Supporting Information

Andalman and Fee 10.1073/pnas.0903214106

SI Methods

Surgery. Animals were anesthetized with 1.5–2% isoflurane in oxygen and placed in a stereotaxic apparatus (MyNeuroLabs) with a head angle of 20° (anterior skull relative to horizontal). LMAN was localized stereotaxically relative to the bifurcation of the sagittal sinus (5.2-mm anterior, 1.75-mm lateral). Untethered microdialysis probes were implanted bilaterally into LMAN (Fig. 2 *A* and *S*3) such that the center of the microdialysis tube was located at a depth of 1.82 mm from the brain surface, and secured in place with dental acrylic (Flow-It ALC, Pentron Clinical Technologies). The speaker tube for distorted auditory feedback (MicroRenathane Implantation Tubing, Braintree Scientific, 0.025-inch O.D., 4-mm length) was surgically inserted through a small hole in the skull into the airsac surrounding the cerebellum (Fig. S1), and sealed into place with dental acrylic. This airsac is continuous with the interior surface of both eardrums. The sound levels in the speaker/airsac system were calibrated by temporarily sealing a small probe microphone (ER-7C, Etymotic Research) into the cranial airsac on the opposite hemisphere. A full transfer function was computed between the speaker drive current and airsac sound level as a function of frequency. After surgery the animal was allowed to recover for several days. The total weight of the implant was 1.6 g.

Conditional Auditory Feedback. In our conditional feedback protocol, the playback of disruptive noise was a function of a measure of syllable pitch. Specifically, the power of the feedback played to the bird as a function of time, $P_F(t)$, was determined by a scalar function, $R(t)$, of the spectral properties of the song as follows: $P_F(t) = \alpha P_S(t)[R(t) - \lambda]_+$, where $P_S(t)$ is the amount of power in the song vocalization at time t, α is a multiplicative factor that sets the loudness of the feedback relative to the song power, and λ is a threshold value (typically set to 0.85) of the scalar $R(t)$. The brackets $[]_+$ indicate that negative values of the argument were set to zero, so that feedback is only played when the scalar function, $R(t)$, is greater than the threshold λ. $P_S(t)$ was calculated as a low-pass filtered version of the squared audio signal. $P_F(t)$ was computed on-line using a digital signal processor (RX8, Tucker Davis Technologies). The loudness factor α was chosen as the minimum value that gives reliable learning during the conditioned feedback protocol, and was typically set at ≈10.

$R(t)$ was constructed to give a large value when the pitch was within a specified range. This was achieved by using a bank of 6 bandpass filters arranged to detect the harmonic peaks in the squared audio signal of the targeted harmonic stack (Fig. S2*A*)

$$R(t) = \frac{F_1(t) + F_3(t) + F_5(t)}{\sum_{i=1}^{6} F_i(t)}$$

where the function $F_i(t)$ is the power measured in the squared song signal by ith bandpass filter. The filter center-frequencies were spaced in an array. The filter shape (equi-ripple 300–450Hz full width) was chosen so that the feedback power fell off rapidly outside of a band of syllable pitches. In this way, a pitch threshold θ_F was defined, on one side of which the feedback power was large (greater than the birds own song) and on the other side of which the feedback power was small (Fig. 1*B* and Fig. S2*C*). This threshold pitch was set each morning before lights-on and remained constant for the rest of the day. The pitch threshold was set based on the last 50 non-TTX syllables of the previous day. Specifically, if the previous day was a TTX day, the threshold was set to average pitch of the last 50 syllables before inactivation. If the previous day was a vehicle day, the threshold was set to the average pitch of the last 50 syllables at the end of the day, before lights-out.

The value of $P_F(t)$ was used to modulate the amplitude of a broadband noise signal shaped to have the same spectral profile as the average spectrum of the experimental bird's song (Fig. S2*B*). Note that the feedback power was always scaled by the instantaneous song power, so that the bird could not reduce the relative feedback power by simply singing louder. All of the above filters were implemented as FIR filters on the digital signal processor. The filter widths were designed to be narrow in time so as to provide a very short delay (<4 ms) between the singing and feedback signal. The loudness scale α was set such that the feedback was between 10 and 20dB louder than the target syllable.

