

A learning program that ensures prompt and versatile vocal imitation

Wan-chun Liu* and Fernando Nottebohm*

Laboratory of Animal Behavior, The Rockefeller University, 1230 York Avenue, New York, NY 10021

Contributed by Fernando Nottebohm, October 23, 2007 (sent for review August 24, 2007)

Here we show how a migratory songbird, the chipping sparrow (*Spizella passerina*), achieves prompt and precise vocal imitation. Juvenile chipping sparrow males develop five to seven potential precursor songs; the normal development of these songs requires intact hearing but not imitation from external models. The potential precursor songs conform with general species-typical song parameters but differ from the song of wild, adult territorial males. As chipping sparrow males return from migration to start their first breeding season, they settle close to an older adult. The young male then stops producing all but one of its precursor songs, retaining the one that most resembles that of its neighbor. This single song then becomes more variable and, in a matter of days, is altered to closely match the neighbor's song. This elegant solution ensures species specificity and promptness of imitation.

auditory feedback | chipping sparrows | precursor song | sensitive period | vocal learning

It is thought that birds that imitate their song first memorize an external model and then use auditory feedback to gradually modify their vocal output until it matches the memorized model (1). This developmental process can take weeks or months in juveniles because the point of departure can bear little semblance to the final imitation (2–5). In many seasonal, migratory songbirds, however, vocal learning often has to be mastered in a shorter period. For example, the song of a yearling migratory male often closely matches that of an adult close neighbor first encountered when the yearling migrant arrives at the spring breeding grounds (6–10). How such fast and precise song matching is achieved remains controversial (11–13). One way to do it would be for a juvenile to imitate a diversity of external models during its hatching year and then during the following spring retain only the imitation that best matches the song of a neighbor with whom it interacts (8, 11, 14, 15). However, songs acquired during the hatching year might not come close to the song of breeding-season neighbors. Here we describe a learning program that can achieve a fast yet precise imitation of the song of a neighbor first encountered during a young male's first breeding season. This program develops a set of “potential precursor songs” that can be altered later in life to precisely match an external model.

We encountered this program while studying the song ontogeny of a seasonal, migratory songbird, the chipping sparrow (*Spizella passerina*). Adult male sparrows sing only one song type during the breeding season; each song type consists of the repetition of a single syllable (10) (Fig. 1). In nature, there are ≈ 25 –30 different syllable types [supporting information (SI) Fig. 7] distributed across the geographic range of the species (10). The very simple but diverse song of the chipping sparrow provides a great model to examine the detailed process of song ontogeny. A previous field study suggested that the simple song is most likely acquired by precise imitation of the song of an immediate adult neighbor at the breeding grounds during the juvenile's hatching summer or the following spring (10).

Results and Discussion

Song ontogeny in juvenile chipping sparrows starts at ≈ 1 month of age with soft and variable vocalizations referred to as “subsongs” (16–18) (Fig. 2). Early in the following spring (7–9 months old),

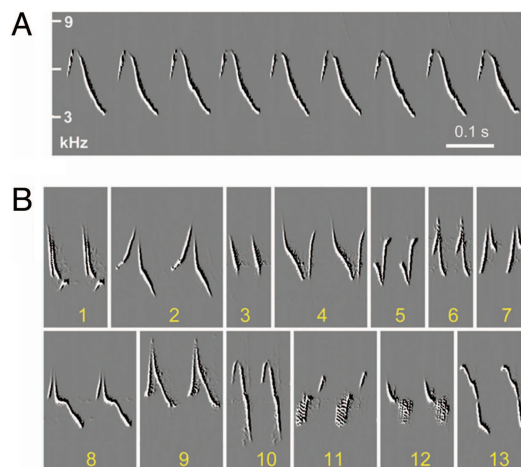


Fig. 1. The song of adult male chipping sparrows. (A) Each song consists of a syllable repeated many times, with variable song duration used under different social circumstances (11, 21). (B) Song variation among individual birds. Shown are song fragments from 13 free-ranging, territorial males, each identified by a different number, recorded in our study area. Each of these birds built its song from a different syllable type.

subsong gradually transforms into “plastic song,” with the appearance of recognizable song syllables and syntax. During this stage juveniles develop a repertoire of several juvenile “precursor” song types (Fig. 2). These multiple song types are delivered serially, each repeated several times before switching to a different one. We call them precursor songs because, as we shall see, in the normal course of events one of them will be modified to crystallized or match tutor song.

