to motor transfer. Obviously, the generality of the conclusions is unknown, since they are limited to observations on just one call in one species.

Nonetheless they exemplify the unique strengths of this approach.

Materials and methods

Experimental animals and behavioral paradigm

Adult male and female zebra finches (*Taeniopygia guttata*) raised in our laboratory breeding colonies were used in the studies described. Long call (LC) discrimination behavior of all subjects was tested using a paradigm (Vicario et al. 2001a) in which LCs were played back to an isolated zebra finch and the bird's vocal responses were measured. Briefly, each bird was isolated overnight in a Lucite sound attenuation chamber exposed to natural light. One hour after sunrise, a set of previously recorded natural LCs of 12 female and 12 male conspecifics were played back through a small speaker in a pseudo-random order, under computer control. These LC stimuli were unfamiliar to the subjects. Ten presentations of each stimulus were included in a stimulus set, yielding a total of 240 stimuli. Inter-stimulus intervals were randomized from 10 s to 20 s; playback of the complete set lasted about 1 h. Vocal responses identified as LCs were quantified as both the probability and quantity of response to each stimulus (see below).

Synthesis of altered and "schematic" LC stimuli

Modification of natural calls was accomplished using a procedure developed by Kim Beeman and Evan Balaban, implemented as a set of commands in the SIGNAL programming language (Engineering Design). Briefly, for each natural LC, the contour of the fundamental frequency in time was determined from the sonogram. This contour was used to derive contours for the frequency and amplitude of all the harmonics in the LC; non-harmonic components are lost in this process. This set of contours was then used to re-synthesize each of the harmonic components as a time signal varying in amplitude and frequency. These in turn could be summed to produce a time signal for playback. By manipulating the contours arithmetically before re-synthesis, we produced sounds that were altered in duration or frequency to test the importance of particular acoustic parameters.

To have complete control of each stimulus parameter, we synthesized schematic versions of call stimuli, using simple mathematical functions to create frequency contours rather than deriving them from natural LCs. Acoustic parameters of these schematic stimuli were varied in equal intervals over the range seen in our natural LCs. We used duration (DUR) values of 150 ms, 190 ms, 230 ms, 270 ms, 310 ms, and 350 ms, and fundamental frequency (FF) values of 450 Hz, 550 Hz, 650 Hz, 700 Hz, 750 Hz, 850 Hz, and 950 Hz. Stimuli with each possible parameter combination were synthesized. For each DUR, we also synthesized stimuli with initial fast frequency modulations (FFM), starting from FFM frequencies (FFM0) of 650 Hz, 850 Hz, 1650 Hz, 3350 Hz, and 5050 Hz. For each of these the FFM0 fell exponentially to 650 Hz according to a function derived from natural FFMs with a wide range of different FFM0s.

Behavioral data analysis

Data were collected over a 5-s period beginning 1 s before the beginning of each stimulus. Each of the 5-s samples was displayed and all responsive vocalizations that occurred more than 50 ms and less than 3 s after stimulus onset were scored by using the display cursor to determine their latency and duration. In this paradigm, by far the most frequent call produced was the LC, and the present analysis is limited to that type. The LCs of adult males are easily identifiable on the basis of male-typical characteristics, usually including at least one of the following features: short, stable duration

(typically <200 ms), elevated fundamental frequency (typically > 550 Hz), and the presence of a FFM (Zann 1984; Simpson and Vicario 1990). In previous experiments, the LC of adult females has been identified as a loud call with a DUR > 125 ms (Simpson and Vicario 1990; Vicario et al. 2001a). In this paradigm, some birds also produce a small number of "medium" calls. Medium calls are simple harmonic stacks typically 60–110 ms long. In males, they are easy to distinguish from LCs because they lack FFMs or elevated FF. In females, there are sometimes ambiguous cases, so an arbitrary threshold of 125 ms was used to distinguish medium from LCs. We also used these criteria to identify the LCs of juvenile male and female zebra finches. To identify the LCs of the youngest male zebra finches who produce calls without the adult male-typical LC acoustic features, we used the criteria for adult female LCs. In birds that received central lesions, as described below, call morphology was affected. Ambiguous calls produced by these birds were classified as LCs if they had a DUR > 125 ms, following the work with female calls (Simpson and Vicario 1990). In all cases, we used sound spectrograms and histograms of the distribution of durations to aid in distinguishing LCs from other calls.

Responses to each of the 24 stimulus calls were quantified in two ways: (1) as the number of trials (TRs) out of the 10 presentations of each call where one or more LC was produced (nTRs), and (2) as the total number of LCs produced across all 10 presentations of a given stimulus (nLCs). The gender preferences of individual subjects were assessed by comparing nTRs (or nLCs) to the set of female stimuli with nTRs (or nLCs) to the set of male stimuli in each bird with the Mann-Whitney *U*-test. These measures were also used to calculate individual preference ratios (nTR to female/nTR to male stimuli or nLCs to female/nLCs to male stimuli). Differences in gender preference ratios before and after lesions were assessed by comparing the individual preference ratios with the Wilcoxon matched pairs or Mann-Whitney *U*-test, as appropriate, because distributions of ratios do not satisfy the assumptions of parametric tests.

Individual birds varied in their absolute rates of calling. In order to enable comparisons between subjects, the raw counts of call responses to each stimulus were normalized by dividing the nTRs given to each stimulus call by the nTRs to all the stimulus calls in each subject. This normalization procedure produced a percentage (%TRs) for each unique stimulus. If all 24 stimuli received an identical response (nTRs), each stimulus would receive a %TRs of 4.17 (100/24). Similarly, the nLCs to each stimulus call was divided by the nLCs to all the stimulus calls in each subject to produce a %LCs for each unique stimulus. These transformations allowed us to measure the relative efficacy of each stimulus call in eliciting LC responses across the subject population.

Relationships of the response measures to stimulus parameters (e.g., DUR) within subject groups were analyzed using standard correlation and linear regression procedures. Regression coefficients were compared using the *F*-test for difference in slope and the *t*-test for difference in position. Group differences in the various relationships were assessed with the ANOVA or ANCOVA, as appropriate. In order to satisfy the assumptions that underlie parametric significance tests, percentages (%TRs, %LCs) were first transformed using the arcsine method. General data analysis, plotting and statistics were carried out using Excel (Microsoft) and Statistica (StatSoft). The criterion for statistical significance was set at $P < 0.05$, two tailed.

Surgery, lesions, and histology

Each bird was anaesthetized with Nembutal (50 mg kg⁻¹ with additional doses of 10 mg kg⁻¹ as needed) and prepared for physiological recording by immobilizing the head in a stereotaxic device, and opening the skull over the appropriate brain region (for a full description see Vicario et al. 2001b). The location of the robustus archistriatalis (RA) was determined by mapping sites that showed the characteristic electrophysiological profile for RA: units with large action potentials and high rates of activity (20–30 Hz). Electrolytic ($n = 19$) or ibotenic acid ($n = 2$) lesions were then di-

rected to these identified sites. Electrolytic lesions were made by passing 30–60 μA of current for 30 s at six sites, typically at two depths in each of three penetration tracks. Ibotenic acid lesions were made by injecting 50–100 nl of 2% ibotenic acid solution through a micropipette. The skull opening was closed and the bird was then monitored during recovery.

