Adult birdsong is actively maintained by error correction

Samuel J Sober & Michael S Brainard

Humans learn to speak by a process of vocal imitation that requires the availability of auditory feedback. Similarly, young birds rely on auditory feedback when learning to imitate the songs of adult birds, providing one of the few examples of nonhuman vocal learning. However, although humans continue to use auditory feedback to correct vocal errors in adulthood, the mechanisms underlying the stability of adult birdsong are unknown. We found that, similar to human speech, adult birdsong is maintained by error correction. We perturbed the pitch (fundamental frequency) of auditory feedback in adult Bengalese finches using custom-designed headphones. Birds compensated for the imposed auditory error by adjusting the pitch of song. When the perturbation was removed, pitch returned to baseline. Our results indicate that adult birds correct vocal errors by comparing auditory feedback to a sensory target and suggest that lifelong error correction is a general principle of learned vocal behavior.

Song learning in many species of birds, similar to language learning in humans, is based on a process of imitation[1,2]. Learning begins when a young bird is exposed to the song of an adult 'tutor'. Subsequently, the young bird refines his initially disordered vocalizations into a mature, 'crystallized' song that is very similar to that of the tutor[3–5]. This process parallels speech acquisition in humans, in which language exposure in infancy is thought to establish perceptual categories that then serve as targets for vocal production[2,6], eventually resulting in a child learning to produce the phonemes of his or her native language. For both adult speech and crystallized song, phonetic structure is highly stereotyped from one rendition to the next and is extremely stable throughout the remainder of the individual's life.

Despite the widely accepted parallels between the acquisition of song and speech[7], it is unclear whether the extraordinary stability of adult vocal behavior in birds and humans results from similarly parallel processes. Speech performance in adult humans is thought to rely on an active process in which sensory signals are used to identify and correct vocal motor errors. Recent laboratory studies have shown that when auditory feedback is manipulated, human adults alter their vocal output so as to reduce the experienced auditory error[8,9], demonstrating a reliance on auditory feedback to correct speech errors.

Here, we tested the hypothesis that, similar to humans, birds maintain adult vocal output by comparing auditory feedback to a long-lived auditory target and using the resulting error signals to adaptively modify their vocalizations. This type of error-corrective mechanism requires several things of the adult song system. First, there must be a stable auditory target representing the desired song output. Second, adults must detect small differences between auditory feedback and the target and evaluate which changes in motor output reduce the sensory error. Third, the motor program for song must be modifiable.

The first two requirements define error correction, whereas the third merely asserts that the song system is plastic in adulthood.

Manipulations that corrupt or completely eliminate auditory experience (via distortions of auditory feedback or deafening) have been shown to drive degradations of crystallized song[10–12], demonstrating the potential for vocal plasticity in adulthood and showing that the third condition is satisfied. Furthermore, differential reinforcement signals provided by an external evaluator can drive directed changes in adult song[13]. To our knowledge, however, no prior studies have shown that adult birds naturally perform error correction by monitoring song output to detect and correct deviations from an auditory target. Deafening and feedback distortion procedures presumably create a mismatch between auditory feedback and the sensory target, but provide no opportunity for the bird to correct these errors, as no alteration of vocal output can restore normal auditory feedback. Reinforcement procedures circumvent error detection entirely by rewarding or punishing birds on the basis of the experimenter's (rather than the bird's) evaluation of vocal performance. A true test of adult error correction therefore requires an experimental procedure in which birds both detect song errors and modify their vocal output to reduce them.