To determine whether these multiple precursor song types were imitated from adult tutors, juveniles ($n = 6$ males) were exposed to a single adult tutor for 3 weeks during their first summer only (1.5–2 months of age). During the following spring no tutor was provided, and each juvenile produced five to six recognizable song types (Fig. 3). Four of these birds imitated the tutor song in one of their multiple song types (SI Fig. 8A); the rest of song types did not match the song of the tutors or any wild-type adult song (Fig. 4). During early spring, at the plastic song stage (early March), the daily amount of time each juvenile spent singing increased significantly (Fig. 5A). Initially, the various song types were each given equal singing time, but in a few weeks one of these song types was produced much more frequently than the rest of song types (Fig.

Author contributions: W.-c.L. designed research; W.-c.L. performed research; W.-c.L. contributed new reagents/analytic tools; W.-c.L. analyzed data; and W.-c.L. and F.N. wrote the paper.

The authors declare no conflict of interest.

*To whom correspondence may be addressed. E-mail: liuw@mail.rockefeller.edu or nottebo@mail.rockefeller.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0710067104/DC1.

© 2007 by The National Academy of Sciences of the USA

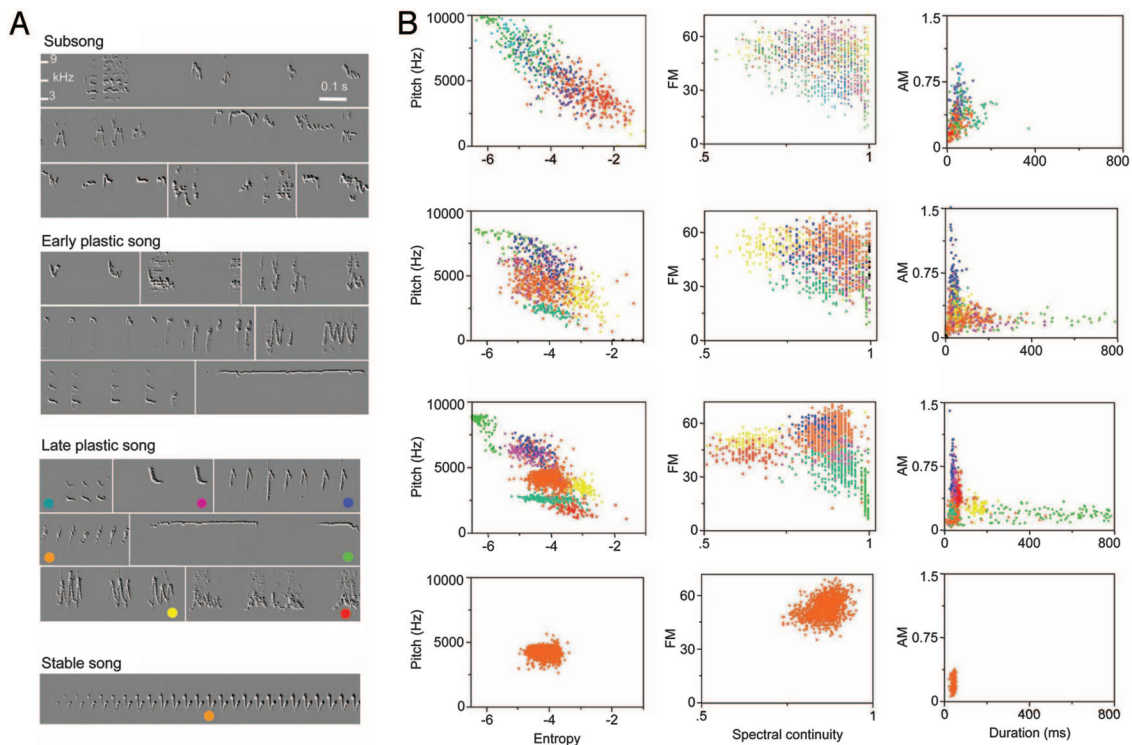


Fig. 2. An example of song ontogeny in chipping sparrows. (A) Recordings from a same individual show subsong (63 days old), early plastic song at 265 days old (as recognizable song and syntax have just started to emerge), late plastic song at 288 days old (when precursor song types are well developed), and stable adult song (325 days old). Each color dot in the sonogram identifies one song type in the cluster analysis. (B) In the syllable clusters from cluster analysis, six song features are measured (see *Materials and Methods* for details). During the early and late plastic song stage, each song type was denoted by a cluster of color dots, but the color dots shown in subsong are arbitrarily assigned. The sounds were obtained from a single day (i.e., 0500–2000 hours) of continuous song recording.