After final behavioral testing, the bird was sacrificed with a lethal overdose of anesthetic (Nembutal, 200 mg kg^{-1}), and perfused with 0.9% saline followed by 4% formol-saline. After fixation, the brain was removed, cut into 50- μm parasagittal sections on a Vibratome, and processed for cresyl-violet histology. In each brain hemisphere, lesion size was quantified on video images of multiple evenly-spaced sections using NIH Image software. The RA is surrounded by a distinct fiber capsule, whose perimeter was traced to provide an estimate of the intact volume of RA on each measured section. The damaged area was then traced. Lesion size was calculated as the percentage of the RA lesioned across all the sections examined. The location of damage to areas outside the RA was also recorded.

Results and discussion

Description of the zebra finch LC

Zebra finches produce a number of different types of calls, but the LC (also known as the “distance call”) is the loudest, longest call they make (Price 1979). Male and female birds use LCs in the same behavioral contexts (Zann 1984). The most common of these is when two birds can hear but not see each other, apparently to establish or maintain contact (Blaich et al. 1996; Zann 1996). However, the LCs of the two sexes differ in form as the result of learning in males, but not in females (Zann 1984, 1985). Zann described three acoustic features that distinguish the LCs of males from the simpler LCs of females: (1) shorter, much more stable duration, (2) higher fundamental frequency, and (3) presence of fast frequency modulations (Zann 1984; see also Simpson and Vicario 1990). The LC of every male contains at least one, and often two or three, of these features; each feature can vary parametrically, with the result that the LC of each male is virtually unique. Although they are unlearned, the LCs of individual females contain subtler idiosyncratic features that also differentiate them acoustically.

Male LCs resemble song syllables and many males include a version of their LC as a song syllable. The general phenomenology of LC learning is similar to that described for song, although there are suggestions that the sensitive period for LCs may be slightly later than that for song in zebra finches (Zann 1990). In addition, in a case where more than one adult tutor is available, young male can choose different adults as their song and LC tutors (Simpson and Vicario 1996).

Contribution of the vocal control pathway to learned features of the LC

We undertook to investigate whether these learned LCs also depend on the telencephalic song control pathway, originally shown by Nottebohm to be necessary for song

production and vocal learning in songbirds (Nottebohm et al. 1976). In zebra finches this pathway is sexually dimorphic. The telencephalic nuclei high vocal center (HVC) and RA are dramatically larger in males than in females, and the projection from HVC to RA appears not to be present in females (Simpson and Vicario 1991). The projection from RA to areas of the brainstem that control respiratory and syringeal vocal effectors is the final common output pathway from the telencephalon for vocal production. We recorded LCs from 11 male and 6 female birds before and after bilateral lesions of the vocal control nuclei RA or HVC. Similar effects were obtained with HVC and RA lesions. These lesions dramatically affected the acoustic structure of the learned LC, but did not abolish it (Simpson and Vicario 1990). Males with bilateral lesions of RA, for example, continued to produce LCs in the correct behavioral context, but their LCs no longer possessed male-typical features. They became completely feminized (Fig. 1). In females, bilateral RA lesions had no significant effect on the LC.

These results suggest that the smaller vocal control structures in the female telencephalon make no essential contribution to the structure of female calls. In males, these structures are necessary for the production of learned male-typical features, but not for call production per se. This leads to the hypotheses that (1) brainstem

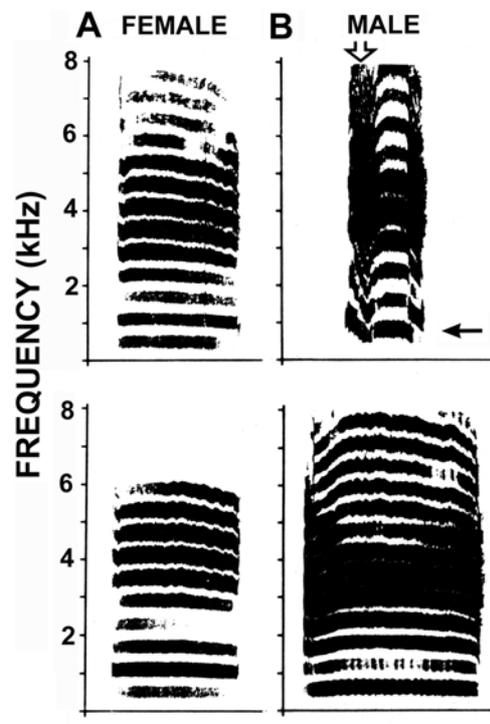


Fig. 1A,B Effect of RA lesions on female and male long calls (LCs). **A** LCs of one female before (*top*) and after (*bottom*) bilateral robusus archistriatalis (RA) lesion, showing no change in the call. **B** Long calls of one male before (*top*) and after (*bottom*) bilateral RA lesion, showing loss of the fast frequency modulation, FFM (*open arrow*), the lowered fundamental frequency, FF (*solid arrow*), and the increase in duration. Time bar: 150 ms (reprinted with minor modifications from Simpson and Vicario 1990)