We tested the hypothesis of adult error correction by introducing small, correctable perturbations to the pitch (fundamental frequency) of auditory feedback to mimic naturally occurring vocal errors. Pitch is a learned and precisely controlled parameter of song and the pitch of individual song elements (or syllables) is refined during song acquisition and is extremely stable in adulthood[5,14]. We predicted that if adult birds maintain their songs by comparing auditory feedback to an auditory target and modifying their songs to correct sensory errors, then shifting the pitch of auditory feedback would cause birds to change the pitch of song in the direction opposite the experimentally imposed feedback shift, thus reducing the experienced auditory error.
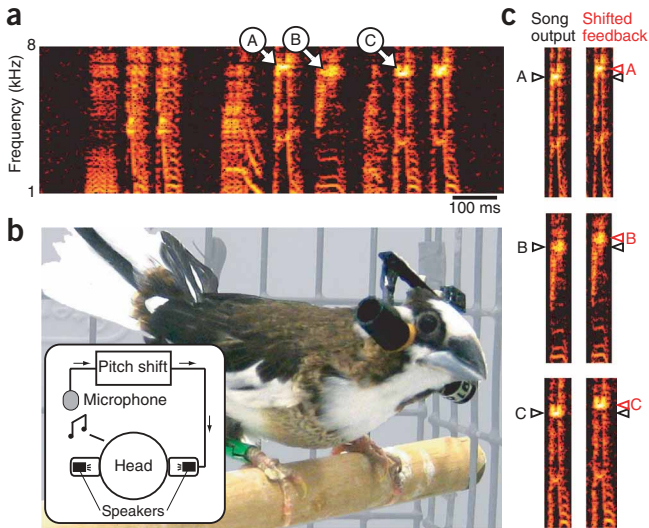
Figure 1 Technique for manipulating auditory feedback. (**a**) Crystallized song from an adult Bengalese finch. Spectrographic representation shows the power at each frequency (color scale) as a function of time. Three harmonic features are labeled A, B and C. (**b**) Each bird was fit with a set of headphones that housed a pair of speakers. A microphone in the cage (see inset) provided input to online sound-processing hardware, which was used to manipulate the pitch of song. Processed acoustic signals were then relayed to the headphone speakers via a flexible cable (not shown in photograph) and played through the speakers. (**c**) An upward (+100 cents) shift in the pitch of auditory feedback was introduced by the headphone system. For each of the harmonic features labeled in **a**, the left spectrogram shows the bird's acoustic output and the black triangle shows the frequency of the harmonic feature. The right spectrogram shows the pitch-shifted auditory feedback played through the headphones and the red triangle shows the frequency of the harmonic feature in the shifted song. Black triangles are repeated next to the spectrograms on the right for comparison.

each syllable, we quantified changes in pitch by measuring changes in either the fundamental frequency or the frequency of one of the higher harmonics (a harmonic feature, see Online Methods). In a typical

RESULTS

A set of lightweight headphones was custom-fit to each bird in the study to generate online shifts in the pitch of auditory feedback (an example of crystallized song from one bird in our study is shown in **Fig. 1a**). A microphone in each bird's cage relayed acoustic signals through sound-processing hardware capable of generating arbitrary shifts in pitch. These pitch-shifted acoustic signals were then played back through speakers in the headphones (**Fig. 1b**) with an average processing delay of ~7 ms. Shifts in the pitch of auditory feedback and the resulting changes in the pitch of song are both measured in units of 'cents' (see Online Methods), where 1,200 cents corresponds to an octave and 100 cents represents the same pitch interval as a semitone (approximately a 6% change in absolute frequency). A 100-cent upward shift in pitch applied to several song syllables is shown in **Figure 1c**.

