Behavioral/Systems/Cognitive

Limits on Reacquisition of Song in Adult Zebra Finches Exposed to White Noise

Jason D. Zevin,[1] **Mark S. Seidenberg,**[2] **and Sarah W. Bottjer**[1]

[1]Program in Neuroscience, University of Southern California, Los Angeles, California 90031, and [2]Department of Psychology, University of Wisconsin–Madison, Madison, Wisconsin 53706

Zebra finches (*Taeniopygia guttata*) learn a specific song pattern during a sensitive period of development, after which song changes little or not at all. However, recent studies have demonstrated substantial behavioral plasticity in song behavior during adulthood under a range of conditions. The current experiment examined song behavior of adult zebra finches temporarily deprived of auditory feedback by chronic exposure to loud white noise (WN). Long-term exposure to continuous WN resulted in disruption of song similar to that observed after deafening. When auditory feedback was restored by discontinuing WN, birds were either tutored using tape-recorded playback or housed with adult conspecific tutors. No evidence of learning new tutor syllables was observed, and recovery of pre-WN song patterns was very limited after restoration of hearing. However, many birds did reacquire some aspects of their pretreatment song, suggesting an adult form of learning that may retain some of the initial aspects of sensorimotor acquisition of song in which vocalizations are shaped to match a stored template representation. The failure to learn novel song elements and the modest degree of recovery observed overall suggest a limit on plasticity in adult birds that have acquired species-typical song patterns and may reflect an important species difference between zebra finches and Bengalese finches.

Key words: sensitive period; behavioral plasticity; songbirds; vocalization; auditory learning; song learning

Introduction

Songbird species vary in the degree of plasticity exhibited in adult song behavior. Canaries and other so-called "open-ended learners" continue to modify their songs throughout adulthood, whereas "closed-ended learners" such as zebra finches and song sparrows typically learn a single song as juveniles that they sing in a stereotyped manner throughout their adult lives (Nottebohm and Nottebohm, 1978; Böhner, 1990). However, a number of recent studies have demonstrated a surprising degree of plasticity in adult closed-ended learners. Changes in adult zebra finch song as a result of deafening (Nordeen and Nordeen, 1992), tracheosyringeal nerve sectioning (Williams and Mehta, 1997), and delayed auditory feedback (Leonardo and Konishi, 1999) reveal considerable plasticity in an ordinarily stereotyped system. Furthermore, lesions of the basal ganglia circuit that supports song acquisition in juveniles can prevent the effects of these manipulations (Williams and Mehta, 1997; Brainard and Doupe, 2000), suggesting that these manifestations of adult plasticity may share a common mechanism with initial song acquisition.

Interestingly, there is also some variability in the timing or degree of plasticity in adult song among species of closed-ended

learners. Nordeen and Nordeen (1992) found that deafening resulted in the gradual disruption of adult song in zebra finches over a period of 4–6 weeks, whereas Okanoya and Yamaguchi (1997) and Woolley and Rubel (1997) found deafening-induced changes in the songs of Bengalese finches after only 1 week. In later work, Woolley and Rubel (2002) deafened Bengalese finches using ototoxic cochlear lesions. Because hair cells regenerate in birds, Woolley and Rubel (2002) were also able to observe changes in song as a result of the gradual restoration of hearing. They found that some birds learned song material from cagemates with which they were housed during recovery. This is an important finding because it suggests that the capacity for neural plasticity present in adult Bengalese finches is sufficient not only to allow changes in previously learned song as a result of distorted or removed feedback, but also to support the acquisition of novel song material.

The range of facts about adult plasticity is consistent with an emerging view that song stereotypy is the result of an active process in which auditory feedback of song output is compared with a stored representation of song (or "template"). Behavioral stereotypy emerges in juvenile birds once this comparison indicates congruence between vocal output and the stored song template (and is dependent on achieving this congruence, as opposed to being driven by a strictly timed maturational mechanism). Maintenance of vocal stereotypy in adult birds is also the result of this active comparison process, and the decline in plasticity exhibited over the lifespan is the result of the "engrainment" of the corresponding representation (Doupe and Kuhl, 1999; Lombardino and Nottebohm, 2000; Troyer and Bottjer, 2002). When auditory feedback is disrupted, this process introduces changes

into the stable comparison that has been achieved, and the net result is a decline in spectral and temporal stereotypy of song behavior.

The current experiment examines song change in adult zebra finches temporarily deprived of auditory feedback via chronic exposure to high-amplitude white noise (WN). The results provide evidence of only limited song reacquisition after WN was discontinued. The majority of syllables that became disrupted because of WN exposure did not recover after restoration of hearing, and lost syllables were never recovered. In addition, adult zebra finches did not learn new syllables from tutors, in contrast to findings for adult Bengalese finches (Woolley and Rubel, 2002), suggesting species-specific differences in adult plasticity.

Materials and Methods

Subjects

Male zebra finches used in this experiment were bred and raised in group aviaries, each of which contained five adult breeding pairs, on a 14/10 light/dark schedule at 25°C. At 55 d post-hatch (juvenile stage), subjects were moved to small flight cages in which they were housed in groups of three. This allowed ample exposure to their father's song to ensure normal song development (Böhner, 1990). At 120 d post-hatch (adulthood), subjects were housed individually in these cages, which were placed within sound-attenuated boxes, and WN exposure began.

WN exposure

Twelve sound-attenuating boxes (23 × 20 × 26 inches) were built in-house and consisted of double plywood walls separated by acoustic foam. Two WN generators (Quan-Tech 420 and General Radio Company 1382; both from Tucker Davis Technologies, Garland, TX) were used to present WN, driven by two six-channel amplifiers (NAD 916). Each sound-attenuating box was equipped with a four-inch, 200 W two-way speaker (Houston Acoustic 4202), mounted on the top side of one edge. To test whether the noise exposure was consistent at all audible frequencies (250 Hz to 20 kHz), WN was recorded using a reel-to-reel tape recorder (TEAC X-300) and analyzed using a digital sound spectrograph (Kay Elemetrics DSP 5500). The WN was found to be flat within 5 dB at all frequencies, measured at 100 dB. Because of near-field acoustics, the precise amplitude of the WN was not constant at all points in the sound-attenuating boxes. Levels were thus set so that the minimum amplitude was 100 dB as measured at five different points in each box with a hand-held sound level meter (Realistic 33-2050). In practice, this meant that the amplitude 2–3 cm from the speaker ranged from 116 to 120 dB. The maximum volume at the level of the perch was ~103 dB. This level of WN does not result in damage to hair cells in zebra finches (Ryals et al., 1999; Sarah Woolley, personal communication).

Four groups of birds were exposed to different schedules of WN. Groups 1 and 2 were both exposed to a mixture of "interrupted" and "continuous" WN schedules. Interrupted refers to a schedule on which song was recorded once weekly for 1 hr or until 30 song samples were recorded. During recording, WN was turned off, allowing the subject to hear its own vocalizations (notwithstanding any short-term effect of WN on hearing; see below). These weekly, hour-long recording sessions were the only time during which noise was not present on the interrupted schedule. For subjects on the continuous schedule, WN was simply never turned off, and vocal behavior of the birds was not recorded. Our initial motivation for including the interrupted condition was to allow us to observe changes in song over the course of exposure to WN. However, we observed no substantial change in song over extended periods during which subjects were exposed to interrupted noise, and these birds were therefore switched to a continuous schedule because the hourly interruption of noise was apparently preventing changes in song. Whether such brief exposure to song can actually serve to counteract the effects of WN exposure observed here remains an open question for future research. Groups 3 and 4 were run later and were maintained on continuous noise from the outset. In cases in which siblings from the same clutch were run, they were evenly distributed across treatment groups.

The conditions were: birds in group 1 were maintained on an interrupted schedule for 2 months, after which they were switched to a con-

tinuous schedule for 4 months (n = 6). Group 2 birds were maintained on an interrupted schedule for 6 months, after which they were switched to a continuous schedule for 4 months (n = 5). Birds in group 3 were maintained on a continuous schedule for 5 months (n = 6). Group 4 birds were maintained on a continuous schedule for 7 months (n = 5). Finally, a control group of five birds was recorded on two different occasions, 3 months apart without any treatment to establish a baseline for measures of temporal aspects of song (see below).

Song recording

Song was recorded using automated recording software (Avisoft SASLab). Microphones (Shure BG 1.1) were positioned inside the sound-attenuated boxes and each channel independently preamplified using a mixing board (Mackie 1604-VLZ). Song was digitized using an eight-channel soundcard (SEK'd ARC 88), allowing simultaneous monitoring of eight birds. Birds were recorded on two separate occasions (1 week apart) before the onset of WN exposure. After WN onset, weekly recording sessions (interrupted conditions only) lasted 1 hr, or until 30 samples of song were recorded. On some occasions, up to 5 hr would elapse without 30 samples being recorded for a given subject. Rather than allow a lengthy interruption of WN exposure, recording was discontinued after 5 hr and WN was restored. After WN was discontinued, recordings were made daily for 2 weeks, then weekly for at least 100 d. The time points discussed in Results include the pre-WN recordings (PRE), the first day out of WN (D1), and after 100 d out of WN (D100).

Tutoring

Potential tutors were recorded, and spectrograms made of their songs. These were compared with the subjects' songs, and pairings were selected to maximize dissimilarity between the tutors' songs and the subjects' PRE songs. In this way, any observed increases in similarity to a tutor song could be distinguished from recovery of PRE song more clearly. Birds in group 1 were tape-tutored using computer playback for 2 hr per day. Sequencing software (Cakewalk Pro Audio 8) was used to present two song samples every 10 min during each hour-long tutoring session, one in the morning and one in the afternoon. Birds in groups 2, 3, and 4 were presented with live tutors immediately after their initial recording on D1 after removal from continuous WN. Tutors were housed continuously with subjects, except during recording sessions.

Evoked potential recordings

To determine whether long-term exposure to WN had lasting effects on hearing, we recorded evoked potentials from auditory brainstem in 15 birds housed in continuous WN under the same conditions as experimental birds for 1 month. Recordings were made at 1 hr, 3 d, and 1 week after removal from WN (five birds in each condition) using procedures modeled after those of Woolley et al. (2001).

Auditory thresholds were determined by evoked potential recordings from auditory brainstem during the presentation of 10 msec pure tones (rise/fall time of 1 msec) at frequencies of 0.25, 0.5, 1.0, 1.5, 2.0, 3.0, 3.5, 4.0, 5.0, and 6.0 kHz, presented at a rate of 10 per second. Stimuli were generated using custom software (Brandon Warren, University of Washington, Seattle, WA) and delivered with a free field speaker (RCA PRO-X55AV), placed at a 90° angle to the head at a distance of ~10 cm. Evoked responses were digitized using the Line In input to the soundcard and recorded by the same software used for stimulus generation.