Transient LMAN Inactivation. After recovery from surgery, birds were connected to the electrical commutator and PBS was infused into the drug reservoir every afternoon until the animals sang consistently after infusion. Drug infusion was carried out using the following procedure: ≈4 h after the onset of singing, the bird was placed in a small foam restraint that left the head free and the microdialysis probe drug reservoir accessible. The caps were removed from the drug reservoirs. For infusion, the contents of the reservoir (vehicle) were extracted with the corner of a Kimwipe. A 1 mL of syringe containing the solution to be loaded (TTX or vehicle) was placed into the reservoir (Luer fitting). The solution was injected into the reservoir until ≈50 μL of solution had flowed out through the outlet tube and the reservoir was filled. The syringe was removed and the caps were replaced. For washout, the contents of the reservoir (TTX or vehicle) were extracted with a Kimwipe, as described above, and 100 μL of PBS was injected through the outlet tube. The reservoir was then emptied and filled again to ensure all TTX was rinsed from the reservoir and dialysis tubing. The reservoirs were left full of PBS, and the caps were replaced.

Typically, singing started within 1 h of lights-on. The average number of syllables sung before drug infusion was 831 (5th, 50th, and 95th percentile were 191, 691, and 1,818, respectively). The amount of time between TTX infusion and singing averaged 78 ± 72 min (SD), which was not significantly different from vehicle infusion (62 ± 48 min, SD, $P = 0.25$). On most TTX days, birds began to sing within an hour of infusion (n = 19/34 infusions). The number of syllables sung during TTX infusion was not significantly different from during vehicle infusion (TTX: mean 263 ± 271 syllables, SD; vehicle: mean = 312 ± 329 syllables, $P = 0.46$). Reversals of pitch contingency (i.e., changing from moving up to moving down) were made every 4–8 days and were performed before lights-on on vehicle days.

To assess the extent to which pitch variability was reduced by TTX infusion, we calculated a reduction in variability based on the last 50 renditions of the target syllable before infusion and the first 50 renditions after infusion. For each set of 50 renditions, the mean time course of pitch was computed. This mean time course was subtracted from the individual pitch time courses within the set—resulting in a residual time course for each of the 50 syllables. Pitch variability was defined as the average of the standard deviations of these 50 residuals. Consistent with the fact that the standard deviation is calculated as variations around the mean value, residuals shown in Fig. S4*A* and *B* were demeaned (i.e., we subtract from each residual the mean pitch of that residual). The reduction in variability was calculated as the ratio of the postinfusion pitch variability to the preinfusion pitch variability. These results were plotted as a function of TTX concentration (Fig. S4*C*). At a dose of 25 μM, TTX infusion produced a saturating reduction in pitch variability, even for syllables produced within 30 min of infusion.

The spatial extent of inactivation produced by TTX infusion into LMAN was determined by electrophysiological recordings in anesthetized birds. First, we confirmed that medial MAN (MMAN) was not inactivated in our experiments. A microdialysis probe was stereotaxically placed in LMAN (1.75-mm LM), and an electrode (Carbostar, 1 MΩ, Kation Scientific) was placed near the lateral edge of MMAN (0.6-mm and 0.7-mm LM, $n = 2$ birds) to record the robust spontaneous bursting activity present throughout the nidopallium. Activity persisted in MMAN for the duration of the recording session, at least 1–1.5 h after infusion of 25 μM TTX into the probe. The electrode was then moved laterally to record in steps of 0.2 mm. Complete loss of spontaneous activity was confirmed at lateral positions >1.0 mm (within 0.75 mm of the probe).

Data Analysis. The time courses of pitch and feedback power during the day (Fig. 1 *E* and *F*) were computed as follows. The mean pitch and feedback power of every rendition of the target syllable was calculated for each day of training ($n = 80$ days, $n = 5$ birds). For each day, a smoothed trajectory of pitch and feedback power was calculated by averaging over blocks of 40 renditions (overlapping blocks, sliding by 10 renditions). Each smoothed trajectory was interpolated in 15 min steps. The initial morning value of pitch was subtracted from the pitch trajectory for each day, and trajectories for down days were inverted. For the feedback power trajectory, the trajectory was normalized by the initial morning value of feedback power. The trajectories for all days were averaged together to produce the plots shown in Fig. 1 *E* and *F*. These plots include learning on TTX days, but only include singing before TTX infusion. The average rate of learning during the first 4 h of the day (quoted in the results section) was computed from Fig. 1*E* as the average pitch change at the 4-h point, divided by 4 h.

The distributions of daytime pitch changes include data from TTX and vehicle days (Fig. 1*H*). On vehicle days, the pitch change was between the first morning songs and the last evening song before lights-out. On TTX days, the pitch change was between the first morning songs and the last songs before TTX infusion. Overnight changes in pitch were quantified as the difference between the morning pitch and that of the final renditions of the preceding day. The overnight analysis did not include nights following TTX infusion ($n = 43$ days).