We consistently found that shifting the pitch of auditory feedback led to adaptive changes in song (that is, changes in the direction opposite the imposed pitch shift). We monitored changes in vocal output by repeatedly measuring pitch at particular times (or spectral frames; **Fig. 2a**) during song syllables. In the spectral frame chosen for
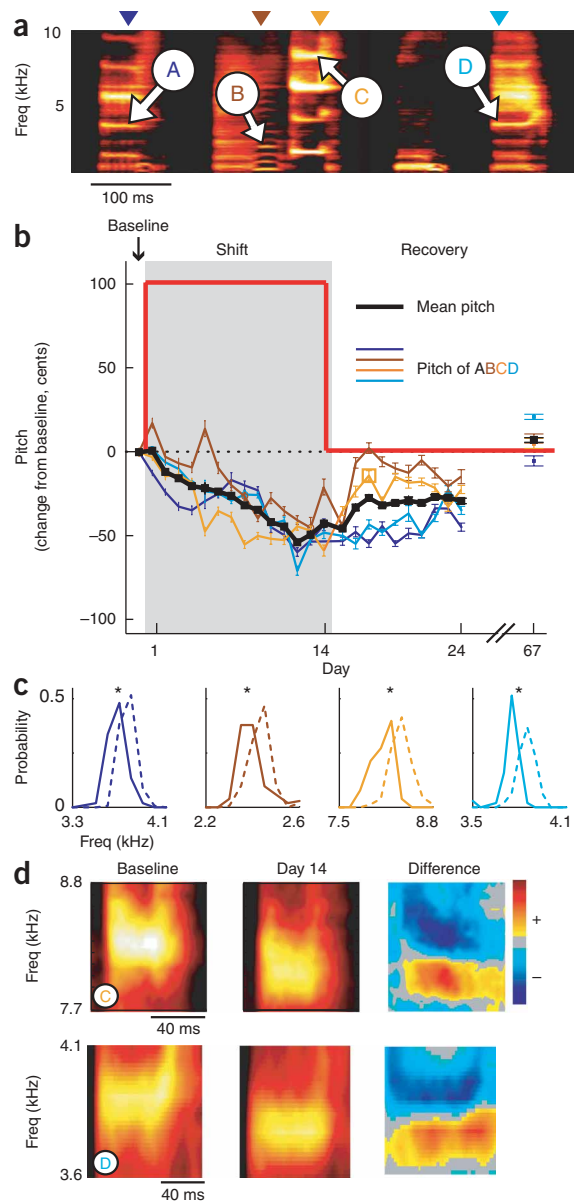

Figure 2 Vocal error correction driven by an upward shift in the pitch of auditory feedback. (**a**) Baseline song of Bird 1 (mean spectrogram). Arrowheads above the spectrogram indicate the spectral frames (measurement times in each syllable) at which four harmonic features (A–D) were measured to quantify changes in the pitch of song. (**b**) Changes in pitch in response to a 100-cent upward shift (red line) in the pitch of auditory feedback and subsequent recovery back to baseline. Colored lines show the mean ± s.e.m. change in pitch (measured in cents, see Online Methods) of each harmonic feature across time and the black line shows the mean change in the pitch of song (mean ± s.e.m. pitch change averaged across harmonic features). After 14 d of shift exposure (gray box), unshifted auditory feedback was restored and the bird was monitored for an additional 10 d. Pitch was also measured on day 67 to assess any long-term changes. (**c**) Pre- and post-shift distributions of the frequencies of the harmonic features shown in **a**. For each feature, the probability distribution of frequencies during baseline (dashed lines) and day 14 (solid lines) differed significantly (* $P < 10^{-5}$, one-tailed t test). Color conventions for each feature as in **b**. (**d**) Pitch shift-induced changes in mean spectral structure. Left, mean spectrograms for harmonic features C (top) and D (bottom) during the baseline epoch. Middle, mean spectrograms for features C and D on shift day 14. Right, difference spectrograms obtained by subtracting the baseline spectrograms from the day 14 spectrograms.