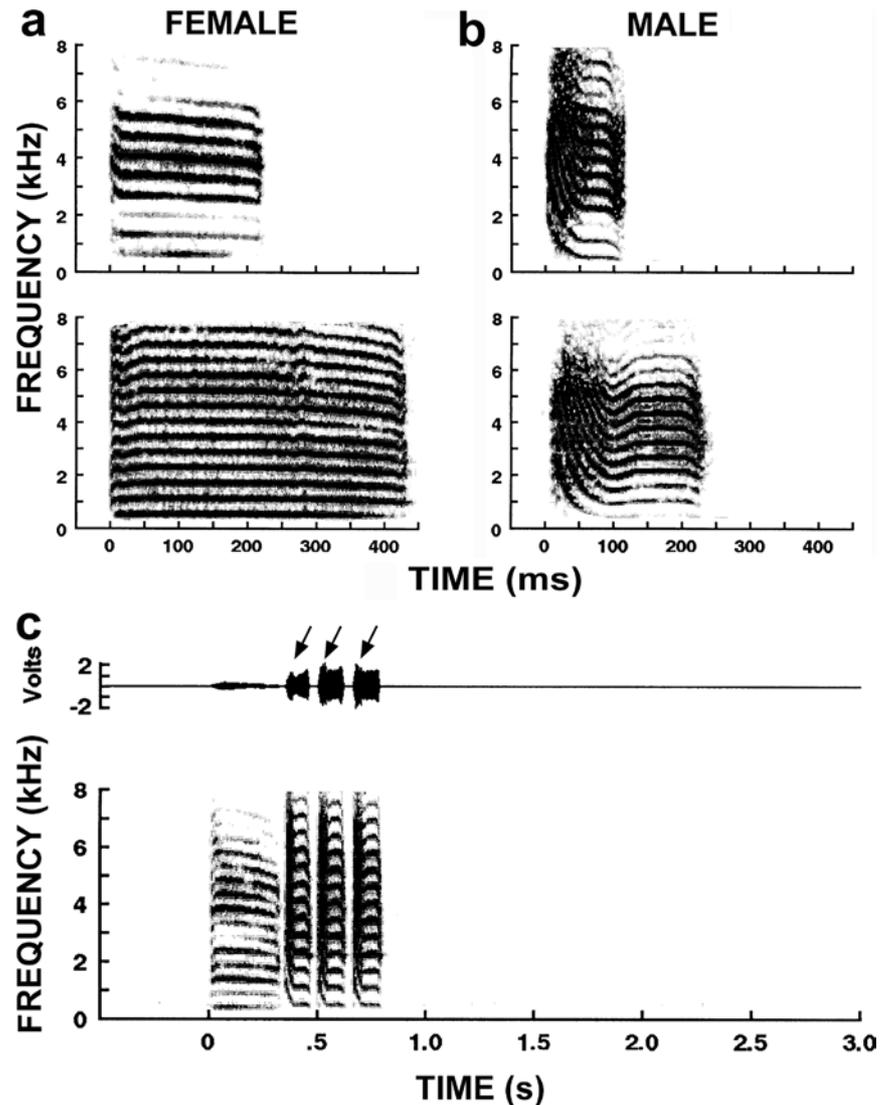
vocal pathways are responsible for the production of the female-type LC in females and in RA-lesioned males, and (2) the learned LCs of males are produced through modification of the female-type call by the action of the telencephalic vocal pathway.

Behavioral discrimination of natural LCs in adult males and females

The process of vocal imitation leads to individual differences in male calls because of the incorporation of different types of male features. As a result, the calls of individual males are virtually unique. Female LCs do not show evidence of vocal imitation, but nonetheless tend to differ in subtle ways between individuals. We were able to take advantage of these forms of natural variation in LCs and the tendency of birds to call back to LCs that they hear to develop a behavioral task that assesses LC discrimination. In this “call response paradigm”, an isolated subject bird hears one of a set of

24 natural LC stimuli (Fig. 2). LCs are the most frequent vocalizations produced in this behavioral situation, although some birds also produce a small number of “medium” calls (see Materials and methods). The analyses presented here are based on the LC data and summarize published results obtained in 21 adult male and 20 adult female zebra finches (Vicario et al. 2001a). The overall level of responding did not differ between males and females. However, some stimuli elicited greater responses than others; female LCs were more effective at eliciting LC responses than male LCs for both male and female subjects. We measured responses in two ways (see Materials and methods): as the number of stimulus presentations (trials) with LC responses (nTRs), which measures the probability of responding to a stimulus and as total number of LCs (nLCs), which measures the amplitude of the response to the stimulus. These gender preferences were significant for both male and female subjects and could be quantified as female preference ratios (nTRs to female divided by nTRs to male stimuli; nLCs to female divided by nLCs to male

Fig. 2 Sound spectrograms of representative male (**A**) and female (**B**) LCs from the stimulus set display the constituent frequencies of each call as a function of time. Male LCs differ from female LCs in one or more of the following ways: shorter duration (DUR), higher FF, and presence of FFMs (see text). Males learn these acoustic features during development and male LCs resemble song syllables. The simpler LCs of females are unlearned. **c** Example of stimulus presentation and LC responses: a female stimulus call (at time zero) elicits three LC responses from a male subject. *Arrows* indicate the amplitude peak associated with the FFM on each response call in this subject. *Upper traces* show the recorded amplitude waveform and *lower traces* show the sound spectrogram for each trial. The latencies and DURs of all responsive calls in the first 3 s after stimulus onset were measured (reprinted with minor modifications from Vicario et al. 2001a)



stimuli). On average, males responded twice as much to female calls as to male calls (for nLCs F/M=2.45; for nTRs F/M=1.94). Females had weaker average ratios (for nLCs F/M=1.40; for nTRs F/M=1.23). The preference ratio for both nLCs and nTRs was significantly greater for males than for females.

In order to determine the basis for the preferential response to female calls, we then carried out a detailed analysis of the gender-typical acoustic features that might be used by the birds to discriminate male and female LCs. We analyzed the responses to each male and female stimulus as a function of the principal acoustic features that distinguish male and female LCs, originally described by Zann (1984): DUR, FF, and presence of FFM.

Longer stimuli received larger responses: response magnitude was positively correlated with DUR in both males and females. However, the relationship differed qualitatively between the sexes (Fig. 3). The responses of female subjects to the male and female stimuli appear to be largely proportional to DUR over the whole range, independent of stimulus gender. In contrast, males respond more vigorously – there seems to be a step increase in responding (Fig. 3A) – to all female LCs, even those that overlap in DUR (between 220 ms and 255 ms) with male LCs. The pattern of responses in the

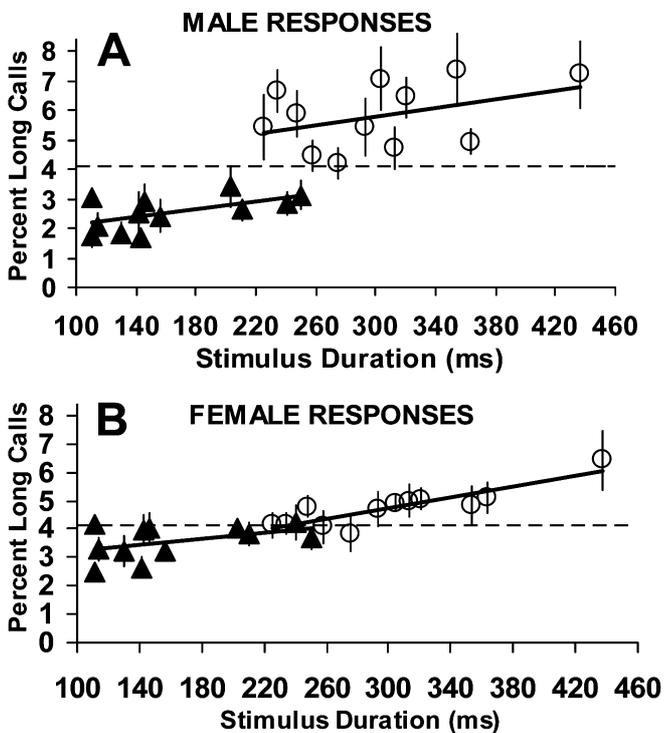


Fig. 3A,B Relationships between stimulus DUR call responses for male and female zebra finches hearing natural stimuli. **A** Mean %LCs (\pm SE) as a function of stimulus duration for male (solid triangles) and female (open circles) stimuli in male subjects. **B** Mean %LCs as a function of stimulus duration for responses in female subjects. If all 24 stimulus calls had received equal responses, the response percentage for each stimulus would have been 4.17 (1/24), indicated by the dashed line (reprinted with minor modifications from Vicario et al. 2001a)

two sexes was tested by analysis of covariance, including DUR as a covariate, for both nTRs and nLCs. In both cases, stimulus gender and DUR were highly significant factors, as was the interaction between stimulus gender and subject sex. The strong interaction between stimulus gender and subject sex suggests that male and female birds are processing male and female calls differently. For females, DUR seems to predict responses independent of the gender of the caller. Males show a much stronger call preference than females because they seem to place the calls of males and females into separate categories, presumably using acoustic cues in addition to duration.