Subjects were anesthetized with urethane (0.002 mg/kg), and their heads were stabilized using a specially designed holder. Differential recordings were made using pin electrodes (Grass Instruments, Quincy, MA). Two electrodes were implanted transcranially into the brain: a recording electrode was placed into the cerebellum just above the auditory brainstem nuclei, and a differential electrode was placed in the contralateral anterior cortex. A third (ground) electrode was inserted into the leg muscle. Responses were preamplified using a 10× stage pre-amp (A-M Systems, Carlsborg, WA), filtered (0.3–10 kHz bandpass), and amplified at 1000× (microelectrode amplifier model 1800; A-M Systems). Stimulus presentations for each frequency were begun at 120 dB and decreased in intensity in 10 dB steps in the range of amplitude that was well above threshold, and in 5 dB steps near threshold. Responses were averaged over 200 traces above threshold and 500 or 1000 traces

Table 1. Analyses of temporal structure of song for all subjects on D1 (immediately after removal from WN), relative to pre-WN (PRE)

	Linearity			Consistency			Syllable probability	Motif probability	Syllables changed	Syllables lost	Proportion lost/changed
	PRE	D1	Diff	PRE	D1	Diff					
Group 1											
Y685(S)	1.00	0.67	0.33	0.93	0.88	0.05	0.40	0.00	3	1	1.00
Y692(S)	1.00	1.00	0.00	0.95	0.94	0.01	0.60	0.00	2	2	0.80
Y693(S)	1.00	0.80	0.20	0.99	0.92	0.07	0.61	0.00	5	1	1.00
Y696(S)	1.00	0.75	0.25	0.91	0.97	−0.06	0.16	1.00	1	5	1.00
Y663	1.00	1.00	0.00	0.98	0.99	−0.01	0.21	0.00	0	0	0.00
Y695	1.00	1.00	0.00	0.97	0.96	0.01	0.05	1.00	3	0	0.38
Mean	1.00	0.87	0.13	0.96	0.94	0.04	0.34*	0.33*	2.00*	1.83*	0.68*
Group 2											
Y686(S)	1.00	1.00	0.00	0.91	0.99	−0.08	0.24	0.50	1	1	0.50
Y665	1.00	1.00	0.00	0.94	1.00	−0.06	0.07	1.00	1	0	0.20
Y666	1.00	0.83	0.17	0.98	0.97	0.01	0.23	0.00	0	1	0.17
Y677	1.00	0.78	0.22	0.99	0.91	0.08	0.50	0.00	1	0	0.25
Y682	1.00	1.00	0.00	0.95	0.95	0.00	0.12	0.50	5	0	1.00
Mean	1.00	0.92	0.08	0.96	0.96	0.05	0.22*	0.40*	1.60*	0.40	0.42*
Group 3											
W696	1.00	1.00	0.00	0.99	1.00	−0.01	0.19	0.00	1	0	0.20
Y701	0.80	0.57	0.23	0.90	0.87	0.03	0.05	0.33	0	0	0.00
Y746	1.00	0.71	0.29	0.98	0.89	0.09	0.17	0.33	0	0	0.00
Y747	1.00	1.00	0.00	0.97	0.99	−0.02	0.04	1.00	0	0	0.00
Y748	0.67	1.00	−0.33	0.84	0.93	−0.09	0.16	0.20	0	0	0.00
Y750	1.00	1.00	0.00	0.96	0.94	0.02	0.21	1.00	0	0	0.00
Mean	0.91	0.88	0.14*	0.94	0.94	0.04*	0.14*	0.48	0.17	0.00	0.03
Group 4											
pu746(S)	1.00	0.75	0.25	0.86	0.78	0.08	0.20	0.67	3	0	1.00
pu750(S)	0.86	1.00	−0.14	0.90	0.93	−0.03	0.37	1.00	3	3	0.86
pu762(S)	1.00	**	**	0.97	**	**	1.00	0.00	0	5	1.00
pu781(S)	1.00	1.00	0.00	0.95	0.95	0.00	0.07	1.00	5	0	1.00
Pu776	1.00	1.00	0.00	1.00	0.98	0.02	0.02	1.00	3	0	0.60
Mean	0.97	0.94	0.10	0.94	0.91	0.03*	0.33*	0.73	2.80*	1.60	0.89*
Control											
Pu767	1.00	1.00	0.00	0.98	0.98	0.00	0.01	1.00	0	0	0.00
R812	1.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0	0	0.00
pu751	1.00	1.00	0.00	0.97	0.95	0.02	0.04	0.50	0	0	0.00
Lg788	1.00	1.00	0.00	1.00	0.99	0.01	0.05	1.00	0	0	0.00
Lg728	1.00	1.00	0.00	0.99	0.96	0.03	0.05	0.50	0	0	0.00
Mean	1.00	1.00	0.00	0.99	0.98	0.01	0.03	0.80	0.00	0.00	0.00

D1, 1 d after WN was discontinued; Diff, difference between PRE and D1 song; syllable probability, the change in syllable probability between PRE and D1; motif probability, the proportion of motifs at D1 conserved from PRE; proportion lost/changed, the sum of deleted and changed syllables divided by the number of syllables in pre-WN song; (S), severely disrupted bird. See text for additional details.

*Significantly different from controls, $p < 0.05$. **No scores were generated for pu762 because song consisted of a single syllable.

near threshold. Threshold was defined operationally as the intensity at which the mean evoked response was twice the amplitude of the baseline electrical signal.

Subjective analyses of song
Subjective analyses of song were conducted as one means of assessing the effects of WN on song and of changes after removal of WN between D1 and D100. Two independent observers examined complete motifs of song and scored syllables as lost or changed. D1 songs were examined relative to PRE songs, whereas D100 songs were examined relative to D1 and PRE songs. Lost syllables simply did not appear in the later song sample, whereas changed syllables were typically in the same sequence in which they had been present and could be visually identified as exemplars of the same syllable despite gross changes in spectral properties. Changed syllables involved changes in fundamental frequency, harmonic structure, loss of specific subsyllabic structures, truncation (loss of notes within syllables), increased noisiness, or a combination of these. We used a conservative criterion for scoring syllables as changed such that only syllables that were quite visibly different between two time points were scored as changed; we did not attempt to assess subtle changes in syllable morphology. In addition, syllables were scored as changed or lost only if judged as such by both observers.

For the subset of syllables scored as changed at D100, two independent observers examined isolated exemplars of each and judged whether the D100 iteration was more similar to the D1 or the PRE version. This served as a preliminary test of whether syllables initially scored as obviously

changed at D100 increased in similarity to their PRE iterations, providing possible evidence of song reacquisition. A second test was conducted in which the identity of the D1 and D100 syllables was unknown to the two observers. In the second analysis, all syllables present at all three time points were examined. Syllable renditions from D1 and D100 were compared with renditions from PRE, and each was scored for similarity on a scale of 1 (similar) to 3 (dissimilar). Syllables scored as 1 were judged as being no different from their PRE iterations, those scored as 2 were judged as being different from their PRE iterations but still fairly similar, and those scored as 3 were judged as being clearly dissimilar from their PRE iterations. Observers were aware neither of the syllable identity nor of the time point at which later recordings were made (PRE samples were labeled to enable comparisons). Scores for D1 and D100 were then compared to establish at which time point song was more similar to PRE, and only those instances in which both observers agreed were counted. This procedure allowed a direct comparison of the blind results with the initial analysis.

In all cases, sound analysis measures of syllable accuracy (described in more detail below) were generated to provide objective scores to supplement the subjective measures described above. Accuracy scores were used to verify the subjective classification of syllables as changed (Tables 1 and 4) and to verify subjective judgments of increases in similarity to PRE.

Analyses of temporal structure of song
Zebra finch song consists of a stereotyped sequence of three to eight discrete vocalizations called "syllables." Syllables range in duration from

~50 to 250 msec and are separated by 10–50 msec of silence. A sequence of syllables is called a "motif," and multiple motifs are often sung in "bouts." Zebra finches generally sing one canonical motif with little variation throughout their adult lives (Zann, 1996; Brainard and Doupe, 2001). Many birds also sing a limited number of "motif variants" (Williams and Mehta, 1997) wherein some subset of syllables is sung optionally. For example, a series of call-like notes may be sung between motifs in a bout but not at the end of a bout. Thus, there is normally some variability among songs with respect to temporal structure, although all birds produce one dominant motif. We used four measures to assess temporal structure of song and its change as a result of WN exposure: linearity, consistency, syllable probability, and motif probability.

To determine whether changes in any of these measures were the result of normal variation over time (Brainard and Doupe, 2001), we also made two sets of recordings from control birds 3 months apart and computed these statistics for both sets of recordings. One-tailed Mann–Whitney summed ranks tests with α levels of 0.05 were applied to comparisons between experimental groups and this set of control animals. One-tailed tests were justified because in all cases, we were expecting a greater degree of change in experimental than control groups. Nonparametric tests were used because of the small sample sizes and the noninterval nature of the measures.

Linearity. Linearity scores measure the degree to which syllables are sung in a specific order. A perfectly linear song is one in which syllable A is always followed by syllable B, syllable B is always followed by syllable C, and so forth. Thus, a nonlinear song is one in which syllables occur in different sequences in different renditions. Zebra finch songs tend to be highly linear because it is relatively rare for syllables to be sung in a variable order or to be repeated within a motif. However, the existence of motif variants may affect linearity if they result in syllables occurring consecutively on some occasions, but not others. For example, if a bird usually sings ABCDE, but CD is sometimes left out to form ABE, this will decrease the linearity score.

We used the following measure of linearity:

$$Linearity = (number\ of\ different\ syllables - 1)/$$
$$number\ of\ syllable\text{-}to\text{-}syllable\ transitions.$$

This measure differs somewhat from the linearity measure developed by Scharff and Nottebohm (1991) in that it does not count transitions at the ends of songs (Foster and Bottjer, 2001; Iyengar and Bottjer, 2002). This means that a perfect linearity score can be achieved even if a bird's song is occasionally truncated at the beginning and end. For example, if a bird sings ABC#BC#AB# (with pound signs representing the ends of motifs), it will receive a linearity score of 1 because there are two syllable-to-syllable transitions (AB and BC) and three syllables (A, B, and C). The same set of renditions would receive a score of 0.75 using the Scharff and Nottebohm (1991) measure because this measure includes the transitions C# and B# in its calculation of linearity.