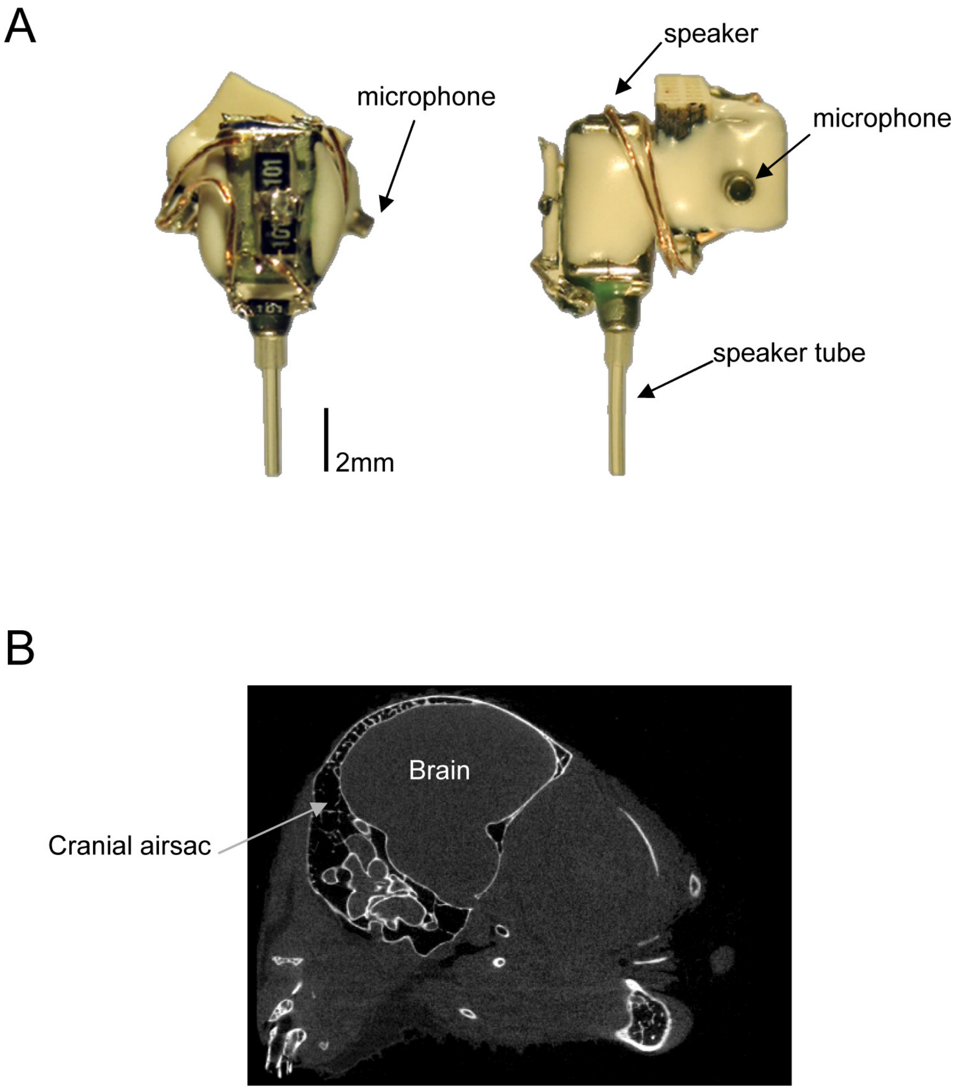

Fig. S1. Disruptive auditory feedback system. (*A*) Photographs of speaker implant. A hearing-aid speaker generated broadband noise to distort the auditory feedback of the bird during singing. The sound was delivered through a speaker tube into the cranial airsac. Song was recorded using a miniature microphone attached to the side of the implant. (*B*) The cranial airsac as seen in a parasagittal CT-section through the middle ear. This airsac is continuous with the inner surface of both eardrums.