A further indication of factors that distinguish the stimuli besides DUR came from examination of the relationship between DUR and response magnitude seen in separate regression analyses for male and female stimuli. In contrast to the responses of females, the responses of males to different female stimuli showed a high degree of variance around the regression on duration (Fig. 3). Inspection of the other acoustic features of these stimuli suggested that stimuli that received responses below the regression line tended to have higher fundamental frequencies and those above the line tended to have lower fundamental frequencies. In this stimulus set of natural female LCs (and in a larger group of LCs we examined), DUR tended to be related to FF over a certain range, so it was difficult to isolate an independent role for FF in the responses of males. In addition, the range of FFs was narrow especially for the female calls. These limitations led us to implement two different procedures (see Materials and methods) for making synthetic copies of natural male and female LCs that enabled us to independently manipulate duration and frequency parameters.

Males classify synthetic LCs by duration and fundamental frequency

In an initial experiment, we used digital techniques to produce variants of a natural call that differed systematically in DUR and FF. Zebra finches responded to these synthetic LCs at rates comparable to those seen with natural calls. The overall pattern of responding was the same as that seen with natural calls. When the ensemble of these data in the two sexes was tested by analysis of variance, as was done for the natural stimuli, both stimulus gender and DUR were significant factors, as was the interaction between stimulus gender and subject sex. Female preference ratios were significantly higher in male than in female birds and only males showed a step in the correlation with DUR.

The behavioral results obtained with these stimuli in males suggested that, in addition to the linear relationship of response level to stimulus DUR, there was a consistent but non-linear relationship to stimulus FF, especially for female calls, and that there was an interaction between these two factors (not illustrated). In

females, as expected, only the relationship to DUR was confirmed. The non-linear relationship seen in males was difficult to analyze because of the limited range of FF in the female stimuli, although the quadratic regression for these data was weakly significant. Responses to female stimuli in female subjects showed no trend; the quadratic regression was not significant. There was no significant relationship for male stimuli in either male or female subjects. These results were suggestive, but not conclusive. To confirm the relative contributions of these stimulus parameters, we needed to be able to test stimuli with a wider range of parameter values than those based on natural LCs, which cannot be excessively modified without distorting harmonic relationships and altering frequency modulations. This led us to simplify our approach by synthesizing highly schematic stimuli from mathematical functions, rather than deriving them from natural LCs.

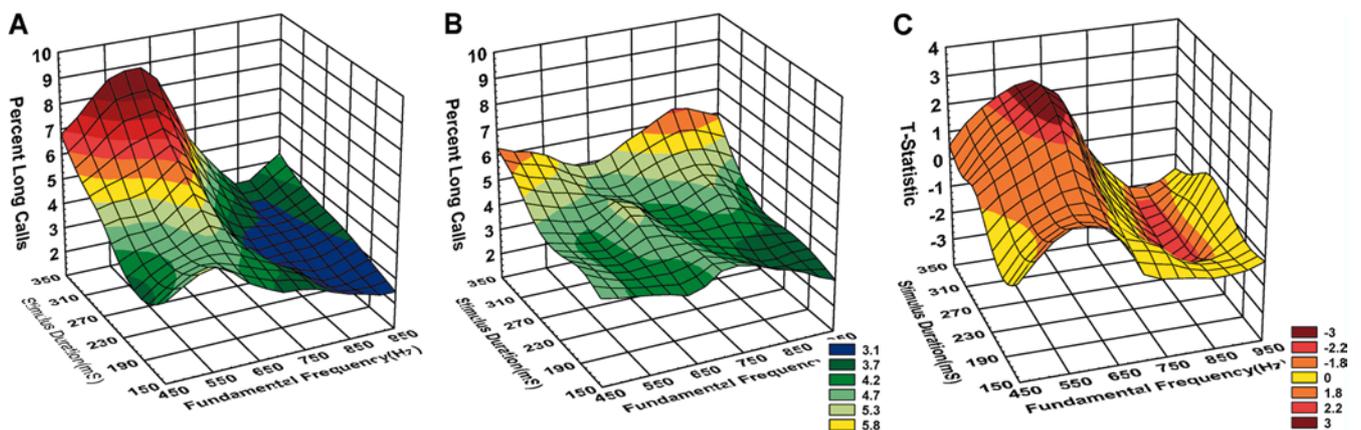
Males process DUR and FF of schematic LCs categorically

We generated a large set of schematic stimuli that varied parametrically in DUR and FF, a subset of which included FFMs with a specified FFM0. Although, to a

human ear, they sound more like a cheap harmonica than a zebra finch, these were effective stimuli. Both male and female birds modulated their response magnitude according to the acoustic features of each stimulus. Qualitatively identical relationships were obtained with both response measures, %TRs and %LCs. In plots of the mean effect of DUR (effectively collapsing the data obtained with different FFs) in males, responses to stimuli shorter than 230 ms showed almost no change with DUR, followed by increases at 270 ms and 310 ms and again little change between 310 ms and 350 ms. Female subjects showed a similar overall pattern. In plots of the mean effect of FF (collapsing the effect of all DURs) in males, responses at 450 Hz, 550 Hz, and 650 Hz showed an increase with increasing FF, followed by a dramatic drop at 700 Hz and again little change at higher FF. In contrast, the responses of female subjects showed little or no relationship to FF. In plots of the mean effect of FFM0 (collapsing the effects of all DURs), only responses for the constant-frequency stimuli at 650 Hz were high; all stimuli with FFMs, even the smallest FFM used (a drop from 750 Hz to 650 Hz) received a dramatically lower response from male subjects. In contrast, females gave the same level of response to all the FFM stimuli as they did to the constant-frequency stimulus.