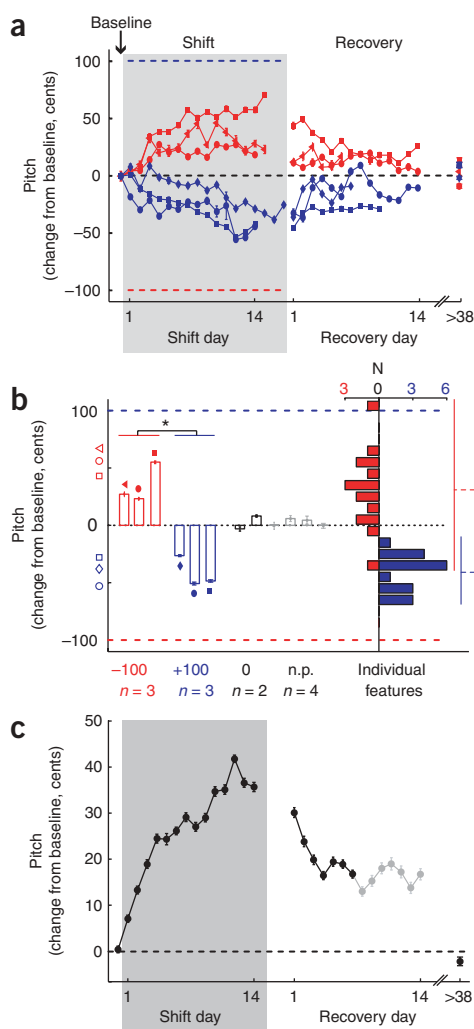
Figure 3 Error correction in response to upward and downward shifts in feedback pitch. (**a**) Mean ± s.e.m. changes in the pitch of song (across harmonic features) as a function of time for three experiments with downward shifts in feedback pitch (−100 cents, red) and three experiments with upward shifts (+100 cents, blue). Recovery day 1 is the first day after unshifted feedback was restored. Symbol shapes identify individual birds. Other conventions are as described in **Figure 2b**. (**b**) Mean ± s.e.m. change in song pitch during shift days 12–14 for downward (empty red bars) and upward (empty blue bars) shifts of auditory feedback. Symbols identify individual birds as in **a**. Empty red and blue symbols at left indicate the magnitude of 1 s.d. of pitch variation in the baseline epoch of each experiment (averaged across harmonic features). Black bars show changes in song pitch across the same interval in birds that wore headphones but did not experience pitch shifts (0 shift), and gray bars show the same measure in a group of birds that did not wear headphones at all (no phones). Longitudinal data from the two 0 shift birds is available as **Supplementary Figure 6** online. Filled bars at right show the distribution of pitch changes for individual harmonic features, combined across all downward (red) and upward (blue) shifts of auditory feedback. Asterisks indicate significant differences between the effects of upward and downward shifts ($P < 0.05$, one-tailed t tests). In addition, the changes in song pitch in each ±100-cent shift group were significantly different from the changes in both of the control groups ($P < 0.05$ in all cases, one-tailed t tests). (**c**) Combined data from all experiments, reoriented so that changes in the adaptive directions are positive. Days from which data were available from all six experiments are plotted in black; recovery days from which data were available from a subset of four experiments are plotted in gray.