Consistency. Consistency provides a slightly different measure of song stereotypy. A song can be linear in the sense that syllables always occur in the same order, but inconsistent in the sense that not all syllables are sung in all renditions. For example, if a bird sometimes sings ABCDE but occasionally sings ABC, this will affect the consistency score but not the linearity score. Furthermore, because linearity is computed by considering the absolute number of transitions, regardless of their frequency of occurrence, it is highly sensitive to any deviations from perfect linearity. In contrast, consistency measures the proportion of syllable transitions that are accounted for by the most frequent, or "dominant," transition for a given syllable. This means that consistency will be less sensitive to infrequent deviations from a canonical pattern, while being more sensitive to deviations that do not affect linearity but are very frequent. In the example given above, for instance, if the bird sang the ABC variant only once, this would have only a moderate effect on consistency. In contrast, if the ABC variant were sung half the time and the ABCDE variant sung the other half, the consistency score would be quite low. Our measure of consistency was taken directly from Scharff and Nottebohm (1991):

$$Consistency = \Sigma(dominant\ transition/all\ transitions)/N.$$

In this measure, we count transitions (#) at both the beginnings and endings of motifs as syllables to capture variability at the beginnings and ends of songs. In the example given above, this measure gives a score of 0.83, reflecting the fact that the syllables #, B, and C participate in two different transitions; for example, for B, the following syllable could be C or #.

Syllable and motif probability scores. Williams and Mehta (1997) observed changes in the probability with which specific syllables and, as a result, specific song motif variants were sung after tracheosyringeal nerve sections (cf. Foster and Bottjer, 2001). We applied their measures of change in syllable probability and motif probability to examine the degree to which both individual syllables and specific song motifs were maintained after chronic exposure to WN.

Changes in syllable probability were defined as the mean difference between the probability with which syllables were sung before and after WN exposure. We computed the probability of each syllable being sung (i.e., the proportion of songs that included that syllable) both before WN exposure and at 1 d (D1) and 100 d (D100) after WN was discontinued. The mean absolute value of the difference between all PRE syllables and corresponding D1 syllables is the syllable probability change score for each bird, and it provides an overall measure of the amount of shift in the regularity with which song syllables were included in the song. Deleted syllables, as well as a decreased frequency of producing specific PRE syllables, can contribute to higher syllable probability scores.

Changes in motif probability were computed by determining the proportion of common motif variants present within each recording session both before and after WN exposure. That is, we identified the motif variant sung most often by each bird and added, in turn, the variants next most often sung until the list of variants accounted for at least 80% of the songs sung within each recording session. We then determined the percentage of original (i.e., PRE) common motif variants retained as common motif variants in subsequent D1 recordings for each bird as a measure of shifts in the composition of songs. Thus, a low score on this measure reflects a low carryover of initial motifs. We also compared D100 motifs to corresponding D1 motifs to examine recovery of song behavior after removal from WN.

Finally, for syllables that had been scored as changed at D100, two independent observers then examined isolated exemplars of each of those syllables and judged whether the D100 iteration of each syllable was more similar to the D1 or the PRE version. In all cases, qualitative assessments of syllable change at D1 and D100, and similarity of D100 syllables to pre-WN versus D1 iterations, were confirmed using sound analysis software to generate measures of syllable accuracy, which quantifies degree of similarity for individual syllables (see below and Tchernichovski et al., 2000).

Spectral analyses

We quantified overall spectral changes in song as well as in individual syllables using Sound Analysis software (Tchernichovski et al., 2000) to compare song behavior before WN (PRE), immediately after WN (D1), and after tutoring (D100). The software analyzes song along four separate dimensions (pitch, Wiener entropy, frequency-modulated noise, and spectral continuity) to arrive at a generic measure of similarity. This method of similarity analysis is superior to methods in which fast Fourier transforms of whole syllables are compared with one another via cross-correlation because it extracts dimensions that are relatively insensitive to normal variability in song production and recording (Tchernichovski et al., 2000). We also performed these analyses on the recordings from control birds made 3 months apart for comparison with experimental birds.

We used three different measures to characterize different aspects of change in song behavior: percentage similarity, accuracy, and spectral stereotypy. All spectral measures were computed by comparing 10 samples of song in each condition to 10 samples in another condition. Tests of significance for all spectral measures were performed using ANOVA with an α level of 0.05.

Percentage similarity. We analyzed overall spectral similarity of whole motifs using the "percentage similarity" measure to compare song samples at D1 and D100 to PRE songs. To generate this score for a pair of song samples, sound analysis compares the spectral characteristics of the first

song sample to those of the second in 20 msec bins and then calculates the percentage of bins in the first song that exceeds the similarity threshold (default value, 92%) in the second, for a maximum percentage similarity of 100%. We compared PRE songs to D1 and D100 songs for each bird independently using the "overall" setting. We also compared samples of PRE song to other samples of PRE song as a control (note that this is the same as the "spectral stereotypy" measure below). At each time point, the dominant motif produced by each bird was used in these analyses.

Accuracy. Sound Analysis provides a second measure, "accuracy," that we found to be better adapted for the analysis of individual syllables. Rather than compute the proportion of one sound that is similar to another sound, it reflects the degree to which two sounds are similar. For all parts of a pair of samples that exceed a given similarity threshold, the accuracy measure calculates a weighted average of the difference across the four dimensions analyzed. To ensure that as much of each syllable as possible was taken into account when computing accuracy, we set the similarity threshold to 0.15. Thus, for each pair of song samples, sound analysis generated an accuracy score, which was a weighted average of the similarity in pitch, entropy, frequency modulation, and continuity for all sections that exceeded the 0.15 similarity threshold. For each group, the mean accuracy score represents the mean for all syllables sung by birds in that group. Accuracy scores for D1 relative to PRE and for D100 relative to PRE were computed for all birds. To determine whether changes in accuracy as a result of WN exposure were within the range of normal variability, we compared these scores to scores generated by comparing samples of PRE song to other samples of PRE song as well as to measures from control subjects.

Spectral stereotypy. Finally, we used Sound Analysis to examine the variability in spectral properties of song within a given recording session, to generate what we call a spectral stereotypy measure. We measured spectral stereotypy by comparing randomly selected samples of song (whole motifs) within PRE, D1, and D100 recording sessions for experimental and control birds using the percentage similarity measure in sound analysis. This allowed us to observe changes in the stability of the spectral properties of song before and after WN and after tutor exposure. If a bird's song is highly stereotyped, renditions of individual syllables recorded at a given time point should be very similar to one another, and the spectral stereotypy score will be high. If there is a great deal of variability in the spectral properties of song, this score will be low.

Results
Evoked potentials

As shown in Figure 1, auditory brainstem responses revealed a substantial shift in hearing threshold immediately after removal from WN. The mean threshold shift at the best frequency (3 kHz) was 27 dB 1 hr after removal from WN, although the general shape of the frequency sensitivity curve was preserved. Hearing showed substantial recovery by 3 d after removal from WN but was still ∼5–10 dB above that of normal controls. Hearing thresholds at best frequencies had recovered to normal after 1 week out of WN, although some deficits remained at very low (750 Hz and below) and very high (>4 kHz) frequencies.

Effects of WN exposure on song
Qualitative description

Chronic exposure to WN exerted substantial disruptive effects on song behavior for the majority of birds, although a great deal of individual variability was observed. Figure 2 shows two typical examples of the most severe forms of disruption to song observed on D1, immediately after removal from WN. The effect of WN on song was comparable with these examples for nine birds: four from group 1, one from group 2, none from group 3, and four from group 4. In these cases, at least half of the syllables present in PRE song were visibly changed (e.g., syllables B and C in the song of subject y685) or deleted altogether (e.g., syllable D in the same subject). Syllables characterized by rapid frequency modulation in the PRE song tended to be replaced by broadband noise or
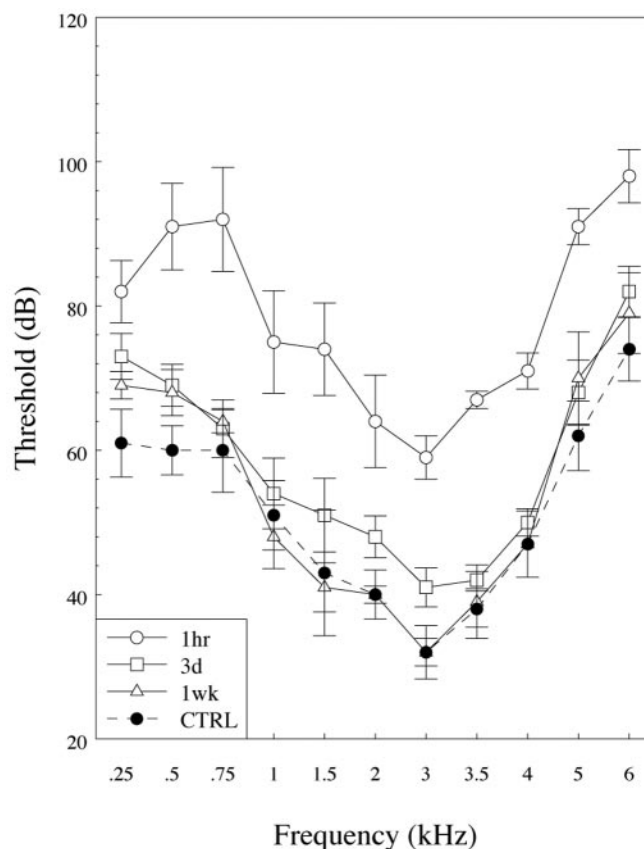

Figure 1. Auditory brainstem recordings reveal a temporary increase in hearing threshold. Error bars indicate SD.

high-pitched whistle-like syllables (such as syllable B in the song of pu746). Furthermore, some call-like notes were produced at a higher pitch in the post-noise song (e.g., syllable C in the song of bird y685). Finally, subsyllabic structures or "notes" were deleted from some complex syllables. For example, compare syllable B in the PRE song of bird y685 to the same syllable in the D1 song: in addition to gross changes in the spectral properties of this syllable, an entire note (the downsweep in the last 10 msec of the syllable) was deleted. This kind of syllable truncation was observed in five of the severely disrupted birds.

The deletion of particular syllables was the most common effect of WN exposure on the temporal structure of song. In the case of y685 (Fig. 2) the temporal sequence of song changed because this bird dropped syllable D altogether and omitted syllable A after a bout onset, which produced C′→B′ transitions that were not present in the PRE song. Other forms of disruption to the temporal structure of song were less commonly observed. For example, in one instance, the order of two syllables was reversed (syllables D and E for y692) (see Fig. 5). In the case of pu746 (Fig. 2), the D1 song lacked temporal structure altogether and bore little resemblance to the PRE song. Two other severely disrupted birds displayed other temporal abnormalities, repeating one or more call-like syllables at the end of a bout (similar to the repeated Cs in the song of pu746).