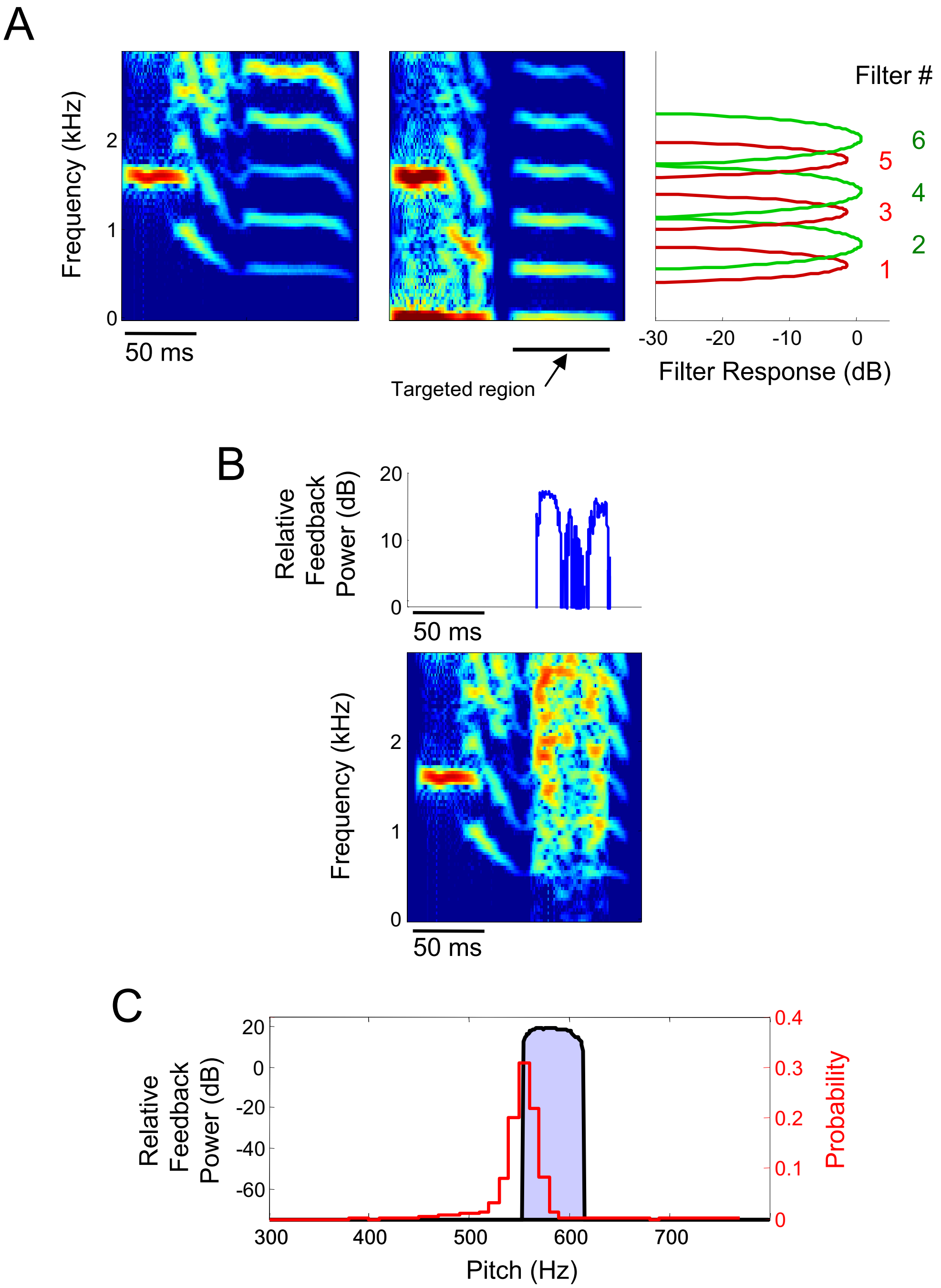

Fig. S2. Method for generating disruptive auditory feedback conditional on pitch. (*A*) Spectrogram of 1 song syllable (*Left*) and the associated spectrogram of the squared audio signal (*Middle*). The latter is a better measure for pitch because it is less affected by timbre (i.e., missing harmonics). A particular pitch range is targeted for feedback using a bank of 6 bandpass filters—three arranged to detect in-band power and three arranged to detect out-of-band power (*Right*; red, in-band; green, out-band). (*B*) The feedback power (relative to song power) is proportional to the sum of the 3 in-band filters, normalized by the sum of all 6 filters (*Upper*). Also shown is the distorted auditory feedback signal superimposed over the syllable spectrogram (with correct amplitude scaling, *Lower*). (*C*) Relative feedback power as a function of pitch (black trace) for the filter settings used in *A* and *B*. Also shown is the distribution of pitches (red trace) for the targeted region of the syllable shown above. Note that the edge of the pitch detection region (shaded) is set approximately in the middle of the distribution of pitches of the targeted syllable.

