

Sensory and Motor Systems

The Effects of Pitch Shifts on Delay-Induced Changes in Vocal Sequencing in a Songbird

MacKenzie Wyatt, Emily A. Berthiaume, Conor W. Kelly, and Samuel J. Sober

DOI: <http://dx.doi.org/10.1523/ENEURO.0254-16.2017>

Department of Biology, Emory University, Atlanta, GA 30322

Abstract

Like human speech, vocal behavior in songbirds depends critically on auditory feedback. In both humans and songbirds, vocal skills are acquired by a process of imitation whereby current vocal production is compared to an acoustic target. Similarly, performance in adulthood relies strongly on auditory feedback, and online manipulations of auditory signals can dramatically alter acoustic production even after vocalizations have been well learned. Artificially delaying auditory feedback can disrupt both speech and birdsong, and internal delays in auditory feedback have been hypothesized as a cause of vocal dysfluency in persons who stutter. Furthermore, in both song and speech, online shifts of the pitch (fundamental frequency) of auditory feedback lead to compensatory changes in vocal pitch for small perturbations, but larger pitch shifts produce smaller changes in vocal output. Intriguingly, large pitch shifts can partially restore normal speech in some dysfluent speakers, suggesting that the effects of auditory feedback delays might be ameliorated by online pitch manipulations. Although birdsong provides a promising model system for understanding speech production, the interactions between sensory feedback delays and pitch shifts have not yet been assessed in songbirds. To investigate this, we asked whether the addition of a pitch shift modulates delay-induced changes in Bengalese finch song, hypothesizing that pitch shifts would reduce the effects of feedback delays. Compared with the effects of delays alone, combined delays and pitch shifts resulted in a significant reduction in behavioral changes in one type of sequencing (branch points) but not another (distribution of repeated syllables).

Key words: Bengalese finch; birdsong; sensorimotor learning; stuttering

Significance Statement

Vocal behavior depends critically on an organism's ability to monitor the sound of its own voice ("auditory feedback"). Studies of both humans and songbirds have demonstrated that successful vocal performance depends critically on the quality and timing of such feedback; however, the interaction between vocal acoustics and the timing of auditory feedback is unclear. Here we used songbirds to examine this interaction by measuring vocal performance during delays and distortions (pitch shifts) of auditory feedback.

Introduction

Learned vocal behaviors depend strongly on auditory feedback. In both birdsong and human speech, adults rely

on auditory feedback to detect and correct errors in vocal production. This reliance on auditory information can be demonstrated by manipulating auditory feedback and

Received August 28, 2016; accepted January 1, 2017; First published January 11, 2017.

Authors report no conflict of interest.

Author Contributions: M.W., E.A.B., and S.J.S. designed research; M.W., E.A.B., and C.W.K. performed research; M.W., E.A.B., and S.J.S. analyzed data and wrote the paper.

National Institutes of Health R01 NS084844 (S.J.S.), National Science Foundation Grant 1456912 (S.J.S.), Woodruff Foundation (M.W.).

M. Wyatt's present address: University of Kentucky College of Medicine, UK Medical Center MN 150, 800 Rose St., Lexington, KY 40506; Emily Berthiaume's present address: David Geffen School of Medicine at UCLA, 10833 Le Conte Ave., Los Angeles, CA 90095; Conor Kelly's present address: Medical College of Georgia, 1459 Laney Walker Blvd., Augusta, GA 30912.

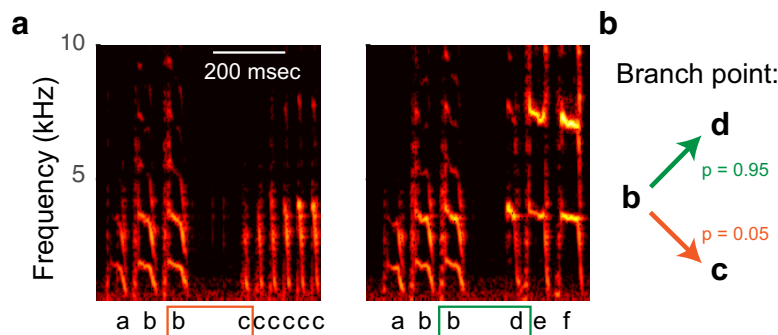


Figure 1. Sequence variation in birdsong. **A**, Spectrographic representations show the power (heat map) at each acoustic frequency (vertical axis) as a function of time (horizontal axis). The two spectrograms show excerpts from different times during the same bout of song from a single bird. Labels below the spectrogram indicate different syllables. Orange and green boxes highlight a branch point in which syllable b can be followed by either syllable d or syllable c. **B**, Schematic quantifies transition probabilities for this branch point.