In contrast to the severely disrupted subjects, exposure to WN induced little or moderate disruption to song in 13 birds, as typified by the examples shown in Figure 3. In the songs of these birds, more than half of the syllables present in PRE song were visible in song samples from D1 and appeared to be relatively unchanged. We refer to these subjects collectively as "moderately
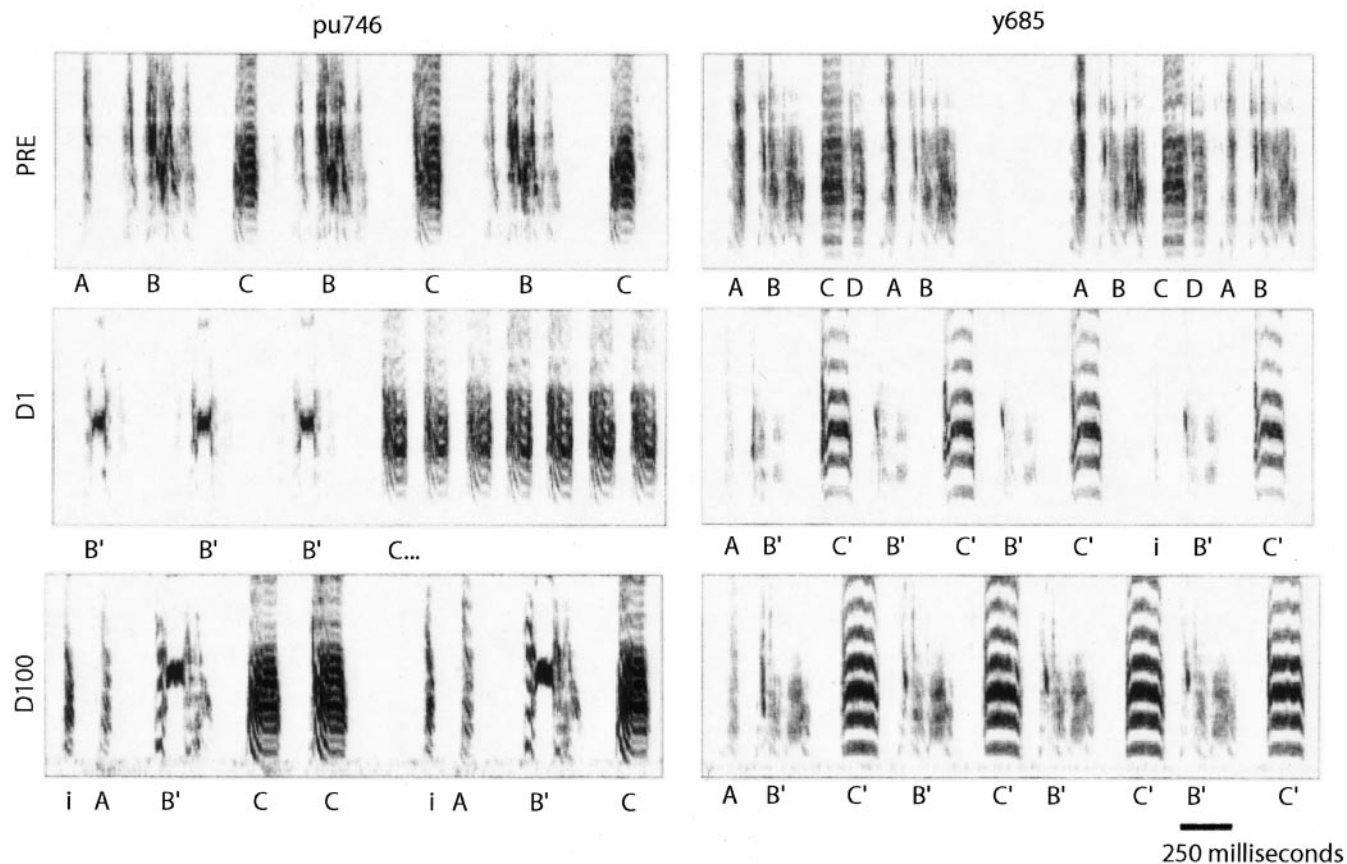
Figure 2. Examples of severely disrupted song: a comparison between song before WN exposure (PRE) and immediately after discontinuation of WN (D1). Syllables labeled with a prime (') were scored as visibly changed between PRE and D1. The *y*-axis is 0 –10 kHz.

disrupted": two were from group 1, four were from group 2, six were from group 3, and one was from group 4. In six of these birds (e.g., y746) (Fig. 3) very little disruption was observed in visual inspection of the spectrograms, although quantitative measurements revealed subtle changes (see below). In the remaining seven birds, changes were more similar to, although less widespread than, those observed in severely disrupted birds. For example, segments of some syllables (notes) characterized by rapid frequency modulation in the PRE song were replaced by broadband noise (e.g., syllable C in the song of subject y686). Syllable truncation similar to that described for severely disrupted birds was observed in two moderately disrupted birds. Temporal changes (aside from loss of entire syllables) among moderately disrupted birds were limited to one instance in which four novel syllables were appended to the end of some motifs.

The large individual differences we observed in degree of behavioral disruption did not bear a direct relationship to differences in treatment. As indicated above, exposure to WN was initially interrupted at weekly intervals for birds in groups 1 and 2 in the hopes of being able to chronicle the time course of changes in song behavior (see Materials and Methods). An initial report of data from groups 1 and 2 alone (Zevin et al., 2000) suggested that this weekly interruption of noise may have prevented the disruptive effects of WN because group 2 was exposed to a longer course of interrupted noise and was, overall, less disrupted. The effect of interruptions of WN, if any, is unclear in light of the data from all four treatment groups, however. Birds in group 3 were exposed to uninterrupted WN for 1 month longer than the group 1 subjects, but none of these birds were severely disrupted. Group 4 was in

continuous WN for 7 months, and nearly all of the subjects in this group were severely disrupted. It may be that, for most birds, long periods (>6 months) of uninterrupted exposure to WN are necessary to induce the extreme changes in song typified by the examples presented in Figure 2. Preliminary analyses of similarity to the father's song, circulating testosterone levels, and syllable types in pre-noise song suggested that these factors were not predictive of the degree of song deterioration in WN (data not shown). Interestingly, this degree of variability in song disruption is also observed in studies using surgical deafening (Nordeen and Nordeen, 1992; Brainard and Doupe, 2001).

Analyses of temporal structure
Data for each subject from song sequence analyses and syllable and motif probability scores are presented along with group means in Table 1 for experimental and control subjects. Exposure to WN did not exert substantive effects on measures of linearity and consistency, even for birds with severely disrupted songs (see below). Some birds actually showed an increase in sequence stereotypy on the first day out of WN, as indicated by higher scores for linearity and consistency. The absolute change in linearity scores was different from controls in group 3 (U = 3; $p < 0.05$), and the change in consistency scores was different from controls in groups 3 and 4 (U = 3; $p < 0.05$ in both cases). There were no significant differences between any other experimental groups and controls. Table 1 shows that even severely disrupted songs had high linearity and consistency scores, indicating that although these songs differed in various respects from the PRE song, they were, nevertheless, stereotyped at the level of temporal
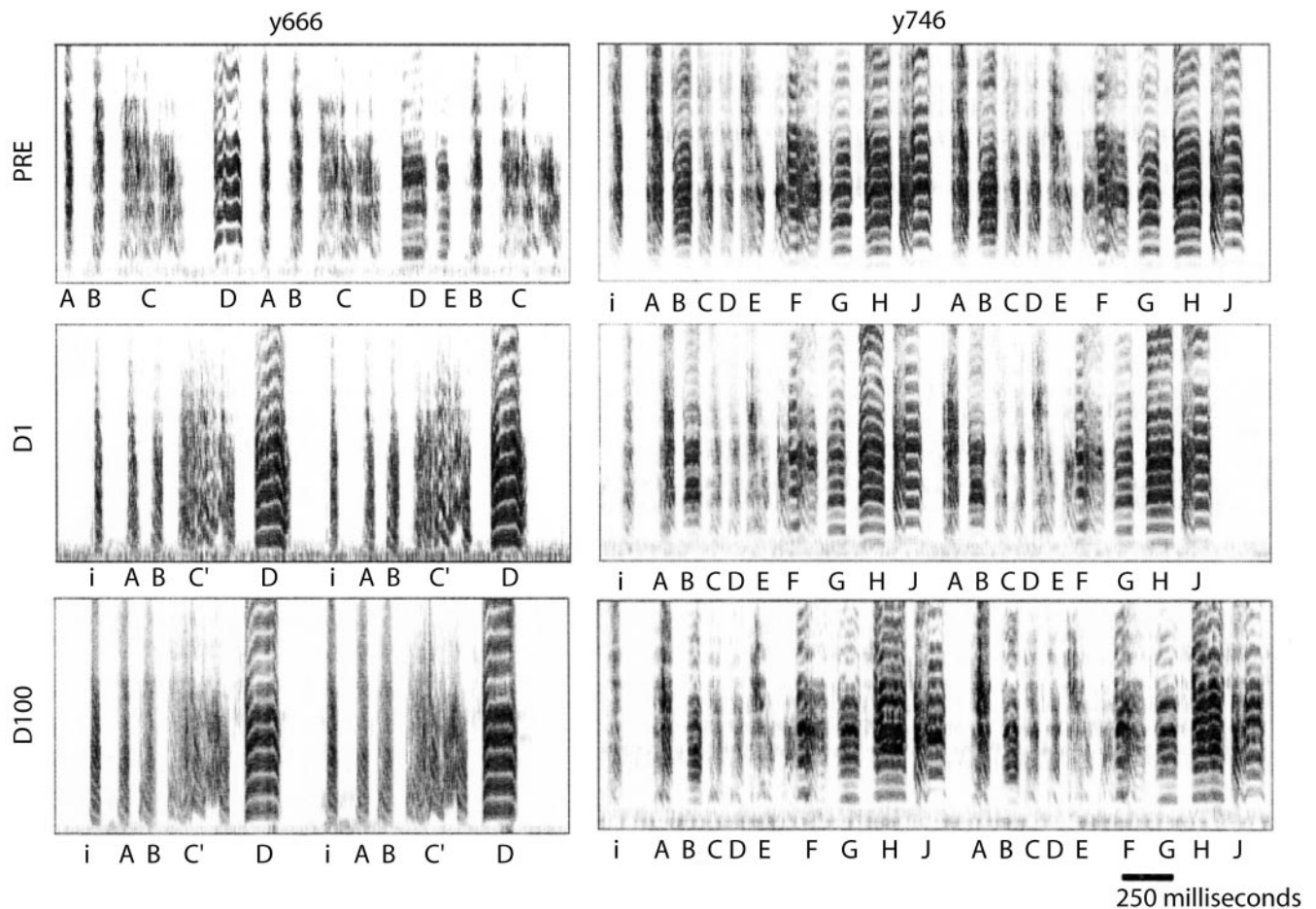
Figure 3. Examples of moderately disrupted song: a comparison between song before WN exposure (PRE) and immediately after discontinuation of WN (D1). The *y*-axis is 0–10 kHz.

structure. These results must be interpreted with caution, however, because the most severely disrupted birds tended to have very few syllables in their motifs. The number of syllables in a motif is inversely related to the lowest possible linearity and consistency scores attainable (Foster and Bottjer, 2001), therefore severely disrupted birds may have artificially inflated linearity and consistency scores (note, for example, that no score is given for pu762, whose song on D1 consisted of a single syllable). Another consideration in assessing changes in song with these measures is that they are not sensitive to wholesale deletions and changes of individual syllables. For example, a bird that consistently sings ABCDE before WN but consistently sings ACD afterward may have comparable PRE and D1 linearity and consistency scores despite obvious changes to song structure.