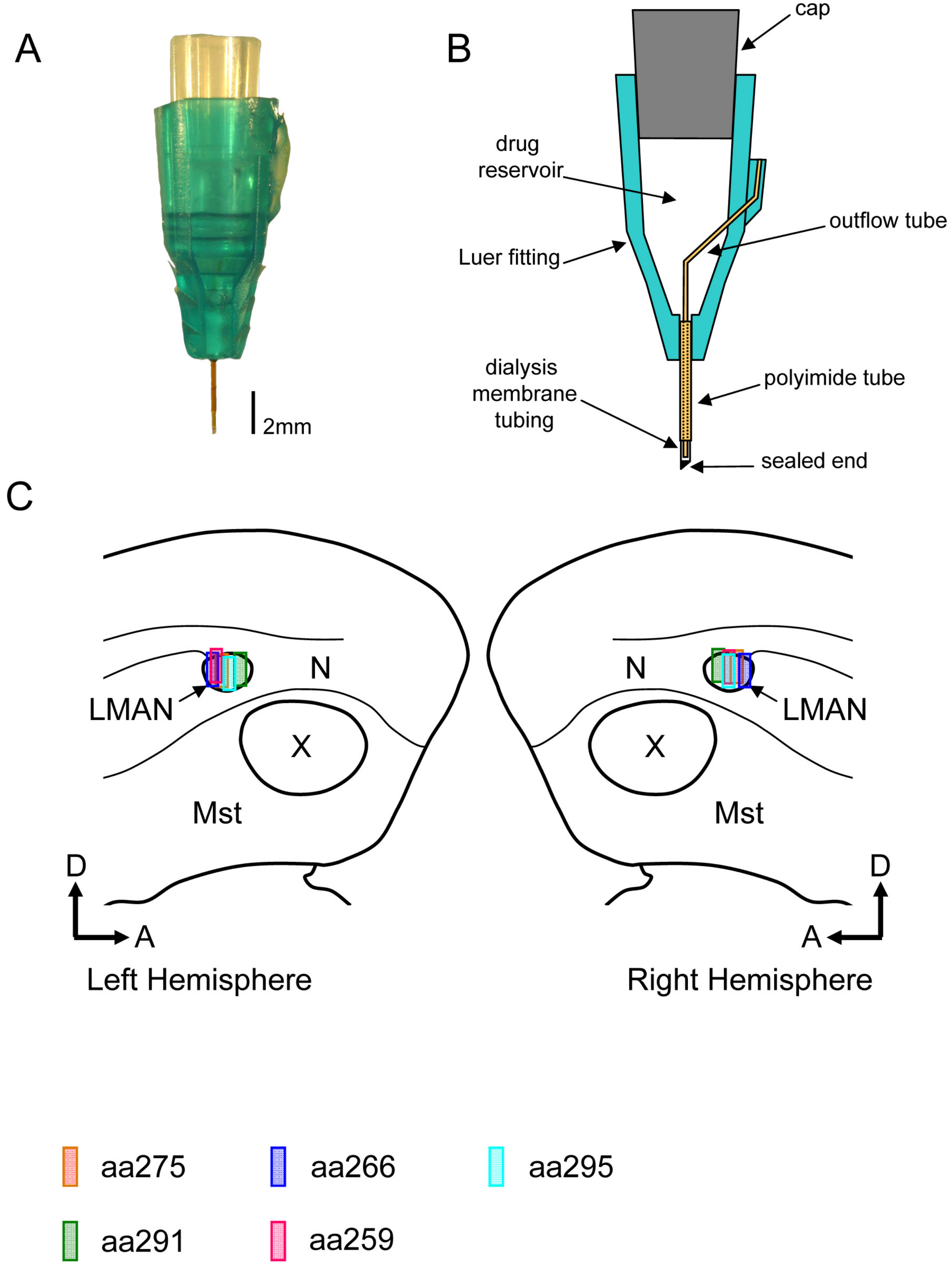

Fig. S3. Untethered reverse microdialysis for inactivation of LMAN. (*A*) Photograph of microdialysis probe. (*B*) Schematic drawing of microdialysis probe showing drug reservoir, dialysis tubing, and outflow tube. The outflow tube is inserted concentrically to the bottom of the sealed end of the dialysis tubing, allowing fluid to be rapidly flowed in and washed out of the dialysis tubing. (*C*) Histologically confirmed location of bilateral microdialysis probes in LMAN in all 5 experimental birds in the parasagital plane. All probes were also within 300 μm of the center of LMAN along the medial-lateral axis. Bird numbers are listed at the bottom. X, Area X; N, nidopallium; Mst, medial striatum; LMAN, lateral magnocellular nucleus of the nidopallium; D, dorsal; A, anterior.