5B). This song type then crystallized as breeding approached (late April to early May), and the rest of song types disappeared (SI Fig. 8A). The crystallized song was not necessarily the song copied from the tutor.

To examine whether the multiple precursor songs of juveniles could be produced without tutors, juveniles ($n = 3$ males) were visually and acoustically isolated from other birds. Each isolate produced five to six different song types in the spring (Fig. 3); these untutored multiple precursor songs were no different from uncopied songs produced by birds in the previous, “live-tutored” group [multivariate ANOVA (MANOVA), Wilks’ $\lambda = 0.917$, $\chi^2 = 185.54$, $P = 0.35$]. The isolated birds also settled on a single adult song type (Fig. 3 and SI Fig. 8B). This crystallized song type was a modification of one of the songs in the juvenile multisong repertoire, and it did not match any of the wild-type adult songs recorded from free-ranging individuals in reproductive condition (Mann–Whitney U test, $P < 0.0001$).

We further tested whether auditory feedback is required to produce the multiple precursor songs by deafening juveniles at 18–28 days of age ($n = 4$ males). Deaf birds produced only one to three different song types during the plastic song stage. The juvenile multisong repertoire of deaf birds was significantly different from that of hearing birds (MANOVA, Wilks’ $\lambda = 0.65$, $\chi^2 = 873.18$, $P = 0.009$) (Fig. 3). Eventually, as in hearing birds, deaf birds culled all juvenile song types except one, which they modified and retained as adult song (SI Fig. 8C).

To test whether these precursor song repertoires occurred also in nature, we trapped juvenile sparrows ($n = 2$) shortly before migration (late September; 2–4 months of age) and housed each one in a sound-proof chamber without hearing or interacting with other birds. Each of them produced five to six precursor song types during the following spring (Fig. 3 and SI Fig. 8D). These precursor songs were very similar to those of the hand-reared, hearing birds

and did not closely resemble adult wild-type song. Multisong repertoires occur also in wild adult males, although we have never heard them in breeding, territorial individuals. We established this in two ways. (i) In nature, territorial adults produce only a single song type during the breeding season. However, when they (five adults, >2 years old) were captured and treated with testosterone, so they would sing a lot and could be used as live tutors, and housed singly, each produced multiple precursor song types (Fig. 3 and SI Fig. 8D), and only one of these songs was the one used during the breeding season. (ii) Another group of four adults was caught at the end of the breeding season and housed indoors. Early in the following spring they produced a multisong repertoire that, in addition to their adult song, included four to six other precursor songs. In both cases, the precursor song types produced by the adults were much like those of juveniles.

The multiple precursor song repertoire developed in juvenile sparrows is different from the wild-type adult songs used by individuals in breeding condition based on several lines of evidence: (i) the juvenile songs of multisong repertoires, except the one copied from the tutor, did not closely match any of the wild-type adult songs (Fig. 4). When we compared the song similarity between the songs from the multisong repertoires of juveniles (from tutored, isolated, and deaf groups) and each of 30 wild-type adult songs (SI Fig. 7), the highest similarity scores ranged from 0% to 36% similarity. When the same comparison was done for songs copied from a tutor, the similarity scores ranged from 81% to 93% ($n = 7$ tutored juveniles) (Fig. 4). (ii) Although the acoustic space of adult wild-type songs fell within the acoustic space defined by juvenile multisong repertoires (Figs. 2 and 3), the space of the former was significantly smaller than that of the latter (MANOVA, Wilks’ $\lambda = 0.6$, $\chi^2 = 1,367.72$, $P < 0.0001$, one-tailed) (3). When songs drawn from juvenile or adult song samples were played to free-living territorial adults ($n = 18$ males), these adults responded