To generate additional measures of change between PRE and D1 song behavior, two independent observers examined complete motifs of song and scored syllables as lost or changed (see Materials and Methods). Whereas no syllables were scored as lost or changed for control birds, 16 of 22 birds exposed to WN lost or changed at least one syllable; a total of 37 syllables were scored as changed and 19 as lost. The accuracy measure, which quantifies the degree to which two sounds are similar (see Materials and Methods), was used to confirm the scoring of syllables as "changed" as determined by visual inspection of spectrograms. Accuracy scores for the 37 syllables scored as changed ranged from 18.0 to 66.6 (57.1 ± 8.2; mean ± SD), whereas accuracy for the remaining 62 syllables ranged from 71.7 to 85.9 (79.1 ± 3.8). We computed the proportion of all syllables in a bird's PRE song

that were lost or changed on D1 (Table 1). On this measure, significant differences from the control group were observed for group 1 (U = 2.5; $p < 0.05$), group 2 (U = 0; $p < 0.05$), and group 4 (U = 0; $p < 0.05$), but not group 3 (U = 7; $p > 0.05$). Figure 4 shows that the proportion of lost or changed syllables was high in severely disrupted birds on D1 and low in moderately disrupted birds (and see below).

We also computed syllable and motif probability scores, which capture probabilistic changes in temporal structure of song not reflected in the categorical distinctions involved in the qualitative measures of song sequence described above. The syllable probability measure is a difference score that reflects changes in the regularity with which specific syllables are sung (see Materials and Methods). All WN groups had higher scores on this measure than did controls, indicating decreased conservation of syllable production after exposure to WN. This difference was significant for all groups [group 1 (U = 1; $p < 0.05$), group 2 (U = 0; $p < 0.05$), group 3 (U = 3.5; $p < 0.05$), and group 4 (U = 0.5; $p < 0.05$)]. The motif probability measure reflects change in the regularity with which the most common motif variants are sung. All WN groups had low scores on this measure, indicating a low carryover of common motifs from PRE song, but these scores were significantly lower than controls only for group 1 (U = 5; $p < 0.05$) and group 2 (U = 2.5; $p < 0.05$). Overall, these measures suggest that WN resulted in changes in the temporal structure of song, even in cases in which gross abnormalities were not observed. However, as indicated above, changes in temporal sequence tended to be secondary to loss of syllables.

Spectral analyses

We also observed effects of deprivation of auditory feedback on the spectral properties of song. Summary statistics for the three measures we generated using Sound Analysis software (Tchernichovski et al., 2001) are presented in Table 2.

The percentage similarity measure reflects the proportion of the most common D1 song motif that was above a specified threshold (0.92) for similarity to PRE song on the dimensions analyzed by sound analysis (pitch, frequency modulation, Wiener entropy, and continuity). Relative to within-subject control comparisons performed within a PRE recording session (Table 2, first column), all experimental groups showed a significant reduction in similarity to PRE song at D1 [group 1 $(F_{(1,5)} = 84.3; p < .001)$, group 2 $(F_{(1,4)} = 19.5; p < 0.05)$, group 3 $(F_{(1,5)} = 17.0; p < 0.01)$, group 4 $(F_{(1,4)} = 17.3; p < 0.05)$]. This decrease in similarity to the PRE song reflects deletion of syllables and various changes to spectral properties of retained syllables, as described above.

The mean syllable accuracy score reflects changes in syllable structure for only those syllables not lost between time points. Accuracy scores derived from comparisons of D1 song to PRE song over all syllables were also lower than scores derived from comparisons performed within PRE recording sessions for all experimental groups [group 1 $(F_{(1,5)} = 72.3; p < .001)$, group 2 $(F_{(1,4)} = 16.3; p < 0.05)$, group 3 $(F_{(1,5)} = 14.9; p < 0.05)$, group 4 $(F_{(1,4)} = 25.1; p < 0.05)$], indicating changes in the spectral characteristics of individual syllables. Taken together, the percentage similarity and accuracy measures reflect substantial changes in the spectral properties of song as a result of exposure to WN, even for subjects whose songs appeared only moderately disrupted on visual inspection of spectrograms.

The spectral stereotypy measure reflects the degree of stability in spectral properties of song at a given point in time. Spectral stereotypy on D1 (immediately after removal from WN) decreased for birds in groups 1 and 4, which contained the highest number of severely disrupted birds, although this change was not statistically reliable (Table 2). When the most severely disrupted birds are considered separately (Table 3), this effect does reach statistical significance, indicating that in the most disrupted subjects, spectral properties of D1 song were more variable than PRE song $(F_{(1,20)} = 8.04; p < 0.01)$, in addition to the spectral changes revealed by the percentage similarity and accuracy measures.

All three of these spectral measures were also examined in control birds, and none of the measures changed significantly between PRE and D1 recording sessions (all $F < 1$). In summary, deprivation of auditory feedback via exposure to WN caused changes in adult song patterns, although the degree of disruption was variable. Some subjects were severely disrupted, such that their D1 song bore little resemblance to their PRE song, whereas in other cases little disruption was observed. Song disruption was characterized by deletions and spectral changes to syllables, indicated by quantitative comparison of PRE and D1 song, as well as visual inspection. Furthermore, spectral properties of song were
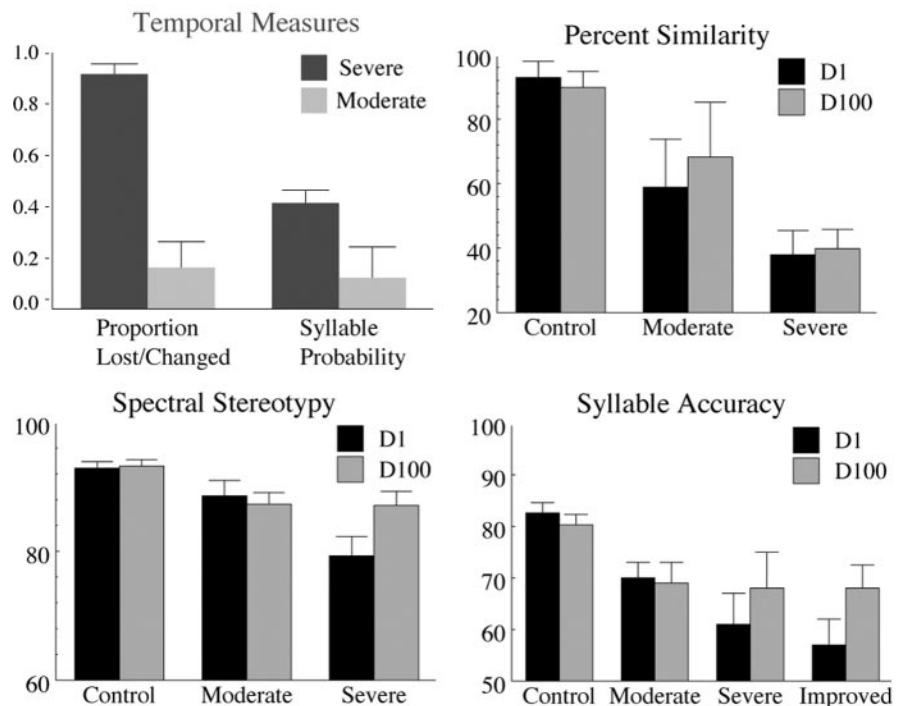
Figure 4. Clockwise from top left panel: mean proportion of syllables lost or changed and syllable probability scores on D1 for severely disrupted (■) and moderately disrupted (▨) birds; mean percentage similarity; syllable accuracy for control, severely disrupted, and moderately disrupted birds at D1 (■) and D100 (▨); and spectral stereotypy. For control birds, results are for two time points separated by 3 months. Accuracy scores are also given for those syllables scored as increasing in similarity to their PRE renditions at D100. Error bars represent SE.

less stereotyped immediately after exposure to WN. In contrast to widespread spectral disruptions to song, the effect of WN exposure on temporal aspects of song structure was much less pronounced. Motif probability and syllable probability measures revealed subtle, but significant, changes in song sequence between PRE and D1, but aside from changes resulting from the deletion of particular syllables, instances of gross disorganization of temporal structure were rare (cf. Leonardo and Konishi, 1999).

Changes in song after removal from WN

After WN was discontinued, birds were exposed to a taped (group 1) or live (groups 2–4) tutor starting immediately after the D1 recording. Songs were recorded again at D100 and compared with D1 and PRE song. As noted above, there was substantial variability in the disruptive effect of WN on song that was not directly related to group membership. In addition, preliminary analyses suggested broad differences between severely and moderately disrupted birds in their recovery after restoration of auditory feedback. Therefore, we decided to analyze recovery as a function of degree of disruption measured immediately after removal from WN. We defined severely disrupted birds operationally as those for whom half or more of all song syllables were lost or changed as a result of WN exposure ($n = 9$ on D1). Moderately disrupted birds were those for whom fewer than half of all song syllables were lost or changed ($n = 13$ on D1).