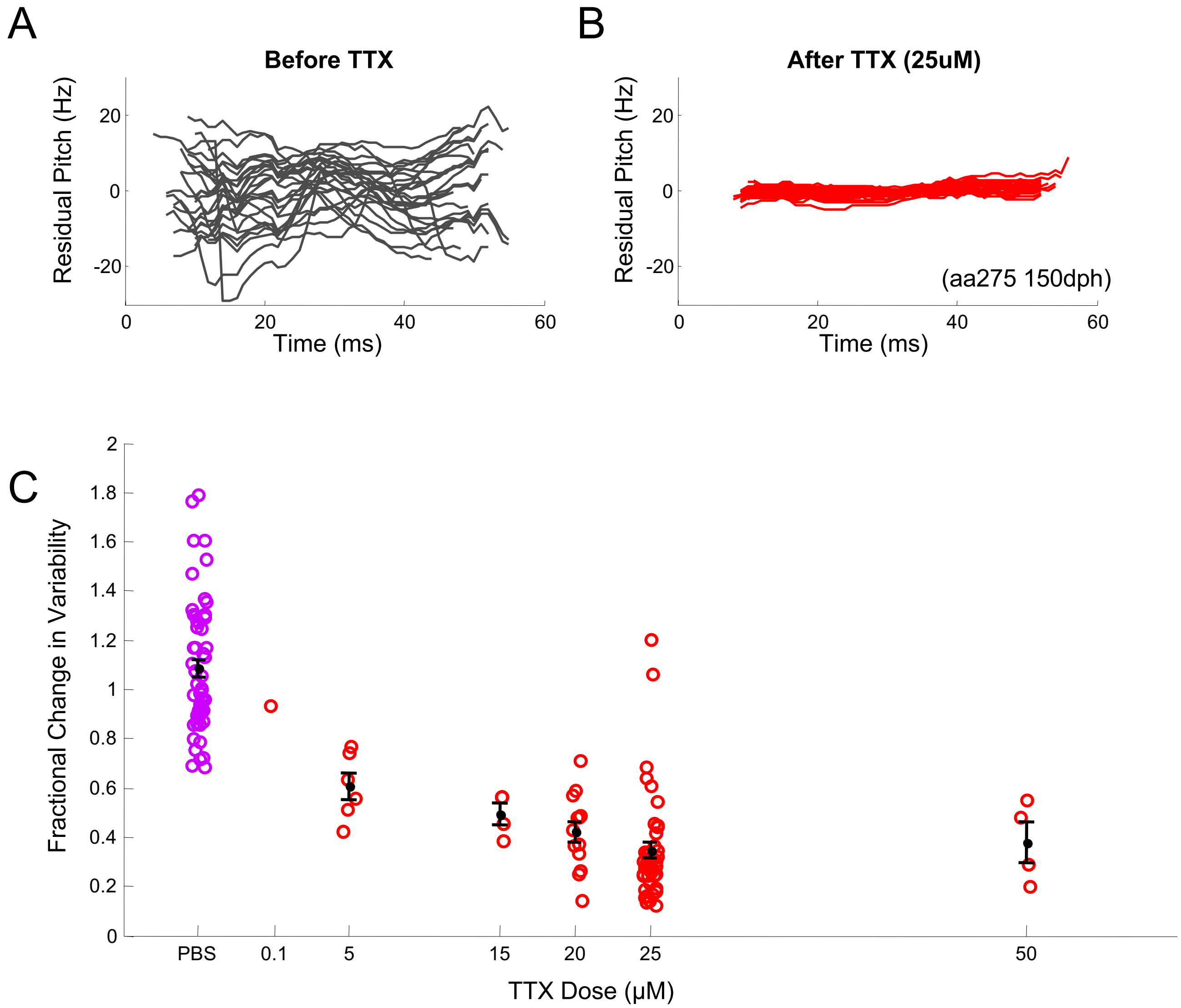

Fig. S4. Effect of tetrodotoxin (TTX) in LMAN on variability of pitch. (*A*) Time course of residual pitch for the last 25 renditions of the targeted syllable before infusion (bird aa275, 150 dph). (*B*) Time course of residual pitch for same targeted syllable after infusion of 25 μM TTX into reverse microdialysis probe (first 25 renditions postinfusion). Pitch variability after infusion was 18% of the preinfusion variability (see *SI Methods*). (*C*) Dose–response curve for reduction in pitch variability for 6 concentrations of TTX, compared to vehicle infusion (PBS). A concentration of 25 μM produced a saturating reduction in pitch variability, and did not affect the rate of singing. Black symbols indicate the mean and standard error for each concentration.