In order to reveal the interactions between DUR and FF, the same data were plotted as separate 3-D response surfaces for male and female subjects (Fig. 4). In both males and females, the general surface slopes up for longer durations, but over a limited range of FFs the effect of DUR is dramatic in males. A clear response ridge develops as FF increases to the sharply preferred FF (650 Hz), rising steeply to a new level for the longer DURs (Fig. 4A). Responses at higher FFs drop abruptly at all DURs, forming a deep valley. Females also show a rise with DUR, but there is only weak modulation with changes in FF (Fig. 4B). Differences between responses to each stimulus in the two sexes were quantified with the *t*-statistic, plotted as a 3-D surface, and colored according to the probability associated with each value of *t* (Fig. 4C). Males responded significantly more than females to stimuli along the top of the ridge in FF and significantly less to stimuli in the valley from 700–850 Hz. Female calls typically are longer than

Fig. 4A–C Response surfaces in male and female subjects hearing schematic calls. The strength of the vocal response to each stimulus (%LCs) is plotted as a function of the FF and DUR of the stimulus. The response surfaces display a composite least squares fit to the data (Statistica). **A** Males ($n=10$) show a prominent response peak for lower FF and longer DUR and a deep valley in the opposite quadrant. **B** Females ($n=10$) respond more to stimuli with longer DUR but have an almost flat response surface in the FF dimension. Data from multiple sessions have been combined and renormalized as necessary, using fiduciary stimuli included in each set. **C** For each stimulus, male and female data have been compared with the *t*-test. The plotted surface shows the value of the *t*-statistic as a function of stimulus DUR and FF; positive values of *t* indicate higher responses in males; negative values indicate lower responses for males than for females. Color code represents the probability associated with each *t* value: $P < 0.01$ – deep red; $P < 0.05$ – red; $P < 0.1$ – orange. Color legends indicate the minimum response value (% Long calls) coded by each color (reprinted with minor modifications from Vicario et al. 2001a)



200 ms (see Fig. 3) and typically have FFs of 500–600 Hz. They thus tend to correspond to the peak area (red) in the 3-D graph for schematic calls. Male calls tend to have much shorter durations and contain higher FFs, and thus correspond to the valley region (blue) in the 3-D graph.

To test the applicability of these findings to natural LC stimuli, we used the values of DUR, FF, and FFM0 from each natural stimulus to estimate the expected response, using our response data for schematic stimuli with the closest parameter values. These predictions were then compared to the observed responses to the natural stimuli. Predicted responses for male subjects showed a correlation with DUR and a step in responses to female stimuli. They also showed a curvilinear relationship to FF for female stimuli. Overall, there was a highly significant correlation ($P < 0.001$) between predicted and expected responses for the stimuli (Fig. 5) in male subjects and a good correlation ($P = 0.02$) in female subjects. At least in males, differences in the way these parameters are processed appear to account for a very large proportion of the variability observed in responses to natural calls.

In summary, the results show that females respond to stimulus calls as a function of their duration, whereas males respond based on a combination of duration, fundamental frequency and the presence of a fast frequency modulation. In effect, males use multiple stimulus features to produce a categorical discrimination (cf. Nelson and Marler 1989) of the calls into male and female types, and do so reliably.

Call discrimination emerges during development in juvenile males

The different patterns of discrimination produced by adult male and female subjects led us to ask whether this

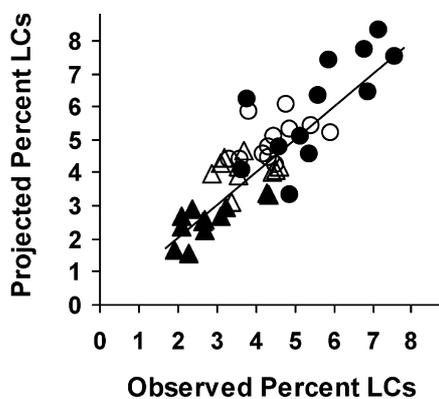


Fig. 5A–C Comparison of predicted and observed responses. **A** Predicted (*open symbols*) and observed (*solid symbols*) responses in male subjects to male (*triangles*) and female (*circles*) stimuli plotted as a function of DUR. **B** Same responses plotted as a function of FF. **C** Scatterplot of predicted versus observed responses to male (*triangles*) and female (*circles*) stimuli for males (*solid symbols*) and females (*open symbols*). The identity line is plotted at 45° (reprinted with minor modifications from Vicario et al. 2001a)

sex difference is always present or whether it emerges in development. To assess the development of the call discrimination behavior observed in adults, juvenile birds were tested using the call response paradigm with the same stimulus set of 24 LCs. We assessed juvenile zebra finches of both sexes ($n = 42$ males, 18 females) at different ages, ranging from 25 days to 86 days old. Although juvenile subjects responded to the LC stimuli at somewhat lower rates than adults, data were obtained from birds as young as 25 days, approximately the age when young birds no longer require parental care. At the younger ages studied, the LC does not have its adult form in either males or females, but is distinguishable from other calls, and could be reliably scored.

We compared the response behavior of these subjects with the adult call response behavior data gathered from the previous study (Vicario et al. 2001b). In 25- to 35-day-old birds, all stimuli received a similar level of responding (Fig. 6A), in dramatic contrast with the adult

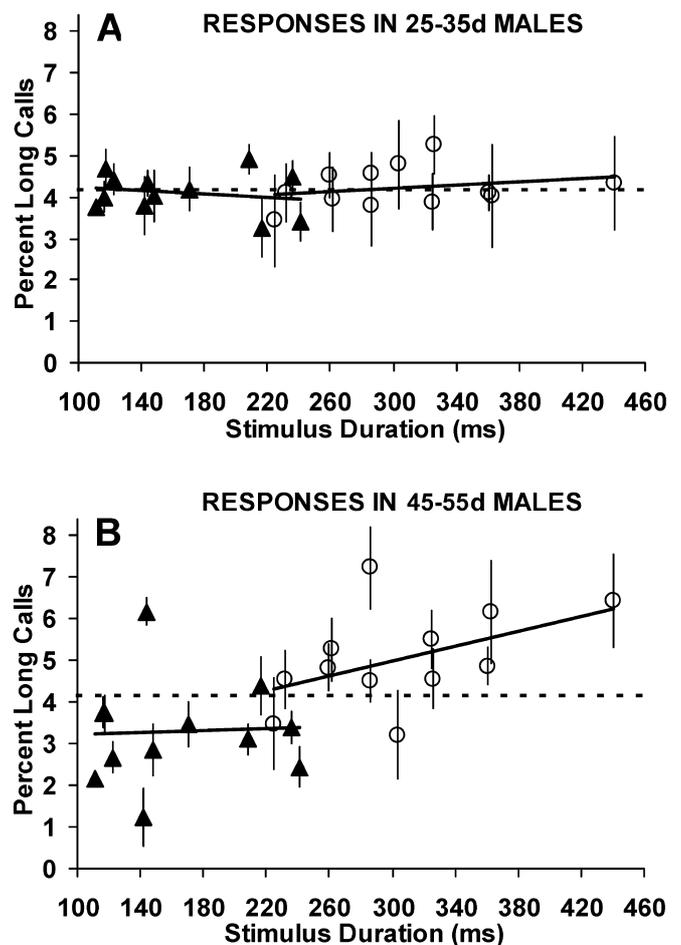


Fig. 6A,B Relationships between stimulus duration call responses for young male zebra finches hearing natural stimuli. **A** For males aged 25–35 days, mean %LCs (\pm SE) as a function of stimulus DUR for male (*solid triangles*) and female (*open circles*) stimuli in male subjects. **B** For males aged 45–55 days, mean %LCs as a function of stimulus duration for responses. Conventions as in Fig. 3