To validate the categorization of birds into "severe" and "moderate" groups, we performed spectral analyses for birds in the two groups by comparing their PRE song behavior to that recorded on D1. As shown in Table 3 (see also Fig. 4), the severe group was more disrupted than the moderate group, as judged by all three measures used: percentage similarity $(F_{(1,20)} = 7.03; p < 0.01)$, accuracy $(F_{(1,19)} = 5.10; p < 0.05)$, and spectral stereotypy

Table 2. Spectral analyses: percentage similarity, accuracy, and spectral stereotypy of pre-WN and D1 song

	Percentage similarity			Mean syllable accuracy			Spectral stereotypy		
	PRE	D1	Diff	PRE	D1	Diff	PRE	D1	Diff
Group									
1	90.7	42.2	48.5***	83.4	61.7	21.7***	90.7	81.9	8.8
2	87.4	52.9	34.5*	80.4	67.8	12.6*	87.4	88.4	−1.0
3	87.4	61.4	26.0**	86.4	73.2	13.2*	87.4	85.0	2.4
4	87.2	43.5	43.7*	82.1	60.5	21.6*	87.2	80.6	6.6
Control	92.9	89.8	3.1	82.6	80.3	2.3	92.9	93.2	−0.3

PRE, Mean within-subject control scores for comparisons made within pre-WN recording sessions; D1, mean scores for comparisons between D1 song and PRE song; Control, control birds not exposed to WN, recorded 3 months apart. The spectral stereotypy is the percentage similarity computed within each recording period (PRE and D1) and the difference between them (Diff); percentage similarity for PRE is thus equivalent to spectral stereotypy. See text for additional details.

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table 3. Spectral analyses of D1 song (relative to PRE) for moderately versus severely disrupted birds

|	Percentage similarity	Syllable accuracy	Spectral stereotypy
Moderate	58.9	70.3	88.6
Severe	37.9	61.1	79.3
Difference	21.0**	9.2*	9.3**

*$p < 0.05$; **$p < 0.01$.

$(F_{(1,20)} = 8.04; p < 0.01)$. Interestingly, categorizing birds in this manner resulted in a particularly pronounced difference in the incidence of syllable loss, as opposed to syllable change: severely disrupted birds lost a total of 18 PRE syllables from their D1 songs, whereas moderately disrupted birds lost only one syllable; in contrast, severely disrupted birds changed a total of 23 of their syllables compared with 14 in moderately disrupted birds. Thus, severe disruption of song because of removal of auditory feedback seems to entail a disproportionate loss of syllables from stable vocal patterns.

Qualitative description
Song change after removal from WN provided little evidence for new learning. Unlike results reported for Bengalese finches (Woolley and Rubel, 2002), we observed no evidence for imitation of either live or taped tutors. We also observed only limited evidence for reacquisition of previous song patterns. Song did gradually become more stereotyped in many instances, but this was often not the result of increasing similarity to PRE song. Instead, D100 song syllables frequently evolved into somewhat "cleaned up" or less variable versions of those produced in the D1 song. Finally, PRE syllables that were deleted entirely in D1 songs were never reintroduced.

Figure 5 shows an instance of song increasing in stereotypy without increasing in similarity to PRE song. Subject y692 produced three novel syllables on D1 (x, y, z), which were noisy and variable. On D100, although the song had not become more similar to the PRE song, the structure of these syllables had become much more regular. Figure 2 also shows an interesting example of increasing stereotypy of spectral properties. Syllable B in the song of pu746 changed in two ways after removal from WN: the first note of this syllable became a relatively high-pitched stack, which was more stereotyped in both duration and spectral characteristics in the D100 song than the low-amplitude noise in the D1 song but did not resemble any note in the PRE song. Interestingly, there was some evidence for reacquisition of elements of PRE song within the same syllable. The last note in syllable B increased in relative amplitude and developed a more structured frequency-modulated component in the D100 song, potentially reflecting reacquisition of an element of PRE song.
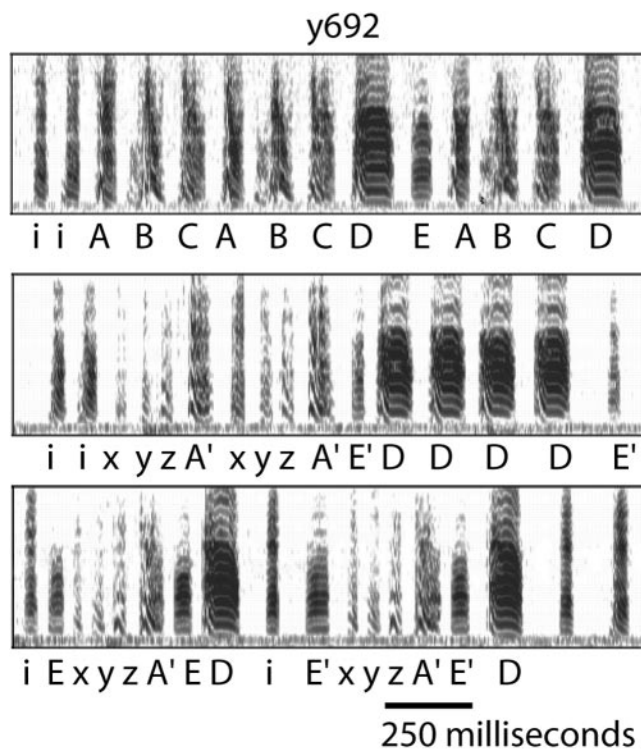
Figure 5. An example of increasing stereotypy without increasing similarity to initial song. The *y*-axis is 0–10 kHz.

We observed other evidence for reacquisition of PRE song, in particular with respect to the spectral properties of individual syllables. In some subjects, subsyllabic structures known as notes had been deleted as a result of WN exposure (see above). At D100, many of these syllables (two of two among moderately disrupted birds, four of five among severely disrupted birds) had regained some of their pre-WN structure. Two examples of this phenomenon are presented in the two left panels of Figure 6, one from a moderately disrupted bird (y665) and another from a severely disrupted bird (y685). The short, frequency-modulated downsweep at the end of each syllable in the PRE song is lacking in the songs of both birds at D1 but is present again in the D100 recordings. Figure 6 (right) shows an example of an apparently similar recovery process in pu781, a severely disrupted bird. Frequency-modulated sections of this syllable tended to be replaced by broadband noise in D1 recordings, in addition to truncation of the end note of this syllable. At D100, much, although not all, of the original structure had returned. A similar but less dramatic example of the same type of recovery is observable in syllable C in the song of y686 (Fig. 3).
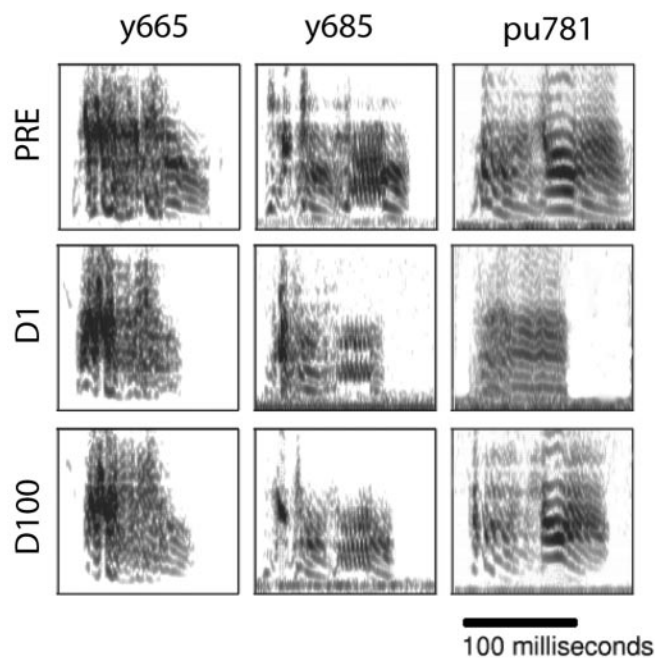
Figure 6. Recovery of individual syllables for three birds. The *y*-axis is 0 –10 kHz.

Analyses of temporal structure

Measures of sequence stereotypy for song behavior on D1 versus D100 are presented in Table 4 for moderately and severely disrupted birds. Linearity scores remained stable or increased between D1 and D100 in all but one bird. This subject (y666) underwent other changes to song, including syllable deletion, that were not typical of song changes observed during the tutoring period in other birds. The change in linearity was not significant for either group. Consistency scores also tended to increase after removal of WN in both moderately and severely disrupted birds. This effect was reliable for severely disrupted birds only (U = 7.5; $p < 0.05$).

As shown in Table 4, very few syllables were lost between D1 and D100; the number of syllables lost was not significantly different from controls (Table 1) for either group. One moderately disrupted bird (y666) lost most of his syllables after 1 week out of WN. This was the only instance in which a moderately disrupted bird became more disrupted after removal from WN. Only one instance of syllable loss was recorded among severely disrupted birds. In contrast, 18 syllables were judged as visibly changed between D1 and D100, including 13 syllables from severely disrupted birds and 5 syllables from moderately disrupted birds. Of the 37 syllables that changed between PRE and D1 (Table 1), fewer than half (15 of 37) changed further between D1 and D100. The number of syllables changed was significantly different from controls for severely disrupted birds only (U = 5; $p < 0.05$). Syllable accuracy scores comparing D100 to D1 iterations confirmed the scoring of syllables as changed: accuracy scores for the 18 syllables judged as changed on D100 ranged from 39.5 to 64.8 (52.9 ± 7.1; mean ± SD), whereas accuracy scores for the 22 syllables that had been judged as changed on D1, but not on D100, ranged from 69.0 to 86.3 (74.9 ± 5.0).

Evaluation of changed syllables

To determine the proportion of syllables that increased in similarity to PRE song between D1 and D100, we undertook two separate analyses (see Materials and Methods). First, as a follow-up of our initial analysis of judging lost and changed syl-

lables in the context of whole motifs of song, we examined isolated renditions of all syllables scored as obviously changed at D100 on an expanded time scale (as in Fig. 6) and for each determined whether typical renditions on D1 or D100 were more similar to a typical PRE rendition. In this analysis, 14 of the 18 syllables scored as changed at D100 were found to have increased in similarity to PRE between D1 and D100 (10 of 13 from severely disrupted birds and 4 of 5 from moderate). In accordance with these subjective judgments, syllable accuracy scores for these syllables were significantly higher on D100 (67.8 ± 15.5) than on D1 (mean, 57.6 ± 17.8) ($F_{(1,13)}$ = 4.75; $p < 0.05$).

This initial analysis was asymmetrical in the sense that it involved syllables that had been identified as being most changed in D1 song, perhaps increasing the likelihood of finding an increase in similarity to PRE on D100. Therefore, we conducted a second analysis in which isolated exemplars of all syllables at all three time points (PRE, D1, and D100) were examined on an expanded time scale, and both observers were blind to the identity of each syllable as well as to the date of the later renditions (PRE renditions were labeled for comparison). Of all 97 syllables scored in this way at all three time points, 50 were scored as changed: 15 on D1 only, 2 on D100 only, and 33 at both time points. Of all 50 syllables scored as changed, 24 were judged as having increased in similarity to PRE between D1 and D100, 5 were judged as having decreased in similarity between D1 and D100, and 21 remained unchanged between D1 and D100.