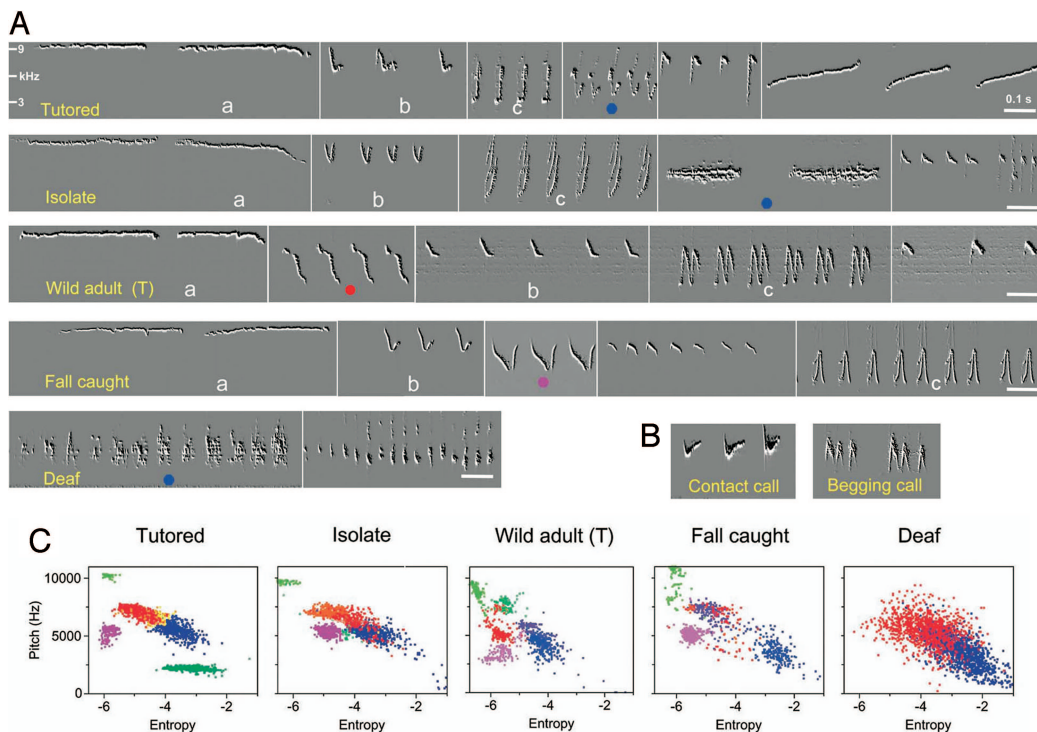


Fig. 3. Examples of multisong repertoires from one bird in each of five different experimental groups recorded during the late plastic song stage. (A) Birds in all hearing groups produced five to six different song types during the plastic song stage, but deaf birds produced only one to three song types. Some song types were very reminiscent of sounds produced earlier in development, such as contact calls (syllable b) and food begging calls (syllable c), as if these earlier sounds were used, with modification, to build some of the songs in the multisong repertoire. Song type a occurred only in hearing birds. The song type later modified to adult stable song is identified by a color dot. (B) Examples of the contact calls and food begging calls. (C) Clusters of song types from cluster analysis; only two features (mean pitch and entropy) are shown here. The sounds were obtained from a single day (i.e., 0500–2000 hours) of continuous song recording during late plastic song or, in the case of the wild adult, 3 days after onset of testosterone treatment. Each color dot represents one syllable. Each song type cluster present in each male's song repertoire occupied its own and different acoustic space. Notice that the deaf bird did not have the high-pitched whistles present in the four other groups.

more aggressively (or differently) toward the speaker broadcasting wild-type adult song than toward the speaker broadcasting songs from juvenile multisong repertoires. This effect persisted even when we took a single syllable from a song in the multisong repertoire of a juvenile and iterated it as in normal adult song so that intrasong syllable variability was similar in the playbacks that used juvenile or adult material (MANOVA, $F = 24.5$, $P = 0.002$) (see SI Table 1). Overall, these results suggest the adult and juvenile songs represent different acoustic signals.