pattern (cf. Fig. 3). After 45 days of age, males start to show increased responding to stimuli with longer DUR. On average both juvenile male and female subjects under 45 days of age responded slightly more to female stimuli than to male stimuli (males: for nLCs $F/M = 1.27$; for nTRs $F/M = 1.16$; females: for nLCs $F/M = 1.29$; for nTRs $F/M = 1.16$). However, for this age group, male and female juveniles did not show a significant difference in call preference ratios. This suggested that the sex difference in call preference ratios seen in adults must appear in development. To investigate this, we tested cohorts of birds of different ages to determine their preference ratios. The results are summarized in Fig. 7. The pattern of responses in the two sexes was tested by analysis of variance using the arcsine transformed preference ratios (nLCs F/M or nTRs F/M) as the response measure, and subject age and gender as factors. Both subject gender and subject age were significant factors, as was the interaction between gender and age. These statistical results confirm that young males undergo a developmental change in their pattern of call response behavior. Very young juvenile male and female birds respond to LCs with low preference ratios. During development, the low female preference ratio of female birds increases only slightly to reach the adult female level (Fig. 7), whereas the female preference ratio of juvenile males increases dramatically with time.

To determine the approximate age when juvenile male subjects begin to exhibit adult male-typical preference behavior we made pair-wise comparisons of the preference ratios of each juvenile age group with the adult male preference ratios (using the Mann Whitney U -test). For both response measures, the preference ratios of young males were significantly different from adult ratios at 25–34 days and 35–44 days, but were no longer different at 45–54 days and all older ages. Inspection of the graph (Fig. 7) suggests that a trend towards a difference may begin before 45 days. We also

determined the approximate age when juvenile male and female preference behavior diverged. In comparisons of the preference ratios of juvenile male and female subjects in the same age cohort, the preference ratios of juvenile male and female birds less than 45 days old do not differ significantly, but begin to diverge significantly thereafter.

Experiments described above showed that adult male birds also use FF as well as DUR to determine their responses to LCs. When the contribution of FF to response level was assessed for juvenile males, it became apparent that the development of male-typical preference ratios was associated with the emergence of a relationship between FF and call responses.

At about the same age as these changes in the call response behavior, young males begin to produce LCs with male-typical features. The first FFMs are typically seen at 40–45 days, and the call becomes less noisy and begins to shorten in the period from 45–55 days of age (D.S. Vicario et al., unpublished observation, cf. Zann 1985). This temporal association leads to the hypothesis that by 40–50 days young males can discriminate the frequency components of the calls they hear, and use those discriminations to guide both the quantity and acoustic quality of the LCs they produce.

Effects of RA lesions on LC discrimination behavior in adult males

The development profile of call preference not only parallels the emergence of male-like vocalization, but also the time when the vocal control pathway matures (Bottjer et al. 1985; Konishi and Akutagawa 1985). Taking this together with the adult sex difference in preference, we speculated that this temporal relationship could reflect the development of neural pathways that might contribute to both vocal production and call discrimination behavior. This led us to carry out experiments in which call preference was assessed in adult males before and after bilateral lesions of the vocal control nucleus RA.

Preoperative testing of call response behavior showed that the experimental males responded more than twice as much to female LC stimuli as to male stimuli, in line with previous observations of normal males (Vicario et al. 2001b). This difference was significant both for the probability of responding, measured as nTRs with at least one LC produced and for the strength of the response, measured as nLCs produced. When tested again 5–7 days after receiving bilateral RA lesions, these same birds no longer showed the large difference in responses to female versus male stimuli. It should be noted that, as described in the first section of the Results, the RA lesions degraded the acoustic form of the LC, even rendering it female-like in the case of the largest lesions. Nevertheless, the LC remained recognizable and could be reliably scored. However, the change in morphology and duration of the call could have affected the ability of

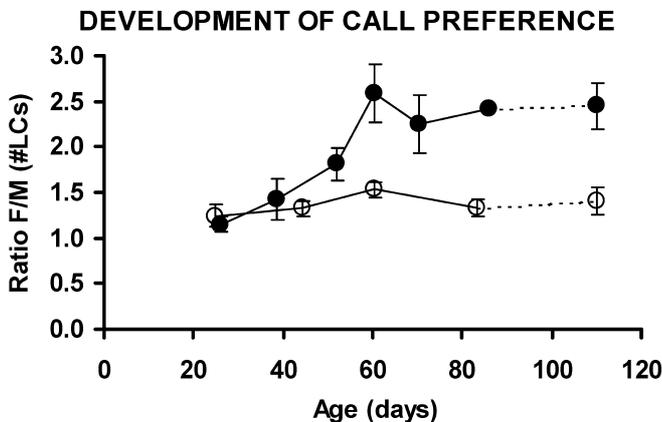


Fig. 7 Call preference ratios for juvenile zebra finches as a function of age. Mean preference ratios (\pm SE) for different age cohorts of males (solid circles) and females (open circles). Ratios for adults have been plotted arbitrarily over the point for 110 days of age. For statistical interpretation see text

the bird to produce multiple calls in succession, so the analyses described here were all performed using the %Trials response measure. This quantifies the probability of a response on a given stimulus presentation, independently of the number of individual calls in the response.

In order to normalize the data for the different absolute rates of responding in different birds, these relationships were analyzed by dividing the total number of trials with LC responses to female stimuli by the total number of trials to male stimuli to form a preference ratio. The preference ratio for nTRs was 2.04 pre-operatively and was significantly lower post-operatively (1.40). Although the lesioned males still showed a significant preference for female calls, these post-operative ratios were not significantly different from the preference ratios of intact adult females determined in previous work. Thus, birds with RA lesions no longer show the male-typical pattern, suggesting that the lesioned tissue may make a necessary contribution to the normal bird's ability to generate vocal responses that show call gender discrimination.

This contribution must be based on some kind of acoustic selectivity, because lesioned birds continue to produce LC responses to the stimuli. In fact, the total number of trials with LC responses was significantly greater after the lesions, indicating that the probability of responding was actually increased. Further analysis showed the number of trials with LC responses given to male stimuli had increased significantly, but that the number of trials with LC responses given to female stimuli had not. The other response measure, total number of LCs produced, showed an increase for male stimuli, a decrease for female stimuli, and a decrease overall, but none of these differences approached statistical significance. Apparently the loss of discrimination behavior, measured as a decrease in the preference ratio, is primarily due to a specific increase in the probability of responding to male stimuli.

The lesion sizes varied from bird to bird. When the preference ratio after lesion was plotted against the mean lesion percent (the mean percentage of RA lesioned on both sides) for the 17 birds with post-operative behavioral data, the correlation was significant for both the nTRs preference ratio and the nLCs preference ratio (Fig. 8). The smaller lesions had variable effects on the preference ratio, perhaps due to subtotal lesions that hit or missed certain critical tissue in RA. Larger lesions were always effective, either because of a volume effect or because they always hit the critical tissue. The former is more likely because the largest lesions ($\geq 55\%$) show a tighter inverse correlation with preference ratio.