All 14 syllables that had been rated as increasing in similarity to PRE in the initial analysis were also rated as more similar to PRE on D100 than on D1 in the blind analysis. Furthermore, the blind analysis of all syllables (as opposed to only those scored as obviously changed at D100) revealed 13 additional syllables (8 from moderate birds and 5 from severe) that changed between D1 and D100 on the basis of differing in similarity scores at these two time points. Of these, 10 (5 moderate and 5 severe) were rated as more similar to PRE on D100 than on D1. As shown in Figure 4, accuracy scores were then computed for all 24 syllables identified as increasing in similarity to PRE and again confirmed the subjective judgments (D1, 59.1 ± 17.2; D100, 71.7 ± 14.0; $F_{(1,23)}$ = 15.47; $p < .001$).

In summary, both analyses provided evidence for some reacquisition of syllable structure. However, evidence for reacquisition of syllable structure was found in only ~25% of all syllables examined (although in ~75% of syllables judged as changed between D1 and D100). Furthermore, these increases in similarity were often quite small, as evidenced by the fact that concomitant increases in accuracy scores did not always result in a complete recovery of the PRE syllable. In 14 cases, accuracy scores on D100 were within the range of syllables not scored as changed on D1, reflecting potentially complete reacquisition. However, eight syllables remained within the accuracy range of syllables scored as changed, and two were intermediate. In some instances, specific notes within a syllable improved (or were reacquired, as described above), whereas other notes remained disrupted, such that the entire syllable did not recover its pre-WN morphology. This gave rise to an overall pattern in which, despite clear isolable examples of recovery of individual syllables, no instances of global song recovery were observed in severely disrupted birds. For example, Figures 3 and 6 contain examples of recovery of syllable structure in the context of songs that remain grossly disrupted.

Table 4. Analyses of temporal structure of song for all subjects on D100 (100 d after removal from WN), relative to D1 (immediately after removal from WN)

	Linearity			Consistency			Syllable probability	Motif probability	Syllables changed	Syllables lost	Proportion lost/changed
	D1	D100	Diff	D1	D100	Diff					
Severe											
y685	0.67	1.00	−0.33	0.88	0.95	−0.07	0.08	1.00	2	0	0.67
y686	1.00	1.00	0.00	0.99	1.00	−0.01	0.01	1.00	1	0	0.25
y692	1.00	1.00	0.00	0.94	0.99	−0.05	0.21	0.50	1	0	0.20
y693	0.80	1.00	−0.20	0.92	0.91	0.01	0.37	0.00	3	0	0.50
y696	0.75	0.67	0.08	0.97	0.90	0.07	0.27	1.00	0	0	0.00
Pu746	0.75	1.00	−0.25	0.78	0.88	−0.10	0.40	0.67	1	1	0.67
Pu750	1.00	1.00	0.00	0.93	1.00	−0.07	0.08	1.00	1	0	0.25
Pu762	**	**	**	**	**	**	**	**	0	0	0.00
Pu781	1.00	1.00	0.00	0.95	0.93	0.02	0.14	0.50	4	0	0.80
Means	0.87	0.96	0.11	0.92	0.95	0.05*	0.20*	0.71	1.44*	0.11	0.37*
Moderate											
y663	1.00	1.00	0.00	0.99	0.99	0.00	0.17	0.00	0	0	0.00
y665	1.00	1.00	0.00	1.00	0.94	0.06	0.09	1.00	1	0	0.20
y666	0.83	0.67	0.16	0.97	0.88	0.09	0.57	0.00	2	4	1.00
y677	0.78	0.78	0.00	0.91	0.90	0.01	0.18	0.50	0	0	0.00
y682	1.00	1.00	0.00	0.95	0.98	−0.03	0.09	0.50	0	0	0.00
y695	1.00	1.00	0.00	0.96	0.94	0.02	0.27	0.00	0	2	0.25
y701	0.57	0.80	−0.23	0.87	0.92	−0.05	0.08	1.00	0	0	0.00
y746	0.71	1.00	−0.29	0.89	0.99	−0.10	0.16	1.00	0	0	0.00
y747	1.00	1.00	0.00	0.99	1.00	−0.01	0.04	1.00	0	0	0.00
y748	1.00	1.00	0.00	0.93	0.94	−0.01	0.02	1.00	1	0	0.25
y750	1.00	1.00	0.00	0.94	0.97	−0.03	0.27	1.00	1	0	0.11
Pu776	1.00	1.00	0.00	0.98	0.99	−0.01	0.01	1.00	0	0	0.00
W696	1.00	1.00	0.00	1.00	0.90	0.10	0.10	0.00	0	0	0.00
Means	0.91	0.94	0.05	0.95	0.95	0.04	0.16*	0.62	0.38	0.46	0.14

D1, 1 d after WN was discontinued; D100, 100 d after WN was discontinued; Diff, difference between D100 and D1 song; syllable probability, the change in syllable probability between D1 and D100; motif probability, the proportion of motifs at D100 conserved from D1; proportion lost/changed, the sum of lost and changed syllables at D100 divided by the number of syllables in D1 song.

*Significantly different from controls, $p < 0.05$. **No scores generated were for pu762 because song consisted of a single syllable.

Table 5. Spectral analyses: percentage similarity, accuracy, and spectral stereotypy for birds with moderately and severely disrupted songs immediately after removal from WN (D1) and after 100 d of tutoring (D100)

	Percentage similarity			Mean syllable accuracy			Spectral stereotypy		
	D1	D100	Diff	D1	D100	Diff	D1	D100	Diff
Group									
Moderate	58.9	68.2	−9.3*	70.3	68.7	1.6	88.6	87.3	1.3
Severe	37.9	39.8	−1.9	61.1	68.3	−7.2*	79.3	87.1	−7.8**

Diff, Difference between D1 and D100. Significance values are for comparisons between D1 and D100 scores. Percentage similarity and syllable accuracy are each computed as comparisons with PRE song (i.e., D1 vs PRE and D100 vs PRE). Spectral stereotypy is computed within each recording period (i.e., this score represents percentage similarity within D1 and within D100). See text for additional details.

*$0.1 > p \geq 0.05$; **$p < 0.01$.

Spectral analyses

Percentage similarity and syllable accuracy measures were generated relative to PRE song at both D1 and D100 (Table 5). For moderately disrupted birds, a marginally significant increase in the percentage similarity to PRE song was observed between D1 and D100 ($F_{(1,12)} = 4.6$; $p = 0.05$). This suggests that although relatively few syllables changed visibly during this time in the songs of moderate birds, overall spectral properties of song evolved to become more similar to the birds' initial songs. No change in mean accuracy was observed ($F < 1$). Because changes in the accuracy measure reflect mean changes in similarity to PRE song across individual syllables, this result further suggests that increases in percentage similarity were the result of changes in general spectral properties of song. No significant change in spectral stereotypy was observed for moderately disrupted birds ($F < 1$), although this may represent a ceiling effect: spectral stereotypy scores at D1 were well within the range of similar scores measured before exposure to WN (Table 2), indicating a stable spectrotemporal organization of song at both D1 and D100.

For severely disrupted birds, no change was observed in percentage similarity to PRE song between D1 and D100 ($F < 1$).

Instead, a relatively large, but nonsignificant, increase in accuracy was observed ($F_{(1,7)} = 2.39$; $0.1 > p > 0.05$). This pattern of results may reflect the fact that whereas many syllables that had changed on D100 became more similar to their instantiations in PRE song, others evolved to become more stereotyped versions of syllables sung on D1 and many other syllables were not scored as changed between D1 and D100 (although a decrease in the variability from rendition to rendition was observed, as noted above). This decrease in variability of syllable production was reflected by a significant increase in the spectral stereotypy score between D1 and D100 ($F_{(1,8)} = 8.0$; $p < 0.01$) for severe birds. Whereas spectral stereotypy scores at D1 were significantly lower than within-subject PRE comparisons, scores from D100 were within normal range (compare Tables 2, 5).

Finally, we examined the influence of tutoring by comparing whole motifs of experimental birds to motifs of their tutors using the percentage similarity measure. Tutors were selected to be maximally dissimilar from tutees, so the accuracy measure was not feasible for these analyses because there were no syllables in the tutor song that corresponded to PRE song for the experimen-

Table 6. Similarity to tutor song immediately after removal from WN (D1) and after 100 d of tutoring (D100)

	Percentage similarity		
	D1	D100	Diff
Group			
Moderate	45.1	44.5	0.6
Severe	40.0	38.5	1.5

Diff, Difference.

tal birds. As shown in Table 6, there was no evidence of increasing similarity between tutor songs and songs of experimental birds among either severely or moderately disrupted subjects.

In summary, change in song during the first 100 d after removal from WN included some general improvement and evidence for limited reacquisition of PRE song elements, but birds whose songs were most disrupted by WN exposure failed to reestablish their previous adult song patterns. In a limited number of cases, aspects of PRE song were reacquired; in particular, subsyllabic structures that had been deleted at D1 were restored, and some syllables that changed after removal from WN became more similar to their PRE iterations. However, most song change could not be characterized as reacquisition of this sort. Of the syllables that were scored as changed on D1, the majority did not change further by D100, but remained disrupted. In addition, severely disrupted birds lost a large number of syllables from their PRE songs, and these deleted syllables were never replaced in D100 songs. In many cases, changes were observed that reflected neither increasing similarity to PRE song nor increasing similarity to the song of live or taped tutors. For example, among severely disrupted birds, spectral stereotypy increased without significant increases in similarity to PRE song.

Discussion

Two critical results emerged from the current study. First, long-term WN exposure affected song in a manner similar to deafening. Second, once WN was discontinued, changes in song behavior were observed that provided only limited evidence for reacquisition of song. Song in many birds became more stereotyped without reversing the disruption induced by deprivation of auditory feedback. In some instances, birds did reacquire aspects of their initial song, particularly the spectral properties of individual syllables. However, the majority of syllables remained permanently altered even after normal auditory feedback was restored, and deleted syllables were never replaced. These permanent deficits suggest limits on the ability of adult zebra finches to recapitulate the sensorimotor phase of song acquisition. The results contrast with the near complete recovery of song observed by Leonardo and Konishi (1999) in zebra finches exposed to different types of delayed auditory feedback and with data from Bengalese finches, which demonstrated some ability to acquire novel song elements in addition to reacquiring their initial song after reversible damage to hair cells (Woolley and Rubel, 2002).