measuring the effects on vocal output. Complete elimination of auditory feedback by deafening in adulthood leads to dramatic vocal performance deficits (McGarr, 1983; Okanoya and Yamaguchi, 1997; Woolley and Rubel, 1997; Lombardino and Nottebohm, 2000). More subtle manipulations of auditory signals reveal the complex influence of sensory feedback on motor programming. Artificially delaying auditory feedback in human speakers can cause vocal sequencing errors, including unwanted repetitions of consonants and words, in normally fluent speakers (Fairbanks, 1955; Chase, 1958; Yates, 1963). Such results suggest that the sequencing errors observed in persons who stutter might result from disorders of auditory feedback processing (Büchel and Sommer, 2004; Hampton and Weber-Fox, 2008). Intriguingly, artificially delaying auditory feedback is sometimes effective as a treatment for stuttering (Ryan and Van Kirk, 1974; Kalinowski and Stuart, 1996), further linking the dependence of vocal sequencing on the timing of auditory feedback and emphasizing the complex relationship between sensory feedback and speech production. Analogously, studies of birdsong have shown that perturbations of auditory feedback timing can degrade vocal production. Delayed playbacks of a bird's own syllable during singing leads to song degradation after chronic exposure in zebra finches (Leonardo and Konishi, 1999; Cynx and von Rad, 2001). In Bengalese finches, a species whose song contains “branch points” where vocal sequencing is probabilistic rather than fixed (Fig. 1), acute changes in vocal sequencing can result from delayed playbacks of a bird's own song syllable while singing (Sakata and Brainard, 2006). The similarities of these results across species

suggest songbirds as a promising animal model for disorders of human speech production.

Other studies in songbirds and humans have explored how the brain uses the acoustic structure of auditory feedback (as distinct from the timing of feedback) to calibrate vocal performance. In both songbirds and humans, manipulations of the fundamental frequency (which we refer to here as pitch) of auditory feedback evoke compensatory responses, for example by increasing the pitch of vocal output in response to a decrease in the pitch of online auditory feedback (Jones and Munhall, 2000; Sober and Brainard, 2012; Hoffmann and Sober, 2014). Notably, vocal pitch changes in birdsong and formant changes in human speech are most robust for smaller shifts in auditory feedback, with larger shifts evoking little or no change in vocal output (Burnett et al., 1998; Liu and Larson, 2007; MacDonald et al., 2010; Katseff et al., 2012; Sober and Brainard, 2012), suggesting that during large pitch shifts, the brain relies less on auditory feedback to influence ongoing vocal behavior. Intriguingly, large (half an octave) pitch shifts cause an increase in speech fluency in some persons who stutter (Kalinowski et al. 1993; Natke et al., 2001), suggesting an interplay between the acoustic structure of auditory feedback and the sequencing of vocal motor commands. However, the relationship between the acoustics and timing of vocal feedback, and their influence on vocal output, remain poorly understood. By examining the interactions between the acoustic structure and timing of auditory feedback in songbirds, we hope to better understand the interaction between sensory and vocal motor signals. Such an understanding will position the songbird as an animal model in which neural activity during normal and disordered vocal production might be examined, and the neural mechanisms by which auditory feedback delays and pitch shifts can improve vocal performance in persons who stutter might be understood.

We used Bengalese finches to investigate whether delay-induced changes in vocal production were influenced by alterations in the pitch of auditory feedback. Although the effects of auditory feedback delays on song have been tested previously using song-triggered play-

Acknowledgments: The authors thank Lukas Hoffmann, David Nicholson, and Kyle Srivastava for helpful conversations.

Correspondence should be addressed to Samuel J. Sober, Emory University Department of Biology, 1510 Clifton Rd. NE, Rm 2006, Atlanta, GA 30322. E-mail: samuel.j.sober@emory.edu.

DOI: <http://dx.doi.org/10.1523/ENEURO.0254-16.2017>

Copyright © 2017 Wyatt et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

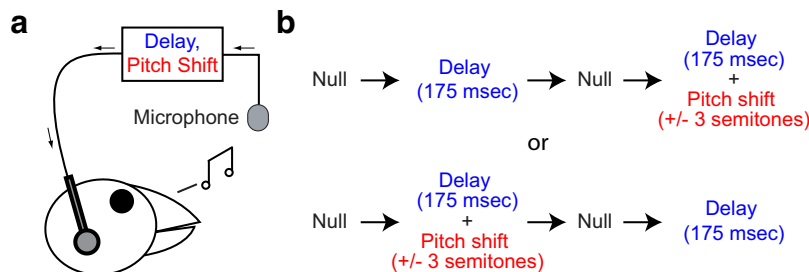


Figure 2. Experimental design. **A**, Auditory feedback manipulated using miniaturized headphones. A microphone transmits a bird's vocalizations to online sound processing hardware, which is used to introduce a 175-ms delay (with or without a ± 3 semitone pitch shift). **B**, Schedule of experimental conditions. Each bird was exposed to the delayed feedback alone (DAF) and delayed feedback plus a pitch shift (DAF+PS). The order of these conditions, as well as the direction of the pitch shift, was randomized across subjects. Before each DAF or DAF+PS epoch, birds sang in a null epoch free of delays or pitch shifts. The four experimental epochs lasted 5 d each.