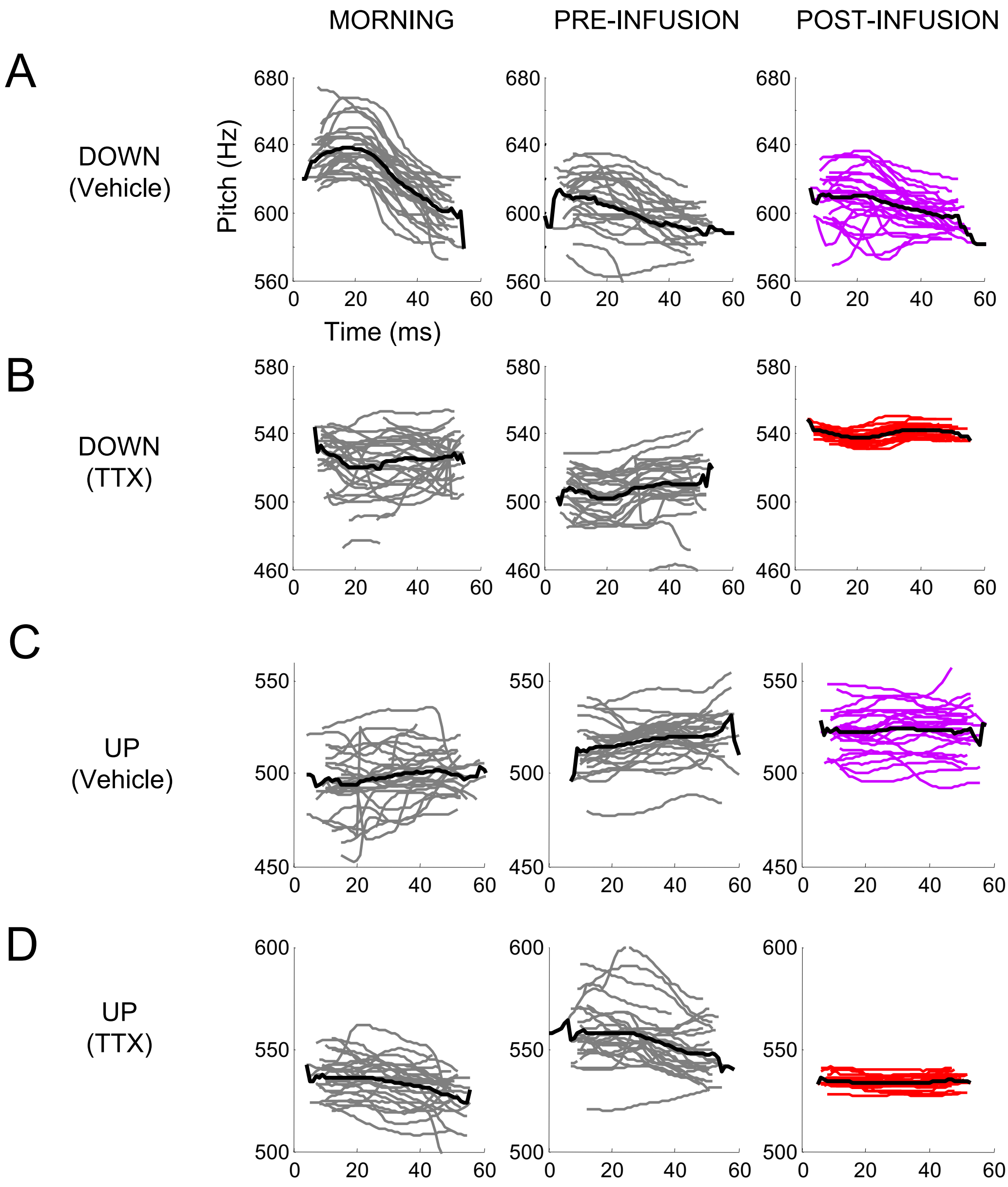

Fig. S5. Examples of pitch time courses on 4 experimental days. (*A*–*D*) Each curve reflects the time course of pitch within 1 rendition of the targeted syllable. Shown are the first 25 renditions of the day (morning, left column), followed by 25 syllable renditions after several hours of learning under conditional feedback (immediately before drug infusion, middle column). Note that during the time between awaking (morning) and manipulation (preinfusion), the distribution of pitches has shifted down for days on which the pitch is being pushed down (DOWN) and has shifted up for days on which the pitch was pushed up (UP). Also shown are pitch traces for the first 25 syllable renditions after infusion (postinfusion, right column; red lines, TTX, *B* and *D*; purple lines, vehicle, *A* and *C*). Note the dramatic reduction of pitch variability during TTX inactivation, and the regression of average pitch back toward the morning value. Mean pitch time course shown in black. *A* and *B* show the pitch time courses from the experimental days shown in Figs. 2 *B* and *C* (bird aa275, 149 dph and 152 dph). *C* and *D* show 2 additional experimental days (bird a275, 141 dph and 142 dph).