Although the lesions extended beyond the borders of RA into the archistriatum in various directions and to varying degrees, the size of these extensions was not correlated with the percentage of RA lesioned. In addition, there was no systematic relationship between the particular area damaged and the magnitude of the behavioral effect in that bird. Data obtained from six birds

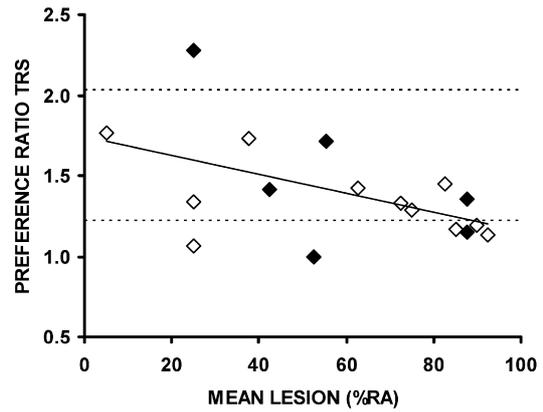


Fig. 8 Female preference ratios in males with lesions of different sizes ($n=17$). Trial preference ratios (quotient of the number of trials to female stimuli divided by the number of trials to male stimuli) after RA lesion plotted as a function of mean lesion size for the two hemispheres in each bird ($n=17$). Upper dashed line shows the mean of preoperative preference ratios; lower dashed line shows the mean preference ratio for females. Solid symbols indicate data from birds in which lesions involved minimal damage to areas outside RA (reprinted with minor modifications from Vicario et al. 2001b)

with lesions that were best restricted to RA (damaging a volume outside RA $<10\%$ of RA's volume) show the same pattern (filled symbols in Fig. 8) as the whole group. A direct test of the contribution of areas adjacent to RA to LC discrimination behavior was attempted, but proved impractical. It was not possible to make control lesions that damaged areas in the archistriatum adjacent to RA without impinging either on RA itself, on the HVC to RA fibers (entering dorsal and caudal RA), the LMAN to RA fibers (entering RA medially) or on the occipito-mesencephalic tract that carries the RA's output (leaving rostro-ventral RA).

The step increase in responding to female stimuli, seen in the intact males, can be largely explained by the pattern of responding to the spectral parameters that distinguish male from female LC stimuli, as described above (see Figs. 3 and 4). In birds with large RA lesions ($>40\%$ of the nucleus), the relationship to stimulus duration no longer showed the sharp increase for female stimuli, because the relationship of responses to spectral parameters was much weaker. Birds with partial lesions ($<40\%$ of the nucleus) showed intermediate patterns. The overall nature of these relationships is most easily visualized as a 3-D response surface that is a function of both stimulus duration and fundamental frequency (Fig. 9). This plot is limited to female stimuli because the presence of FFMs on male stimuli dominates the responses to the FFs of male calls. In the intact males (Fig. 9A), for stimuli with short durations, only a narrow FF band (600–650 Hz) received a high probability of responding. As DUR increased, the band of effective FFs broadened. This highly inflected surface can be contrasted with the almost flat one based on call response data from normal adult females (Fig. 9B). In birds that received smaller RA lesions ($<40\%$ of the

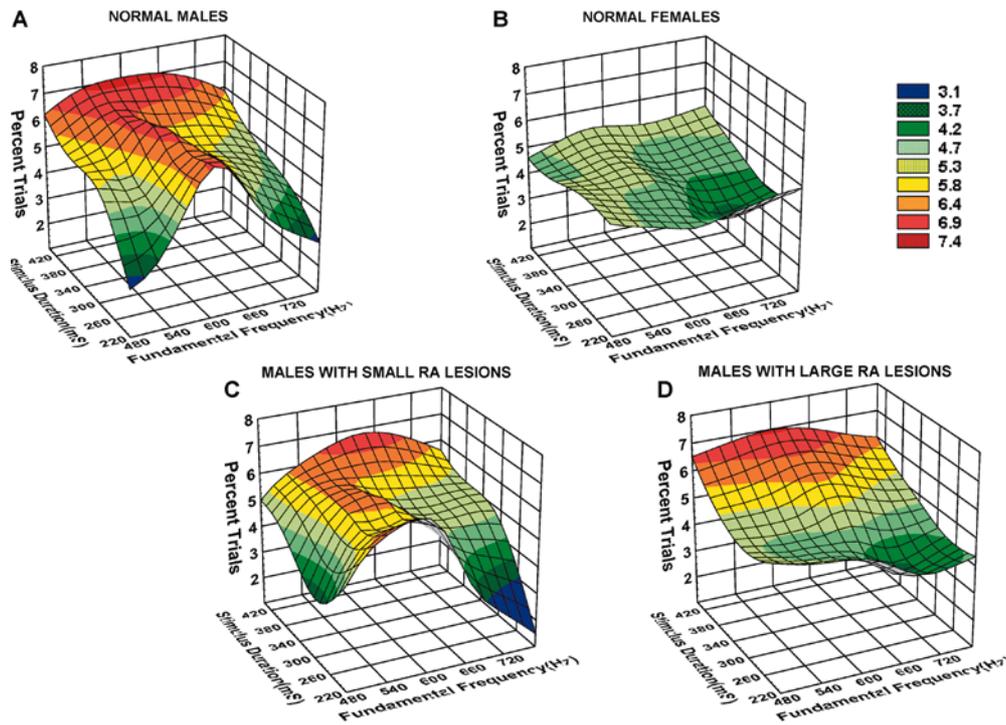


Fig. 9A–D Response surfaces showing response probability in intact and lesioned male birds and female birds hearing conspecific LCs. The normalized probability of a LC response to each stimulus trial (%TRs) is plotted as a function of the fundamental frequency (FF) and duration (DUR) of the stimulus. The response surfaces represent composite least squares fits to the data (Statistica). **A** Intact males ($n=14$) show a prominent response ridge for intermediate FF and longer DUR. **B** Female subjects ($n=10$) respond more to longer stimuli but show almost no variation in the FF dimension. **C** After RA lesion, males with $<40\%$ mean lesion ($n=5$) yield a surface qualitatively similar to that of pre-lesion males, but the FF and DUR slopes are less pronounced. **D** After RA lesion, males with $>40\%$ mean lesion ($n=9$) have a response surface similar to that of females, with no preferred FF and a gradually higher probability of response to stimuli as DUR increases. Color scale as in Fig. 4 (reprinted with minor modifications from Vicario et al. 2001b)

nucleus), the degree of inflection was decreased, but the basic topography of the response surface was unchanged (Fig. 9C). In birds that received larger RA lesions ($>40\%$ of the nucleus), the inflection due to FF was almost completely lost, leaving only the relationship to stimulus DUR, shared with females (Fig. 9D).