backs of previously recorded songs, technical challenges have prevented the use of continuous delayed feedback. We overcame this obstacle using miniaturized headphones (see Methods), which provided continuous, delayed feedback in real time (Hoffmann et al., 2012). Experimental conditions included a null condition (no delay or pitch manipulation), delayed auditory feedback (DAF) without any pitch shift, and a condition in which auditory feedback was both delayed and pitch shifted (DAF+PS). We predicted that, in agreement with prior findings using playbacks of short segments of song (Sakata and Brainard, 2006), delayed feedback would induce changes in the syllable transition probabilities. In particular, we hypothesized that, as reported by Sakata and Brainard (2006), delayed feedback would result in the most common transition (the "primary" transition) becoming less prevalent and the nonprimary transition becoming more common. We further hypothesized that the large pitch shift in the DAF+PS condition would reduce the magnitude of the changes observed in the DAF condition.

Materials and Methods

Four adult (>100 d old) male Bengalese finches (*Lonchura striata* var. *domestica*) obtained from an outside vendor were used as the experimental subjects. During experiments, birds were housed individually in an isolated sound-attenuating chamber, and all song was undirected (i.e., produced in the absence of female birds). The light/dark cycle was maintained for 14:10 h, with lights on beginning at 7 am and ending at 9 pm. All procedures were approved by the Emory University Institutional Animal Care and Use Committee.

Experimental procedure

Miniature, lightweight headphones were custom-built out of lightweight carbon fiber and custom-fitted to each bird's head (Hoffmann et al., 2012). A condenser microphone in the bird's cage (Fig. 2A) captured the bird's acoustic output, which was routed to online sound-processing hardware (Eventide H7600), which provided perturbations of the pitch and/or timing of auditory feedback in real time. This manipulated feedback was then relayed to miniaturized speakers (EH-7157-000, Knowles) inside the headphones. In addition to the speakers, the

headphones apparatus included a miniaturized microphone (EM-3046, Knowles) placed between the speaker and the opening of one of the ear canals. This microphone allowed us to monitor the performance of the headphones apparatus. The miniaturized microphone was used to calibrate the system such that the acoustic signal played through the headphones speakers was ~ 2 log units greater than auditory feedback leaking through the carbon fiber frame. The headphones therefore shielded the birds' airborne vocalizations, allowing the altered feedback to replace the natural version. As described in detail below, auditory feedback conditions included a null condition in which no manipulation was introduced, a DAF condition in which auditory feedback was delayed by 175 ms, or a DAF+PS condition in which both a 175-ms delay and an upward or downward pitch shift were applied simultaneously. As detailed previously, the sound-processing hardware relayed the online acoustic signal to the headphones with a minimal delay (i.e., when auditory feedback was not being intentionally delayed) of ~ 10 ms, a delay which does not in itself evoke any measurable changes in vocal behavior (Sober and Brainard, 2009; Hoffmann et al., 2012; Kelly and Sober, 2014). Note that subjects might receive some unmanipulated auditory feedback via bone conduction, although such a scenario is very unlikely to account for our results (see Discussion).

The experimental design is outlined in Fig. 2B. Once birds habituated to the headphones, all birds sang during a period with zero pitch shift or delay for 5 d. After this null period, in the example shown at top in Fig. 2B, the birds' auditory output was altered with delayed auditory feedback (DAF block) for 5 d. The birds heard their natural vocalizations at 175-ms delay relative to output. We selected this delay magnitude after preliminary studies in two birds (not used in the present study) suggested that this delay consistently evoked changes in vocal sequence; however, different delay values were not tested systematically. After the altered auditory feedback block, birds were subjected to a second 5-d null period of singing with zero pitch shift and no introduced delay. The birds were then subjected to a delayed auditory feedback and pitch shift block (DAF+PS) lasting 5 d. During the DAF+PS block, delayed feedback at 175 ms was concur-

rently pitch-shifted up or down by three semitones. Both the sign of the pitch shift (i.e., upward or downward) in the DAF+PS condition and the order of the DAF and DAF+PS blocks were varied across birds to counterbalance any learning order effects.