and downward shifts ($P = 0.62$, two-tailed t test) (**Fig. 3b**). We observed minimal changes in pitch in two control groups consisting of birds that wore headphones but did not experience a pitch shift and birds that were never fitted with headphones (**Fig. 3b**). In addition, shifting the pitch of auditory feedback did not significantly affect the amount of song produced ($P = 0.54$), the ordering (syntax) of song syllables ($P = 0.46$) or the relative spectral power found in low versus high harmonics ($P = 0.16$; see **Supplementary Figs. 1–3** online).

Analysis of pitch changes in individual harmonic features revealed consistent adaptive learning and heterogeneity in the songs of individual birds. In the majority (31 of 33) of cases, the pitches of individual features changed in the adaptive direction ($P < 10^{-9}$, one-tailed t-test; **Fig. 3b**). The magnitude of pitch change (assessed on shift day 14) varied significantly across the harmonic features in each bird's song in all six experiments (ANOVA, $P < 10^{-9}$ in all cases). This heterogeneity suggests that the pitch of individual syllables can be adapted separately even if the pitch of auditory feedback is shifted uniformly across song. No correlation was observed between the baseline fundamental frequency of syllables and the magnitude of adaptive learning (**Supplementary Fig. 4** online).

The time course and magnitude of adaptation to pitch shifts suggest that birds compromise between the correction of auditory errors (adaptability) and the maintenance of the established motor program (stability). We combined the data from all shift experiments to determine the average magnitude and time course of learning (**Fig. 3c**). Our results suggest that birds balance adaptability and stability by changing their songs relatively slowly in response to auditory errors. This time scale is comparable to the rate of pitch refinement in juvenile birds that are first learning their songs[14]. In addition, the acoustic structure of adult Bengalese finch song syllables begins to degrade in 1 week following deafening[12,15], possibly as a result of error signals generated in the absence of auditory feedback[16]. These results suggest that the intrinsic dynamics of the adult song system only allow for slow modifications of the motor command, perhaps to prevent the destabilization of song during brief periods of aberrant sensory input.

experiment, a 100-cent upward shift in the pitch of auditory feedback led to a downward change in the pitch of the measured harmonic features over a 2-week period (**Fig. 2b**). We used the mean pitch change across all harmonic features to measure the daily mean change in the pitch of song. This change in the bird's song served to reduce the difference (error) between the pitch of auditory feedback and the baseline pitch. By the end of the shift epoch, the pitch of song had dropped by ∼50 cents. On removal of the pitch shift, song returned slowly toward baseline pitch, with a complete return to baseline when the song was assessed 67 d after the beginning of the experiment. We examined the distribution of frequencies for each harmonic feature in the baseline song and after 14 d of exposure to the upward pitch shift (**Fig. 2c**). All four harmonic features had significantly lower frequencies following shift exposure ($P < 10^{-5}$, one-tailed t test). The gross structure of song was preserved despite there being significant changes in pitch (**Fig. 2d**).

Across experiments, the direction of the behavioral response consistently opposed the direction of the imposed shift (**Fig. 3a**). Both upward and downward shifts in the pitch of auditory feedback led to adaptive changes in the pitch of song and song returned to baseline pitch following shift offset. Across six pitch-shift experiments, changes in response to upward versus downward shifts were significantly different ($P = 0.002$, one-tailed t test) and we found no significant difference between the magnitudes of adaptive responses to upward

Our combined data (**Fig. 3c**) also illustrate that on average birds compensate for only ~40% of the imposed pitch shift by the end of the shift epoch. Incomplete adaptation is also consistently observed in human subjects in response to perturbations of auditory feedback[8,9] and is the norm in other sensorimotor systems[18]. Here, the incompleteness of adaptation may reflect a partial reliance on nonauditory sources of information, such as proprioceptive signals[19,20] or the output of an internal (forward) model used to predict the sensory consequences of motor commands[21,22]. Furthermore, although the auditory feedback played through the headphone speakers was substantially louder than auditory feedback that reached the bird's ears directly (that is, leaked through the headphone frame, see Online Methods), some unshifted feedback signals probably reached the bird via bone conduction[9]. Therefore, the observed incomplete adaptation might also have resulted from a conflict between two distinct auditory signals.

The adaptive changes in the pitch of song that we found fall well within the range of both the baseline variability in pitch and the upper limits of pitch plasticity. Comparison of the mean change in song to the baseline variability showed that birds changed the pitch of song by roughly 1 s.d. (**Fig. 3b**). This comparison reveals that the mean pitch of these syllables could have shifted even further in the adaptive direction without leaving the range of baseline variation, which has a s.d. of 48.0 cents (averaged across birds). In addition, the observed magnitude of pitch adaptation does not reflect a physical limitation of the peripheral motor system, as we found in a separate set of experiments that gradually increasing the size of a shift can drive 100-cent (or larger) changes in the pitch of song[17] (**Supplementary Fig. 5** online) and prior studies using reinforcement techniques have demonstrated that birds can shift the pitch of individual syllables by more than 10 s.d[13].

DISCUSSION

Our results show that adult birds use auditory feedback to correct small errors in the pitch of song, indicating that adult song is maintained by a continual process of error correction. Maintaining vocal output through error correction has several important advantages. First, it allows the bird to adapt to changes in vocal output resulting from age-related changes in the strength of vocal muscles and the physical properties of the vocal periphery. Such changes alter the relationship between premotor activity and acoustic output, necessitating changes in the motor command. Second, error correction allows for changes in the neural structures controlling song as synaptic connectivity changes, old neurons die and new neurons are added in adulthood[23–25]. A stable sensory representation of the desired song, therefore, provides for stable behavioral output as both the body and brain change over a bird's lifetime.