Effects of WN are similar to deafening
Long-term exposure to WN resulted in changes in spectral and temporal aspects of adult song behavior similar to those observed in surgical deafening studies (Nordeen and Nordeen, 1992; Brainard and Doupe, 2001). In the most severe cases, a majority of the syllables sung before WN exposure were lost or radically changed. Although song sequence changed as a result of lost syllables, changes in the ordering of syllables were much smaller in magni-

tude than changes in syllable phonology. Birds varied considerably in the severity of disruptions induced by WN. Interestingly, the age at which our subjects were exposed to WN corresponds to the time point at which the most variability in the effect of deafening has been observed in studies examining the influence of age at deafening on the rate and severity of song disruption (Lombardino and Nottebohm, 2000; Brainard and Doupe, 2001).

Changes in song after restoration of normal auditory feedback
Song change after removal from WN was highly variable. Although certain aspects of song were reacquired after removal from WN, recovery of original song patterns was incomplete, especially in birds with severely disrupted songs. For example, approximately half the syllables that changed as a result of exposure to WN did not improve by 100 d after noise removal. Furthermore, reacquisition of individual syllables was often partial and never resulted in complete recovery of a severely disrupted song.

The reacquisition of syllables that had been disrupted, and the overall increase in similarity to PRE song in some moderate subjects (Table 5; Fig. 6), suggests that some form of sensorimotor template is maintained in adult birds. However, it is also possible that some of these changes were the result of adjustments of fairly general parameters of song (Fee et al., 1998). For example, the deletion of the downsweep in the syllable from subject y665 on D1 (Fig. 6) could have resulted from decreased air flow alone (i.e., the syrinx may still have been producing the same gesture as in PRE and D100 song). If this were true, the reacquisition of this note would not require any new learning but only an increase in airflow between D1 and D100. This seems less likely in the case of the syllable from subject pu781 (Fig. 6); in this case, changes occurred at a number of levels, including not just general properties such as source quality and the bandwidth of harmonics, but the overall structure of the syllable as well, from downsweep–stack–downsweep to downsweep–stack–stack (for methods of classifying syllables, see Zann, 1996).

Failures to reacquire initial song structure also provide some insight into the basic mechanisms that underlie learning in the song system. For example, increases in song stereotypy that do not reflect reacquisition of song after restoration of feedback may be guided by a generic species-specific template (Marler and Peters, 1982) in addition to (or instead of) a learned template. The data are also consistent with the notion that song learning is achieved by a reinforcement mechanism, as suggested by Troyer and Doupe (2001a,b; Dave and Margoliash, 2001). Reinforcement learning proceeds by strengthening synaptic connections that lead to an outcome consistent with some target. As such, it requires that the target be approachable by successive approximation given the variability in the behavior itself. In severely disrupted birds, song output may deviate enough from the target stored in the template that this target cannot be reacquired. Instead, the reinforcement signal leads to a stabilization of highly abnormal song.

Another possibility is that the template itself remains susceptible to change even in adulthood. For example, the fact that lost syllables were never recovered could indicate that such syllables were actively "unlearned" as a result of disruption of auditory feedback, perhaps by being eliminated from the bird's initially acquired template. Mechanisms for active template changes in adult birds based on usage or experience could be similar to those used during initial song development; for example, Solis and Doupe (1997) found that responses in lateral magnocellular nu-

cleus of the anterior neostriatum (lMAN) evolved during development to be selective to a bird's own song, even when this song was manipulated to be dissimilar from the putative template song.

Temporal structure and spectral properties of syllables exhibit plasticity under different circumstances

As indicated above, the failure to reacquire lost syllables may indicate elimination of syllables from the template. This pattern is consistent with results from an experiment by Hough and Volman (2002) in which beads were implanted into the syrinx, resulting in the temporary distortion of some syllables and silencing of others. In some instances, the interval between syllables adjacent to the silenced syllable was shortened. The only permanent effect of bead implants was the permanent deletion (even after bead removal) of omitted syllables for which this silent "gap" had been closed. This observation is also consistent with a key finding of Tchernichovski et al. (2001), who studied song acquisition in juvenile birds exposed to a highly efficient tutoring regimen: when a syllable at a given position in the tutor's song was similar to a syllable at a different position in the juvenile's song, the juvenile would change a syllable in the same position rather than move the similar syllable to the appropriate position.

Thus, it may be that the temporal structure of song is less plastic under most circumstances than individual syllable phonology. This would also explain why, leaving aside changes resulting from the loss of individual syllables, temporal changes observed as a result of WN exposure in the current work were minimal and rarely resulted in the reordering of song elements. Furthermore, even disrupted songs maintained a high degree of linearity and consistency, suggesting that they were sung in a stereotyped sequence despite gross spectral disruptions.

Data from the delayed auditory feedback study of Leonardo and Konishi (1999) suggest that there are circumstances under which the temporal structure of song may be more plastic than spectral structure. Their procedure left normal auditory feedback intact, such that birds heard a superposition of normal and introduced sounds. This procedure produced substantial disruption of the temporal organization of song, although birds continued to produce a low incidence of their normal songs even when overall song disruption was maximal. After removal of delayed auditory feedback, birds in the study by Leonardo and Konishi (1999) gradually recovered their previous song patterns.

The differential plasticity of syllable phonology and higher-order temporal structure may arise because these two parameters of song are associated with anatomically distinct neural substrates. One possibility is that high vocal center (HVC) codes primarily for higher-order temporal structure, whereas robust nucleus of the archistriatum (RA) codes for syllable phonology (Vu et al., 1994; Yu and Margoliash, 1999; Hahnloser et al., 2002), and these nuclei lie on different pathways with respect to the MAN–basal ganglia circuit that supports song learning. Whereas RA receives direct input from lMAN, HVC receives input from medial magnocellular nucleus of the anterior neostriatum (mMAN). Foster and Bottjer (2001) demonstrated that lesions to mMAN result in selective disruption of temporal aspects of song, particularly at the beginnings of song bouts.

In contrast to Bengalese finches, zebra finches do not learn new song elements

Woolley and Rubel (2002) found that some syllables in the songs of Bengalese finches reversibly deafened with ototoxic hair cell

lesions gradually evolved to resemble those of birds with which they were housed during recovery, whereas we found no evidence for learning from live or taped tutors in adult zebra finches. Bengalese finches sing more complex songs than zebra finches (they repeat each syllable a variable number of times from rendition to rendition, whereas zebra finches typically sing each syllable once per motif) and have much higher rates of neurogenesis (at least after deafening) (Scott et al., 2000). Bengalese finch song may therefore be more plastic than zebra finch song, as evidenced by the relatively short period of deafening required to induce song deficits in Bengalese finches (cf. Nordeen and Nordeen, 1992, with Woolley and Rubel, 1997). Of course, we must be cautious in concluding that the difference between our results and those of Woolley and Rubel (2002) are attributable solely to species differences in plasticity. Whereas we exposed our subjects to WN, their subjects were deafened by cochlear lesion, which could promote injury-induced plasticity in auditory neural circuitry.

References

Böhner J (1990) Early acquisition of song in the zebra finch, *Taeniopygia guttata*. Anim Behav 39:369–374.

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404:762–766.

Brainard MS, Doupe AJ (2001) Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. J Neurosci 21:2501–2517.

Dave AS, Margoliash D (2001) Song replay during sleep and computational rules for sensorimotor vocal learning. Science 290:812–816.

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631.

Fee MS, Shraiman B, Pesaran B, Mitra PP (1998) The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. Nature 395:67–71.

Foster E, Bottjer SW (2001) Lesions of a telencephalic nucleus in male zebra finches: influences on vocal behavior in juveniles and adults. J Neurobiol 46:142–165.

Hahnloser RH, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419:65–70.

Hough GE, Volman SF (2002) Short-term and long-term effects of vocal distortion on song maintenance in zebra finches. J Neurosci 22:1177–1186.

Iyengar S, Bottjer SW (2002) The role of auditory experience in the formation of neural circuits underlying vocal learning in zebra finches. J Neurosci 22:946–958.

Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. Nature 399:466–570.

Lombardino A, Nottebohm F (2000) Age at deafening affects the stability of learned song in adult male zebra finches. J Neurosci 20:5054–5064.

Marler P, Peters S (1982) Developmental overproduction and selective attrition: new processes in the epigenesis of birdsong. Dev Psychobiol 15:369–378.

Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. Behav Neural Biol 57:58–66.

Nottebohm F, Nottebohm ME (1978) Relationship between song repertoire and age in the canary, *Serinus canarius*. Z Tierpsychol 46:298–305.

Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (*Lonchura striata* var.domestica) require real-time auditory feedback to produce normal song syntax. J Neurobiol 33:343–356.

Ryals BM, Dooling RJ, Westbrook E, Dent ML, MacKenzie A, Larsen ON (1999) Avian species differences in susceptibility to noise exposure. Hear Res 131:71–88.

Scharff CS, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 11:2896–2913.

Scott LL, Nordeen EJ, Nordeen KW (2000) The relationship between raters of HVc neuron addition and vocal plasticity in adult songbirds. J Neurobiol 43:79–88.

Solis MM, Doupe AJ (1997) Contributions of tutor and bird's own song experience to neural selectivity in the songbird anterior forebrain. J Neurosci 19:4559–4584.

Tchernichovski O, Nottebohm F, Ho CE, Bijan PE, Mitra PP (2000) A procedure for an automated measurement of song similarity. Anim Behav 59:1167–1176.

Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291:2564–2569.

Troyer TW, Bottjer SW (2002) Birdsong: models and mechanisms. Curr Opin Neurobiol 11:721–726.

Troyer TW, Doupe AJ (2001a) An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. J Neurophysiol 84:1204–1223.

Troyer TW, Doupe AJ (2001b) An associational model of birdsong sensori-

motor learning II. Temporal hierarchies and the learning of song sequence. J Neurophysiol 84:1224–1239.

Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. J Neurosci 14:6924–6934.

Williams H, Mehta N (1997) Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. J Neurobiol 39:14–28.

Woolley SM, Rubel EW (1997) Bengalese finches Lonchura striata domestica depend upon auditory feedback for the maintenance of adult song. J Neurosci 17:6380–6390.

Woolley SM, Rubel EW (2002) Vocal memory and learning in adult Bengalese finches with regenerated hair cells. J Neurosci 22:7774–7787.

Woolley SM, Wissman AM, Rubel EW (2001) Hair cell regeneration and recovery of auditory thresholds following aminoglycoside ototoxicity in Bengalese finches. Hear Res 153:181–195.

Yu A, Margoliash D (1999) Temporal hierarchical control of singing in birds. Science 273:1871–1875.

Zann RA (1996) The Zebra finch: a synthesis of field and laboratory studies. London: Oxford UP.

Zevin JD, Seidenberg MS, Bottjer SW (2000) Song plasticity in adult zebra finches exposed to white noise. Soc Neurosci Abstr 26:723.