The magnitude of the pitch shift (± 3 semitones) was selected based on prior research on the effects of auditory manipulations on both songbirds and humans. As described in the Introduction, smaller shifts in the pitch or formant structure of auditory feedback evoke robust changes in vocal output. Larger shifts, on the other hand, evoke less robust changes in the pitch of human speech (Burnett et al., 1998; Liu and Larson, 2007). Importantly, large pitch shifts have been shown to reduce stuttering in human speech (Kalinowski et al., 1993; Natke et al., 2001). Because one motivation for our songbird studies was to develop an animal model of how the acoustic and temporal features of auditory feedback interact during human speech, we chose a pitch shift magnitude (± 3 semitones) that has been shown previously to be too large to evoke changes in vocal pitch in Bengalese finches, although smaller shifts do evoke robust pitch changes in song (Sober and Brainard, 2012). Our selection of this pitch shift magnitude therefore represents an attempt to maximize the correspondence between our work in songbirds and prior studies of speech.

Measuring song syntax features

As in previous uses of the headphones paradigm (Sober and Brainard, 2009), we analyzed songs produced during a fixed time window (here, 8 am to noon). In cases in which birds produced more than 60 bouts of song during this interval, we used only 60 bouts (spaced evenly across the interval) in the analysis. Syllable onsets and offsets were determined using an amplitude threshold, and song syllables were assigned arbitrary labels (e.g., a–f in Fig. 1) by visual inspection. Note that the use of the same letters for labeling syllables across different birds does not indicate acoustic similarities between the birds' syllables.

We examined syllable sequencing in two contexts: branch points and repeated syllables. At a branch point, a single syllable can be followed by multiple different syllables. Such sequence variability is a hallmark of Bengalese finch song (Okanoya, 2004; Wohlgemuth et al., 2010; Matheson and Sakata, 2015), and branch point probabilities are actively maintained during vocal learning (Warren et al., 2012). At each branch point, we quantified the probability of each transition (e.g., Fig. 1B). We used a z-test for proportions to compare probabilities from altered auditory feedback conditions to the null condition immediately preceding it. For a group analysis of the effects of DAF and DAF+PS on branch point probabilities across birds, we used a one-sided Wilcoxon signed rank test to evaluate our hypothesis that the effects of delayed auditory feedback would be reduced if the delay were performed in the presence of a pitch shift. We performed this statistical test only on changes in the probability of the primary (most common) transition, for two reasons. First, and most importantly, the probabilities of primary

and nonprimary transitions at a single branch point are not independent. For example, if there are only two transitions and one increases by 10%, then the other must decrease by the same amount, so it would be incorrect to consider changes in two transitions at a single branch points as separate measurements. Second, we focused on the primary transition to evaluate our hypothesis that delayed auditory feedback (in both the DAF and DAF+PS condition) would lead to a reduction in the probability of the primary transition, as observed previously in a similar experiment (Sakata and Brainard, 2006). In total, the four birds used in our studies yielded a total of nine branch points (one to four per bird), consisting of four cases in which one syllable could be followed by one of two different syllables, and five cases in which one syllable could be followed by one of three different syllables.

Bengalese finches also commonly produce repeated syllables (e.g., syllable c in Fig. 3A), which are produced multiple times in succession. We quantified the distribution of repeat numbers for each repeated syllable in each tested auditory feedback condition (for example, the excerpt of song shown in Fig. 3A contains a case in which syllable c is repeated four times). We used a Kolmogorov–Smirnov test to determine whether the repeat distributions of individual syllables differed significantly across feedback conditions, comparing the repeat distribution in the DAF or DAF+PS condition with that of the null period immediately preceding it. As in the analysis of branch point probabilities, we used a Wilcoxon signed-rank test to determine whether the change in mean repeat number from null to DAF was significantly larger than the changes induced by the DAF+PS. In all statistical tests of both branch points and repeated syllables, we used data from only the last 3 d of each auditory feedback condition. The four birds examined yielded a total of 20 repeated syllables (three to nine per bird). In all analyses described above, for each branch point or repeated syllable, we combined data across the last 3 d of the 5-d feedback epoch (null, DAF, or DAF+PS) when computing the effects of each feedback condition.