Apparently, in males, an intact RA is necessary for normal call discrimination behavior in our paradigm. This result in lesioned adults is consistent with the observation described above that young males also show the female-typical response pattern when RA is immature. Of course, we do not know whether RA itself (1) contributes to sensory discrimination, (2) is where the salience signal that results from discrimination gets transformed into response probability and number of LCs, or (3) whether that transformation occurs earlier and RA merely receives commands to produce one or more calls. However, the non-RA pathway that produces calls in lesioned birds is not privy to the com-

mands that result from discrimination of FF and FFM. One reason to suspect that the salience signal might reach the RA is that, under at least some circumstances, RA neurons do have auditory responses that are selective for conspecific vocalizations (Doupe and Konishi 1991; Vicario and Yohay 1993; but see Dave and Margoliash 1998). Further, there is the interesting fact that the LCs of lesioned males lack the very same spectral features that these birds no longer appear to discriminate. Whether this is purely a coincidence awaits further experimentation, but the possible role for a relationship between production and discrimination mechanisms used for vocal communication has a long pedigree in the study of human speech (Studdert-Kennedy et al. 1970) and may operate in animal vocal systems as well (Williams and Nottebohm 1985; Brenowitz 1991; Margoliash 1997; Pytte and Suthers 1999; Solis and Doupe 2000).

Conclusions

1. The learned LC in males can be thought of as a convolution of learned acoustic features (DUR control, FF, and FFM) with a simpler unlearned LC that is shared with females. Production of the learned features depends on the same specialized vocal control system known to be essential for song in males.
2. In the call response situation, the learned acoustic features that distinguish calls are discriminated and determine whether and how much the listener will respond. This behavioral method tests the natural salience of the stimuli in a particular context, not the birds' ability to discriminate differentially-reinforced stimuli in an operant task. The results reveal some

simple aspects of the grammar of vocal communication in this species, and also demonstrate the power of using a natural behavioral response to quantify the processing of species-specific signals.

3. The acoustic frequency parameters (FF and FFM) are used by males and not by females in determining the probability of responding. The responses of females reflect only stimulus duration and thus do not differentiate male and female calls that overlap in duration. Discrimination of all three acoustic parameters enables males to categorize the calls that they hear by gender very reliably and thus to achieve the perceptual task of gender identification.
4. One further inference from the above is that males' ability to discriminate these features underlies their ability to imitate them during vocal learning. In fact, the adult male call response pattern emerges during development in parallel with the production of male-typical features in the LC.
5. Behavioral discrimination based on acoustic frequency parameters is abolished by lesions of vocal control nucleus RA. This suggests that the RA either is involved in some aspect of sensory processing specific to frequency discrimination and/or is a necessary relay along a pathway that transforms information about the frequency content (but not the duration) of a stimulus into the probability and quantity of responding.
6. In summary, study of learned vocal communication calls can provide new insight into mechanisms of vocal production (Simpson and Vicario 1990), vocal signal perception (Vicario et al. 2001a) and their interaction (Vicario et al. 2001b).

Acknowledgements The experiments described herein were performed under an approved Rockefeller University animal protocol and comply with the "Principles of Animal Care", publication 86-23, reviewed 1985, of the National Institutes of Health and also with the current laws of the United States. Supported by NIH R01-MH-40900.

References

- Blaich CF, Norman M, Syud FA, Benitez G, Frost J, Ravenscroft J, Smith T, Tansinsin S, Ware P (1996) The use of distance calls to maintain pair contact in zebra finches (*Taeniopygia guttata*). *Bird Behav* 11:25–30
- Bottjer SW, Glaessner SL, Arnold AP (1985) Ontogeny of brain nuclei controlling song learning and behavior in zebra finches. *J Neurosci* 5:1556–1562
- Brenowitz EA (1991) Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. *Science* 18:303–5
- Catchpole C, Slater PJB (1995) *Bird song*. Cambridge University Press, Cambridge
- Dave AS, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. *Science* 282:2250–4
- Doupe AJ, Konishi M (1991) Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci USA* 88:11339–11343
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22:567–631
- Konishi M (1985) Birdsong: from behavior to neuron. *Annu Rev Neurosci* 8:125–170
- Konishi M, Akutagawa E (1985) Neuronal growth, atrophy and death in a sexually dimorphic nucleus in the zebra finch brain. *Nature* 315:145–147
- Kroodsma DE, Miller EH (eds) (1996) *Ecology and evolution of acoustic communication in birds*. Cornell University Press, Ithaca, New York
- Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. *J Neurobiol* 33:671–93
- Marler P (1991) Song-learning behavior: the interface with neuroethology. *TINS* 14:199–206
- Nelson DA, Marler P (1989) Categorical perception of a natural stimulus continuum: birdsong. *Science* 244:976–978
- Nottebohm F (1991) Reassessing the mechanisms and origins of vocal learning in birds. *TINS* 14:206–211
- Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, *Serinus canarius*. *J Comp Neurol* 165:457–486
- Price PH (1979) Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93:260–277
- Pytte CL, Suthers RA (2000) Sensitive period for sensorimotor integration during vocal motor learning. *J Neurobiol* 42:172–189
- Simpson HB, Vicario DS (1990) Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J Neurosci* 10:1541–1556
- Simpson HB, Vicario DS (1991) Early estrogen treatment of female zebra finches masculinizes the brain pathway for learned vocalizations. *J Neurobiol* 22:777–793
- Simpson HB, Vicario DS (1996) Male zebra finches can learn male-typical vocalizations from hormone-treated female tutors. *Animal Behav* 52:1119–1127
- Solis MM, Doupe AJ (2000) Compromised neural selectivity for song in birds with impaired sensorimotor learning. *Neuron* 25:109–121
- Studdert-Kennedy M, Liberman AM, Harris KS, Cooper FS (1970) Theoretical notes. Motor theory of speech perception: a reply to Lane's critical review. *Psychol Rev* 77:234–249
- Thorpe WH (1958) The learning of song patterns by birds, with special reference to the song of the Chaffinch, *Fringilla coelebs*. *Ibis* 100:535–570
- Vicario DS, Yohay KH (1993) Song-selective auditory input to a forebrain vocal control nucleus in the zebra finch. *J Neurobiol* 24:488–505
- Vicario DS, Naqvi NH, Raksin JN (2001a) Sex differences in discrimination of vocal communication signals in a songbird. *Anim Behav* 61:805–817
- Vicario DS, Naqvi NH, Raksin JN (2001b) Behavioral discrimination of sexually dimorphic calls by male zebra finches requires an intact vocal motor pathway. *J Neurobiol* 47:109–120
- Williams H, Nottebohm F (1985) Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science* 229:279–82
- Zann R (1984) Structural variation in the zebra finch distance call. *Z Tierpsychol* 66:328–345
- Zann R (1985) Ontogeny of the zebra finch distance call. I. Effects of cross-fostering to Bengalese finches. *Z Tierpsychol* 68:1–23
- Zann R (1990) Song and call learning in wild zebra finches in south-east Australia. *Anim Behav* 40:811–828
- Zann R (1996) *The zebra finch*. Oxford University Press, Oxford