The acoustic “space” defined by multisong repertoires was very similar in socially tutored ($n = 6$ birds), socially isolated ($n = 3$),

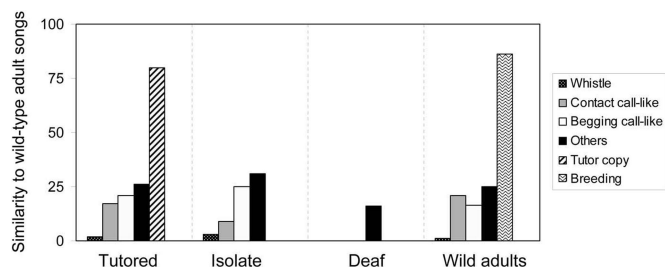


Fig. 4. Juveniles' precursor songs do not match wild-type adult songs. Each of the 30 wild-type adult songs was compared with each of the multisong repertoires produced by four experimental groups: tutored ($n = 6$ birds), isolated ($n = 3$), deaf juveniles ($n = 4$), and wild-caught adult males ($n = 5$) treated with testosterone. The precursor song repertoires were visually categorized as “whistle” song, contact-call-like song, begging-call-like song, song copied from an adult tutor in tutored birds, and the song type used during the breeding season by wild-caught adults (“Breeding”).

wild-caught juveniles ($n = 2$) and wild-caught adults ($n = 5$) (MANOVA with four groups; $\chi^2 = 132.6$, $P = 0.583$) (Fig. 4) but differed significantly in the early-deafened birds (see above). Visual inspection of the multisong repertoire revealed that birds in the first four groups shared several song syllable types, including a high-pitched whistle (Fig. 3 and SI Fig. 8, song type a), a syllable reminiscent of “contact calls” (Fig. 3, type b), and one similar to “begging calls” (Fig. 3, type c).

Why do juvenile chipping sparrows develop a diverse song repertoire that is not acquired by imitation, that is sung profusely early in the spring, and that is not used during the breeding season? We speculated that one or more of these juvenile songs could be modified to match adult song types in the spring. To test this idea, seven juvenile males were not exposed to an adult tutor until the following spring, when each juvenile was tutored for 5 days. Soon after tutor exposure, one or two of the multiple song types in three of the tutored birds became unstable and then was modified to match that of the tutor; the other syllables remained more stable but over a period of 1–2 weeks stopped being produced. Within 5 days after first exposure to the tutor, a recognizable approximation to the tutor's song was already in place, with a stronger, stable match by day 10 (Fig. 6). Cluster analysis, discriminant function analysis, and song similarity tests (Fig. 6 and SI Fig. 9) revealed that the songs (i.e., syllables) used for modification in the three birds that imitated the external model were the ones that, before exposure to that tutor, had the closest acoustic distance to the tutor song syllable. Of the four remaining birds, one copied the song of another juvenile in an adjacent cage, which retained one of its juvenile songs, and the other two also retained one of their juvenile songs without matching their tutor (SI Fig. 9). The four birds that copied an external model

song to reduce the variability in syllable delivery found in juvenile song types. Altogether, 12 playbacks of precursor song types were prepared. The playback experiments were conducted at the Rockefeller University Field Research Center and at the nearby Institute for Ecosystem Studies. Each subject ($n = 18$ territorial males with 27 trials) heard a playback of precursor song or wild-type adult song in the morning (0900–1100 hours). We placed the playback speaker (SME AFS-A70), connected to a Sony MZNH1 hi-minidisc recorder, near the center of each focal male's territory. The volume of playbacks was set to match the volume of diurnal adult wild-type, territorial song (≈ 75 –80 dB). During each playback trial (3 min of song playback and 7 min of silence), only one (adult or precursor) song type was played to the subject. Behavioral data included closest distance to speaker, time latency to produce first vocalization, number of vocalizations, and time period spent within 10 m of speaker. We used principal-components analysis to analyze these variable with SPSS 14.0 (SPSS).

Sound Analysis and Statistical Analysis. Song imitation was defined as instances in which the song of a juvenile reached a similarity index of 72% or better when compared with the tutor song. This measure of similarity uses Euclidean distances across six sound features: pitch, FM, AM, Wiener entropy, duration, and spectral continuity (for more details, see ref. 21). Measures of similarity index were obtained by taking each of 300 syllables of a same kind from the pupil's adult song and comparing them with 300 syllables of a kind from the tutor. For these comparisons we used Sound Analysis Pro (SAP) (21), with the minimum silent gap separating syllables adjusted to 5 ms. The justification of 72% as the criterion for imitation was based on the fact that, when two blind judges looking at sonograms of pupil and tutor songs (not identified) were asked to identify which pupils copied which tutors, these positive identifications carried similarity scores of 72% or higher, whereas comparisons with songs that in their judgment were not copies yielded scores of 32% or less.