Results

As hypothesized, delaying auditory feedback often induced changes in syllable sequencing. Fig. 3 shows data from one branch point. In the first null period, syllable b was followed by syllable c >95% of the time (green trace, Fig. 3B), and was therefore the primary transition at that branch point (see Methods). Syllable b was followed by syllable a <5% of the time (orange trace, Fig. 3B). During the DAF condition (blue shaded region, Fig. 3B), transition probabilities gradually shifted, with the b-to-c transition becoming less common and the b-to-a transition becoming more common. Fig. 4A summarizes the effects of DAF on transition probabilities across all branch points examined. Our dataset contained nine branch points with a mean of 314 iterations (range 71–739) of each branching sequence. In seven of nine cases, delayed feedback led to a significant reduction in the probability of the primary transition (Fig. 4A, filled symbols, $p < 0.05$, z-test for proportions). When considered as a group, transition

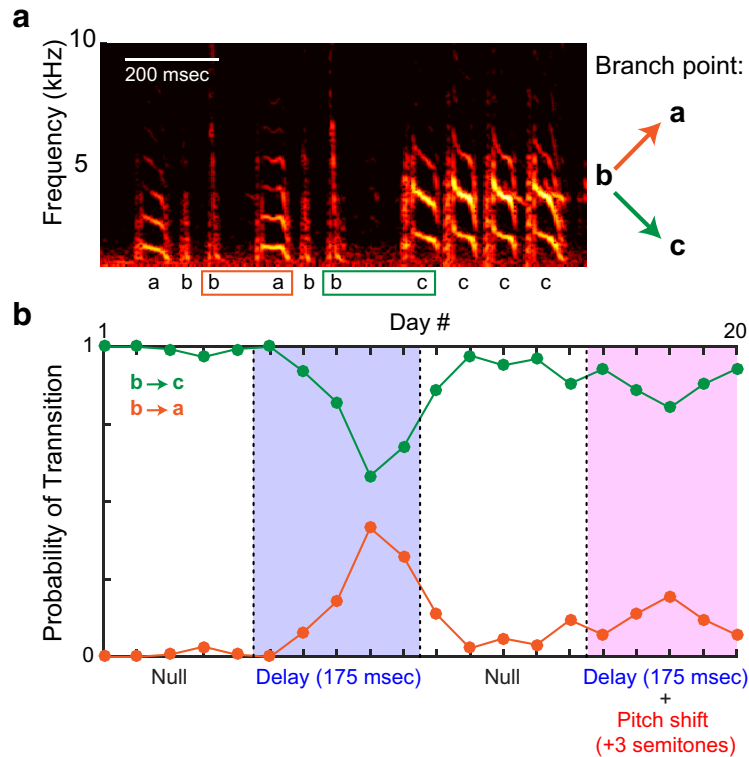


Figure 3. Effects of auditory feedback manipulations on branch point probabilities. **A**, One branch point (in which syllable b can be followed by either syllable a or c), of nine total branch points in our dataset. Spectrogram plotting conventions as in Fig. 1A. **B**, In this experiment, after a null period of unmanipulated auditory feedback, the bird experienced DAF followed by another period of unmanipulated feedback, followed by DAF+PS. Green and orange traces show the probability of the b-to-c and b-to-a transitions, respectively.

probabilities decreased significantly as a result of DAF being applied ($p < 0.01$, one-sided Wilcoxon signed rank test).

In the example shown in Fig. 3, the DAF+PS condition (pink shaded region) had a similar, but smaller, effect on syllable sequencing than DAF, with the b-to-c transition becoming slightly less prevalent in the DAF+PS epoch compared to the preceding null period. Fig. 4B shows the effects of DAF+PS across transition points. Similar to the DAF condition (Fig. 4A), DAF+PS induced significant changes in most cases (Fig 4B, filled symbols, $p < 0.05$, z-test for proportions) and as a group exhibited a significant reduction in the probability of the primary transition ($p < 0.01$, one-sided Wilcoxon signed rank test).

We then asked whether, consistent with our hypothesis, the changes induced by DAF+PS were smaller than those induced by DAF alone. Fig. 4C compares the change in transition probability induced by DAF (Δp , delay) with that induced by DAF+PS (Δp , delay + shift). As hypothesized, the effects of DAF+PS were significantly smaller ($p < 0.05$, one-sided Wilcoxon signed rank test).

Notably, although overall DAF+PS produced significantly smaller changes in transition probability than DAF, in one case (triangle symbols in Fig. 4), much larger changes were observed in the DAF+PS condition. Data from this branch point are shown in Fig. 5. Therefore, it is important to emphasize that although the group analysis demonstrated smaller changes once pitch shifts were

added to delays, the opposite was seen in one individual case.