As we found for songbirds, human adults gradually adapt vocal (speech) production to compensate for shifts in the pitch of auditory feedback[9] and have been shown to compensate for more complex distortions of the structure of individual phonemes[8]. Taken together, these results suggest that the use of long-lived perceptual targets to correct auditory errors is a general principle of learned vocal behavior. Notably, vocal adaptation in humans is substantially more rapid than that described here, suggesting that the sensorimotor strategy used to maintain speech performance puts relatively greater emphasis on adaptability (and less on stability) than that used to maintain song. Alternately, this difference in time scale might result from experimental differences between our study and prior studies on human speech, which have used bone oscillators or whispered speech to minimize unshifted feedback signals transmitted to the ear via bone conduction[8,9].

Our results provide direct evidence that adult birds maintain a sensory target representing the desired song output. Such a representation is presumably formed when juvenile birds are exposed to the tutor's song[1,3], and it is possible that this target persists unchanged into adulthood and drives the adaptive pitch changes that we observed. Alternately, auditory feedback might not be compared to a memory of the tutor's song, but rather to a slowly updated memory of the bird's own (relatively stable) vocalizations. Although our experiments do not allow us to distinguish between these possibilities, the adaptive changes in pitch that we observed and the subsequent return to baseline weeks after pitch shifts were removed demonstrate that the sensory target is unchanged even after weeks of abnormal sensory experience.

Our results also demonstrate that birds can reduce auditory errors by shifting the pitch of song away from a normally stable baseline (**Fig. 3b**). Previous studies have shown that song sometimes reverts to its original state following recovery from partial deafening or the cessation of distortion experiments[10,26]. However, recent lesion studies suggest that recovery in these settings may be based on the persistence of the baseline motor program to which aberrant neural signals have been added by basal ganglia circuitry[27]. Furthermore, recovery following deafening or feedback distortion might rely on a somatosensory or motor memory of the baseline song[1,20] rather than on a process of auditory error correction. A demonstration of adult error correction therefore requires that birds modify their crystallized song to reduce experienced auditory errors, as we observed here.

The songbird system provides a unique opportunity to study the brain mechanisms underlying the learning and control of vocal behavior. After exposure to the tutor's song, juvenile birds can practice and refine their songs in complete social isolation and without any external source of reinforcement[14,28]. On the basis of this and other findings (such as the inability of young birds to acquire normal song if they are deafened after tutor song exposure[1]), it is widely assumed that vocal refinement in young birds reflects a process of error correction that is similar to the one we found here in the adult[29,30]. Numerous studies have documented anatomical and molecular changes that accompany song acquisition and therefore distinguish between the brains of juvenile and adult birds[31,32]. Our results demonstrate that the capacity for adaptive vocal plasticity persists far into adulthood and suggest that the error-correcting processes that are thought to underlie vocal plasticity during song acquisition are active in the adult. If juvenile song acquisition and adult song maintenance are indeed based on the same process of error correction, a common neural mechanism for sensory-guided plasticity may persist throughout the marked neural changes that occur at the close of the critical period for song acquisition.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureneuroscience/.

Note: Supplementary information is available on the Nature Neuroscience website.

AUTHOR CONTRIBUTIONS
S.J.S. and M.S.B. designed the experiments and wrote the paper. S.J.S. designed and built the experimental apparatus, conducted the experiments and analyzed the data.