Cluster Analysis. We used cluster analysis to recognize song clusters. Recorded songs were first segmented as syllables by using SAP. We used the statistical software from SAS (JMP 5.1; SAS Institute) and SPSS to classify syllables for cluster analysis. We used hierarchical clustering analysis with equal weighting on each

variable and with a dendrogram output. Based on the clustering dendrogram, we then selected the squared Euclidean distance to determine the distance between clusters; between groups of syllables (22), the further the Euclidean distance between two syllables, the more dissimilar the two syllables. We combined two independent ways to determine the number of clusters: (i) the cluster of syllables can be selected based on the sudden jump of Euclidean distance between groups; this way we can determine the approximate number of clusters (SPSS user manual). (ii) Two blind judges visually inspected the sonograms to classify the number of clusters. In some cases where these two methods did not agree, we reexamined each of the six song features of those ambiguous clusters and then used discriminant function analysis (see below) to determine whether these acoustic features were significantly different to be classified as two clusters.

Discriminant Function Analysis and MANOVA. After different song type clusters were identified by cluster analysis, we used discriminant function analysis (SPSS 14.0 and JMP5.1) to compare the potential precursor songs and wild-type adult songs, with a measure of acoustic separation between the two. We also used this analysis to separate populations of song types and to decide which of the potential precursor syllables was acoustically closer to the tutor song. For discriminant function analysis, we used the same six acoustic features as independent variables that we had used for cluster analysis. MANOVA was used to test the significance of the discriminant functions obtained. In the statistical output, Wilks' λ was used to determine whether the discriminant model as a whole was significant, or what song features were significant when discriminating between two or more groups, information that we could then use to compare with the similarity scores obtained by SAP.

ACKNOWLEDGMENTS. We thank Tim Gardner for providing his sound recording software and for his comments on the manuscript. Peter Marler, Mark Konishi, Mike Beecher, Heather Williams, and Michale Fee read our manuscript and made helpful suggestions. We thank Debbie Gardner and Dania Vallinova for helping hand-rear juvenile birds, and we thank Daun Jackson, Sharon Sepe, and Helen Ecklund for birdkeeping. The Institute of Ecosystem Studies provided the land for some of the field work. This work was supported by National Institutes of Mental Health Grant MH 18343 (to F.N.) and a Li Memorial Fellowship (to W.-c.L.).

- Konishi M (1965) *Zeit Tierpsychol* 22:770–783.
- Immelmann K (1969) in *Bird Vocalization*, ed Hinde RA (Cambridge Univ Press, London), pp 61–74.
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) *Science* 291:2564–2569.
- Marler P (1987) in *Imprinting and Cortical Plasticity*, eds Rauschecker JP, Marler P (Wiley, New York), pp 99–135.
- Baptista LF, Petrinovich L (1984) *Anim Behav* 32:172–181.
- Kroodsma DE (1974) *Zeit Tierpsychol* 22:770–783.
- Payne RB (1981) *Anim Behav* 29:688–697.
- Nelson DA (1992) *Behav Ecol Sociobiol* 30:415–424.
- Bell DA, Trail PW, Baptista LF (1998) *Anim Behav* 55:939–956.
- Liu W-C, Kroodsma D (2006) *Condor* 108:509–517.
- Marler P (1997) *J Neurobiol* 33:501–516.
- Nelson DA (1997) in *Social Influences on Vocal Development*, eds Snowdon CT, Hausberger M (Cambridge Univ Press, Cambridge, UK), pp 7–22.
- Kroodsma DE (2004) in *Nature's Music*, eds Marler P, Slabbekoorn H (Elsevier Academic, San Diego), pp 108–131.
- Nelson DA, Marler P (1994) *Proc Natl Acad Sci USA* 91:10498–10501.
- Nordby JC, Campbell SE, Beecher MD (2007) *Anim Behav*, in press.
- Thorpe WH (1955) *Ibis* 97:247–251.
- Marler P, Peters S (1982) in *Acoustic Communication in Birds*, eds Kroodsma DE, Miller EH (Academic, New York), Vol 2, pp 135–147.
- Nottebohm F (1972) *J Exp Zool* 179:35–50.
- Marler P, Sherman V (1983) *J Neurosci* 3:517–531.
- Kroodsma DE (1990) *Anim Behav* 40:1138–1150.
- Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) *Anim Behav* 59:1167–1176.
- Sokal RR, Rohlf FJ (1995) *Biometry* (Freeman, New York), 3rd Ed.