We also examined the effect of auditory feedback manipulations on the distribution of repeated syllables (Fig. 6A shows an example containing four different repeated syllables). Fig. 6B shows an example from our dataset in which DAF induces a significant change in the distribution of repeats of syllable g ($p < 0.05$, Kolmogorov–Smirnov test). As shown in Fig. 7A, DAF frequently led to significant changes in repeat distribution (filled symbols), although there was no significant bias toward increases or decreases in mean repeat number ($p = 0.13$, two-sided Wilcoxon signed rank test). The DAF+PS condition (Fig. 7B) also induced significant changes in the repeat distribution in many cases. Interestingly, in the majority of these cases (16/20), DAF+PS reduced the mean number of repeats ($p < 0.01$, two-sided Wilcoxon signed rank test). Fig. 7C, D shows the same data as Fig. 7A, B, respectively, but represented as a change in repeat number between the null and DAF/DAF+PS conditions.

We next evaluated our hypothesis that the DAF+PS condition would evoke smaller changes in repeat number than the DAF condition. Comparing these changes (Fig. 7E) did not reveal any significant difference between the two conditions ($p = 0.65$, two-sided Wilcoxon signed rank test). We further asked whether any differences existed between the data from the two conditions shown in Fig. 7E by performing a two-sample Kolmogorov–Smirnov

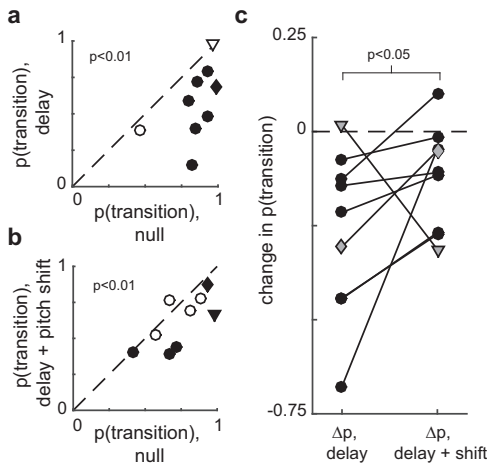


Figure 4. Effects of auditory feedback manipulations on branch points (group data). **A**, Probability of the primary (i.e., most common; see Methods) transition in the null versus DAF conditions. Filled circles indicate that the difference between each probability in the null and DAF condition was statistically significant ($p < 0.05$, z-test for proportions). Across all cases, transition probabilities significantly decreased as a result of DAF ($p < 0.01$, one-sided Wilcoxon signed rank test). **B**, Transition probabilities in the null versus DAF+PS condition. Other plotting conventions as in **A**. **C**, Comparison of the change in transition probability induced by the DAF and DAF+PS conditions. As hypothesized, the change in probability was significantly smaller in the DAF+PS condition ($p < 0.05$, one-sided Wilcoxon signed rank test). In all plots, diamond symbols indicate data from the example shown in Fig. 3, and triangle symbols indicate data from the example shown in Fig. 5.

test, which similarly failed to detect any significant difference ($p = 0.77$). Therefore, although in individual cases both branch point probabilities (Fig. 4A, B) and repeat number (Fig. 7A, B) were often significantly modulated by DAF or DAF+PS, the effects of these two alterations of auditory feedback differed significantly only for branch point probabilities (Fig. 4C) and not for the distribution of repeated syllables (Fig. 7E).

Discussion

Manipulation of auditory feedback induced robust sequence changes in the song of adult Bengalese finches. As hypothesized, both DAF and DAF+PS induced changes in transition probabilities and repeat length distributions in a substantial number of individual cases (filled symbols, Figs. 4A, B and 7A, B). At branch points, both feedback manipulations induced a reduction in the probability of the primary transition (Fig. 4A, B). In contrast, whereas DAF did not significantly bias changes in mean repeat number upward or downward (Fig. 7A, C), DAF+PS caused a reduction in mean repeat number in a significant fraction of cases (Fig. 7B, D). Together, these results demonstrate that continuous, real-time manipulation of auditory feedback can strongly modulate vocal performance and that in some contexts, the addition of a pitch shift can significantly reduce the vocal changes induced by auditory feedback delays.

A number of studies have examined the consequences of subjecting birds to song-triggered playbacks of (previously recorded) samples of the bird's own song, a manipulation that approximates delayed auditory feedback. In contrast, our technique uses miniaturized headphones to continuously stream manipulated auditory feedback. It is possible that this methodological difference accounts for some apparent discrepancies between our results in the DAF condition and prior findings. Notably, Sakata and Brainard (2006) used playbacks of one to three song syllables at specific times during Bengalese finch songs. Similar to our findings, they found that playbacks targeted to branch points reduced the probability of the primary transition (Sakata and Brainard, 2006). In contrast to our findings, however, they noted that the effects of feedback manipulation were observed very soon after the manipulation was introduced and did not increase with continued exposure, whereas in many of our experiments (Figs. 3B and 5), the magnitude of DAF effects on sequence grew steadily over the first few days of exposure. Although the variation in behavioral effects might reflect the different methods of altering auditory feedback, additional studies