1. Konishi, M. The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* **22**, 770–783 (1965).
2. Kuhl, P.K. Learning and representation in speech and language. *Curr. Opin. Neurobiol.* **4**, 812–822 (1994).
3. Arnold, A.P. The effects of castration on song development in zebra finches (*Poephila guttata*). *J. Exp. Zool.* **191**, 261–278 (1975).
4. Olveczky, B.P., Andalman, A.S. & Fee, M.S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol.* **3**, e153 (2005).
5. Kao, M.H., Doupe, A.J. & Brainard, M.S. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* **433**, 638–643 (2005).
6. Kuhl, P.K. *et al.* Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Phil. Trans. R. Soc. Lond. B* **363**, 979–1000 (2008).
7. Doupe, A.J. & Kuhl, P.K. Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631 (1999).
8. Houde, J.F. & Jordan, M.I. Sensorimotor adaptation in speech production. *Science* **279**, 1213–1216 (1998).
9. Jones, J.A. & Munhall, K.G. Perceptual calibration of F0 production: evidence from feedback perturbation. *J. Acoust. Soc. Am.* **108**, 1246–1251 (2000).
10. Leonardo, A. & Konishi, M. Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* **399**, 466–470 (1999).
11. Nordeen, K.W. & Nordeen, E.J. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. Neural Biol.* **57**, 58–66 (1992).
12. Woolley, S.M. & Rubel, E.W. Bengalese finches *Lonchura striata domestica* depend upon auditory feedback for the maintenance of adult song. *J. Neurosci.* **17**, 6380–6390 (1997).
13. Tumer, E.C. & Brainard, M.S. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* **450**, 1240–1244 (2007).
14. Tchernichovski, O., Mitra, P.P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
15. Okanoya, K. & Yamaguchi, A. Adult Bengalese finches (*Lonchura striata var. domestica*) require real-time auditory feedback to produce normal song syntax. *J. Neurobiol.* **33**, 343–356 (1997).
16. Brainard, M.S. & Doupe, A.J. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762–766 (2000).
17. Linkenhoker, B.A. & Knudsen, E.I. Incremental training increases the plasticity of the auditory space map in adult barn owls. *Nature* **419**, 293–296 (2002).
18. Choe, C.S. & Welch, R.B. Variables affecting the intermanual transfer and decay of prism adaptation. *J. Exp. Psychol.* **102**, 1076–1084 (1974).
19. Sober, S.J. & Sabes, P.N. Flexible strategies for sensory integration during motor planning. *Nat. Neurosci.* **8**, 490–497 (2005).
20. Suthers, R.A., Goller, F. & Wild, J.M. Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proc. Natl. Acad. Sci. USA* **99**, 5680–5685 (2002).
21. Wolpert, D.M., Ghahramani, Z. & Jordan, M.I. An internal model for sensorimotor integration. *Science* **269**, 1880–1882 (1995).
22. Troyer, T.W. & Doupe, A.J. An associational model of birdsong sensorimotor learning. I. Efference copy and the learning of song syllables. *J. Neurophysiol.* **84**, 1204–1223 (2000).
23. Nottebohm, F. The road we travelled: discovery, choreography, and significance of brain replaceable neurons. *Ann. NY Acad. Sci.* **1016**, 628–658 (2004).
24. Wilbrecht, L. & Kirn, J.R. Neuron addition and loss in the song system: regulation and function. *Ann. NY Acad. Sci.* **1016**, 659–683 (2004).
25. Mooney, R. Synaptic mechanisms for auditory-vocal integration and the correction of vocal errors. *Ann. NY Acad. Sci.* **1016**, 476–494 (2004).
26. Woolley, S.M. & Rubel, E.W. Vocal memory and learning in adult Bengalese finches with regenerated hair cells. *J. Neurosci.* **22**, 7774–7787 (2002).
27. Thompson, J.A., Wu, W., Bertram, R. & Johnson, F. Auditory-dependent vocal recovery in adult male zebra finches is facilitated by lesion of a forebrain pathway that includes the basal ganglia. *J. Neurosci.* **27**, 12308–12320 (2007).
28. Tchernichovski, O., Lints, T., Mitra, P.P. & Nottebohm, F. Vocal imitation in zebra finches is inversely related to model abundance. *Proc. Natl. Acad. Sci. USA* **96**, 12901–12904 (1999).
29. Adret, P. In search of the song template. *Ann. NY Acad. Sci.* **1016**, 303–324 (2004).
30. Derégnaucourt, S. *et al.* Song development: in search of the error-signal. *Ann. NY Acad. Sci.* **1016**, 364–376 (2004).
31. Clayton, D.F. Songbird genomics: methods, mechanisms, opportunities, and pitfalls. *Ann. NY Acad. Sci.* **1016**, 45–60 (2004).
32. White, S.A. Learning to communicate. *Curr. Opin. Neurobiol.* **11**, 510–520 (2001).