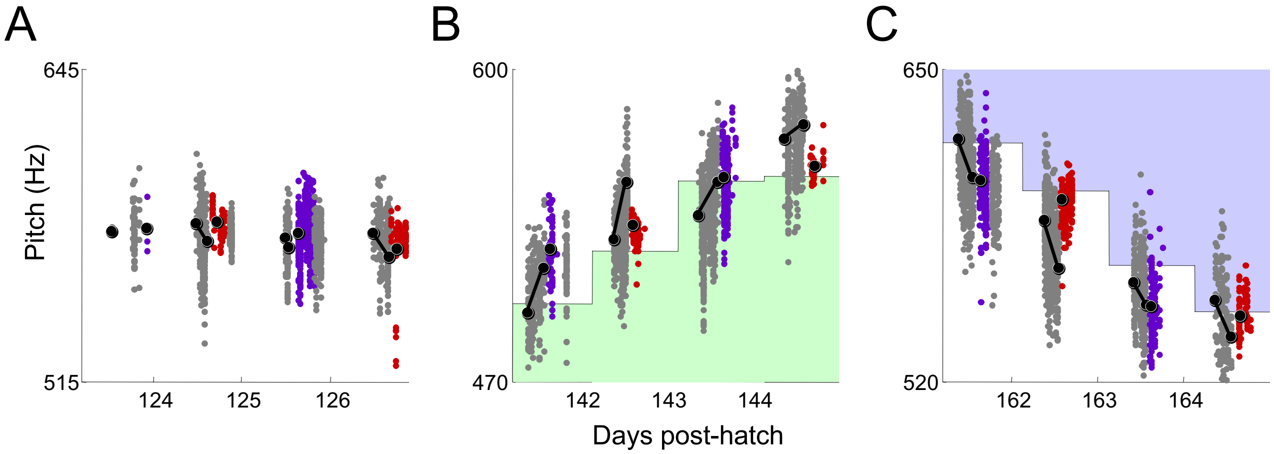

Fig. S6. Sequential experimental days. (*A*) Four sequential days of drug infusion during the baseline period, before application of conditional feedback. TTX (red) and vehicle (purple) infused on alternate days. Black dots indicate average pitch in the morning, preinfusion, and postinfusion periods. (*B* and *C*) Four sequential days of drug infusion during experimental days when conditional feedback was used to move pitch up (*B*) and down (*C*).

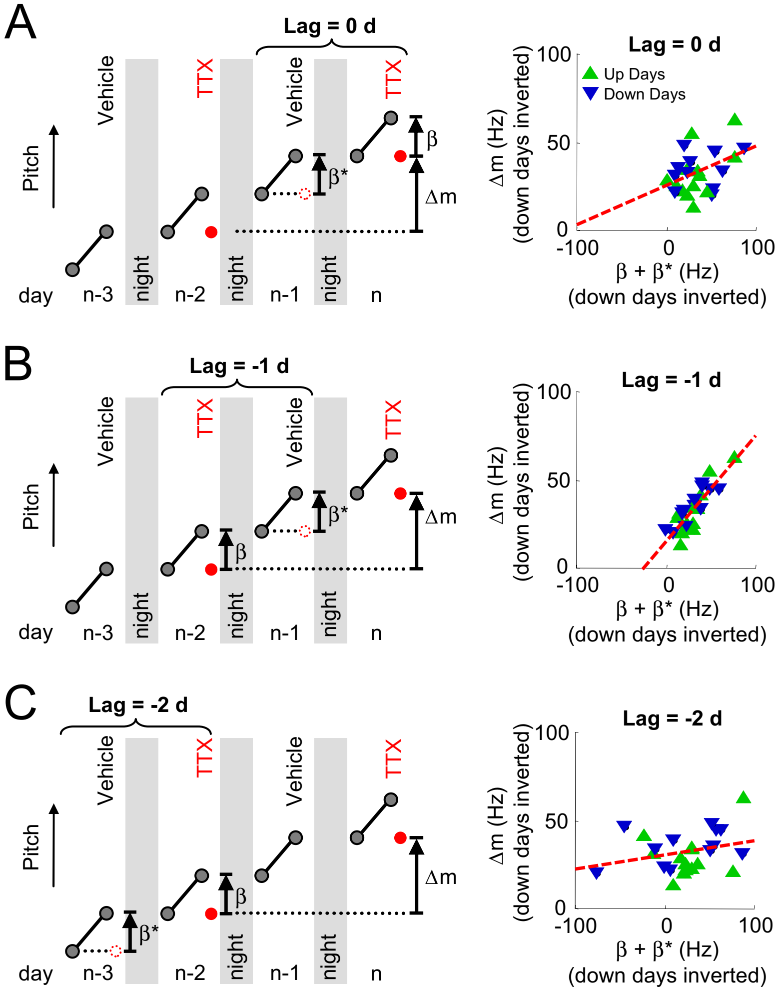

Fig. S7. Detailed illustration of the how the correlation between variations in motor pathway plasticity and estimated AFP bias was computed at different lags for Fig. 5. (*Left*) Schematics showing the quantities used to calculate the change in LMAN(−) pitch (Δm) and the estimated AFP bias ($\beta + \beta^*$) at lag $= 0$ (*A*), lag $= -1$ (*B*), and lag $= -2$ (*C*). Note the quantities used to estimate AFP bias shift in time according the lag value. (*Right*) Scatter plots of Δm versus $\beta + \beta^*$ at the corresponding lag.