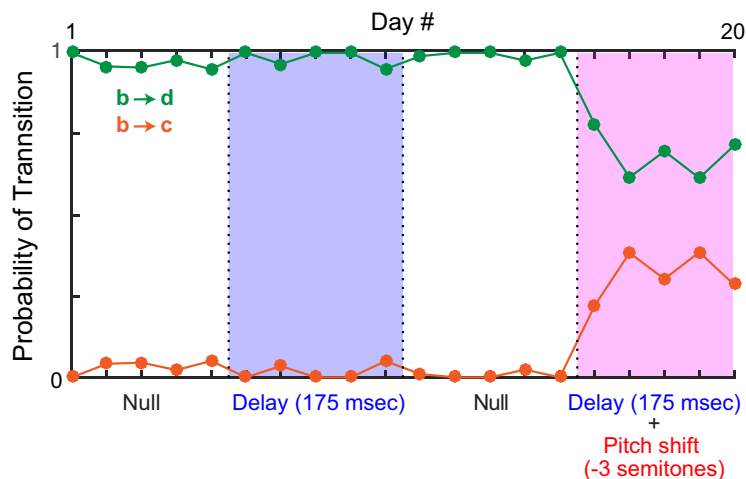


Figure 5. Additional example of effects of feedback manipulations on branch point probabilities. Shown is data from one branch point (of nine total branch points in our dataset). Plotting conventions as in Fig. 3. Data are from the branch point also shown in Fig. 1.

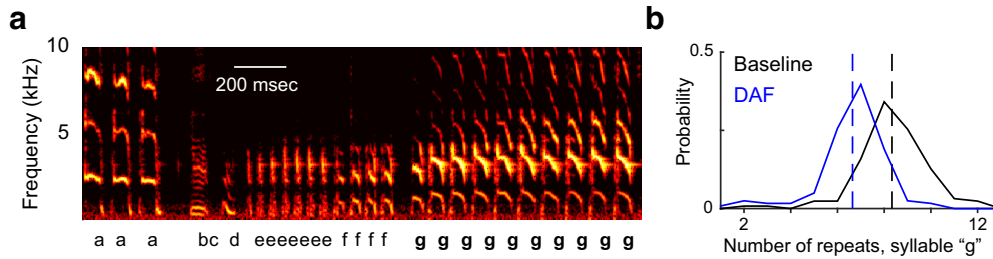


Figure 6. Effects of auditory feedback manipulations on repeat lengths. **A**, Spectrogram shows excerpt of a song that contains three repeated syllables (e, f, and g). In this excerpt, syllable g is repeated 10 times. Spectrogram plotting conventions as in Fig. 1A. **B**, Distribution of repeat lengths of syllable g in the null condition (black solid line) and DAF condition (blue solid line). Dashed lines show the mean of each distribution.

would be required to isolate the effects of continuous, real-time feedback (our study) versus intermittent, prerecorded feedback (Sakata and Brainard, 2006) from other methodological differences between the two studies, including the total time of exposure to altered feedback and the magnitude of the feedback delay.

The headphones apparatus greatly attenuates airborne transmission of a bird’s vocalization, replacing it with the manipulated version played through the headphone speakers. However, as discussed elsewhere, subjects might receive unmanipulated acoustic feedback via bone conduction, in which sound is transmitted via body tissues rather than air (Sober and Brainard, 2009). Although we cannot rule out some influence of bone conduction, we note that this factor presumably applies in both the DAF and DAF+PS conditions, and therefore seems unlikely to account for the differing effects of these results. We further note that potential bone conduction signals are only one of several sensory modalities that can convey unmanipulated feedback, with proprioceptive/somatosensory systems additionally providing information about the birds’ actual motor output.

Our findings highlight the importance of the characteristics of auditory feedback on vocal behavior. As shown in Fig. 4C, DAF elicited significantly larger changes in branch point transition probability than did DAF+PS, as hypothesized. This finding is significant for two reasons. First, it parallels similar findings in persons who stutter. The vocal sequencing errors that typify stuttering can be reduced by the application of an online pitch shift (Kalinowski et al., 1993; Natke et al., 2001; Büchel and Sommer, 2004). Our analogous finding in songbirds (i.e., that delay-induced sequencing changes can be partly reversed by pitch shifts) suggests that songbirds might be used as an animal model of how temporal and acoustic properties of auditory feedback might be manipulated to enhance the fluency of human speech. Second, our findings suggest that pitch shifts reduce songbirds’ reliance on auditory feedback when sequencing vocal behavior. A prior study employing pitch shifts, but not delays, found that while smaller (± 0.5 or 1.0 semitone) pitch shifts evoke compensatory changes in vocal pitch, ± 3.0 semitone pitch shifts did not evoke robust changes in vocal acoustics (Sober and Brainard, 2012). The present findings suggest that