ONLINE METHODS

Subjects. Adult (>190 d old) male Bengalese finches (*Lonchura striata* var. *domestica*) were used. All procedures were approved by the University of California, San Francisco Institutional Animal Care and Use Committee. Birds were individually housed in sound-isolation chambers throughout the experiments and were maintained on a 14-h:10-h light/dark cycle, with lights on from 7 a.m. to 9 p.m.

Manipulating auditory feedback. Headphones that were constructed from lightweight carbon fiber (Hobby Lobby) were custom-fit to each bird and held a miniature speaker (EH-7157-000, Knowles) within 3 mm of the entrance to each ear canal. In addition, a miniature microphone (EM-3046, Knowles) was placed between ear and speaker on one side to calibrate and monitor the performance of the pitch-shifting hardware. The amplitude of the acoustic signal played through the speakers was ~2 log units greater than direct auditory feedback leaking through the carbon fiber frame. A condenser microphone inside each isolation chamber relayed all cage sounds (including, but not limited to, song) through sound-processing hardware (Ultraharmonizer DSP7000, Eventide) that introduced upward, downward or null shifts in pitch. Acoustic signals were then played back through the headphone speakers with an average loop delay of ~7 ms. All recordings are from undirected song (that is, no female was present).

Quantifying behavioral changes. Bengalese finch song consists of distinct syllables that can be identified across multiple renditions of song. For each syllable that had a well-defined pitch (3–8 in each bird), we identified an 8- or 16-ms segment of time (or spectral frame) to use for our pitch measurements. In rare cases where a complex syllable was composed of two spectrally distinct notes, we measured spectral frames for each distinct note (for example, **Fig. 2a**). The beginning of each spectral frame was defined relative to the onset of the syllable. In each spectral frame, we measured pitch by quantifying the frequency of a harmonic feature (either the fundamental frequency or one of the higher harmonics). The frequency of this harmonic feature was then repeatedly measured over the course of each experiment and changes in its frequency were analyzed as described below.

After being fitted with headphones, birds sang for 3–7 d with zero shift. Each day we analyzed all of the songs produced in a 2-h window between 10 a.m. and 12 p.m. (we obtained similar results when song produced in the evening was used for analysis, see **Supplementary Fig. 7** online). The baseline pitch of each harmonic feature was defined as the mean pitch in the last two analysis days preceding shift onset. After the last baseline window, a ±100-cent pitch shift was introduced in a single step and maintained for 14–17 d, after which the pitch shift was reset to zero for 7–14 d. Headphones were subsequently removed and birds were maintained in social isolation for 25–123 d. Headphones were then reattached (with zero pitch shift) to assess the long-term effects on pitch. The pitch of each iteration of a harmonic feature was quantified in units of cents

$$c_x = 1200 \log_2 \frac{h_x}{b}$$

where c_x is the pitch (in cents) of the feature, h_x is the pitch (in Hz) of the feature and b is the baseline pitch (in Hz) of that feature. Note that because the frequency of harmonics are related to each other by integer multiples, when the pitch of a syllable changes, the frequency of every harmonic changes by the same number of cents. Therefore, the computed change in cents describes the change in the fundamental frequency. Also note that this method is equivalent to measuring the frequency of a given harmonic, dividing that frequency by the appropriate integer to compute the fundamental frequency and computing the change in cents as described above.

Mean spectrograms. Mean spectrograms were calculated for each syllable by first computing the spectrogram for each iteration of the syllable being analyzed. Spectrograms were then normalized and aligned to a single prototypical example of the syllable. Alignment was accomplished by shifting and linearly stretching each spectrogram until the cross-correlation between the temporal power profile (total spectral power as a function of time) of the exemplar and the prototype was maximized. After all syllables had been aligned, the mean power at each frequency and time point was computed. Note that mean spectrograms are used for display purposes only (**Fig. 2**) and that mean changes in pitch were computed from measurements of individual syllables.