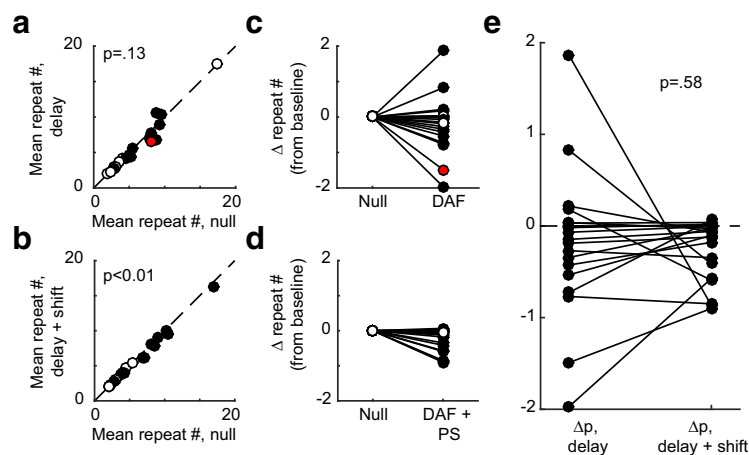


Figure 7. Effects of auditory feedback manipulations on repeat lengths (group data). **A**, Mean repeat length in the null versus DAF conditions. Filled circles indicate that the difference between the repeat distribution in the null and DAF condition was statistically significant ($p < 0.05$, two-sample Kolmogorov–Smirnov test). Across all cases, mean repeat numbers did not differ significantly as a result of DAF ($p = 0.13$, one-sided Wilcoxon signed rank test). **B**, Mean repeat lengths in the null versus DAF+PS condition. Other plotting conventions as in **A**. **C**, **D**, The same data as panels **A** and **B**, respectively, displayed at the absolute difference in mean repeat number. Red dot in **A** and **C** corresponds to the data shown in Fig. 6. **E**, Comparison of the change in mean repeat number induced by the DAF and DAF+PS conditions. No significant difference was detected between the changes in repeat number in the DAF and DAF+PS conditions ($p = 0.65$, one-sided Wilcoxon signed rank test).

pitch-shifted auditory feedback is similarly disregarded when animals program upcoming vocal sequences.

Our analyses did not reveal any significant difference in the effects on repeat number evoked by DAF and DAF+PS. Although further refinements of our technique, such as testing of other delay magnitudes, might reveal such a difference, it is also possible that these two forms of variable sequencing (branch points and repeated syllables) differ in their reliance on the acoustic structure of auditory feedback (Wittenbach et al., 2015). Future studies could examine this possibility by investigating the effects of sensory perturbations on behavior and neural activity during vocal production.

References

- Büchel C, Sommer M (2004) What causes stuttering? PLoS Biol 2:E46. [CrossRef Medline](#)
- Burnett TA, Freedland MB, Larson CR, Hain TC (1998) Voice F0 responses to manipulations in pitch feedback. J Acoust Soc Am 103:3153–3161. [Medline](#)
- Chase RA (1958) Effect of delayed auditory feedback on the repetition of speech sounds. J Speech Hear Disord 23:583–590. [Medline](#)
- Cynx J, von Rad U (2001) Immediate and transitory effects of delayed auditory feedback on bird song production. Anim Behav 62:305–312.
- Fairbanks G (1955) Selective vocal effects of delayed auditory feedback. J Speech Hear Disord 20:333–346. [Medline](#)
- Hampton A, Weber-Fox C (2008) Non-linguistic auditory processing in stuttering: evidence from behavior and event-related brain potentials. J Fluency Disord 33:253–273. [CrossRef Medline](#)
- Hoffmann LA, Sober SJ (2014) Vocal generalization depends on gesture identity and sequence. J Neurosci 34:5564–5574. [CrossRef Medline](#)
- Hoffmann LA, Kelly CW, Nicholson DA, Sober SJ (2012) A lightweight, headphones-based system for manipulating auditory feedback in songbirds. J Vis Exp e50027.
- Jones JA, Munhall KG (2000) Perceptual calibration of F0 production: evidence from feedback perturbation. J Acoust Soc Am 108:1246–1251. [Medline](#)
- Kalinowski J, Stuart A (1996) Stuttering amelioration at various auditory feedback delays and speech rates. Eur J Disord Commun 31:259–269. [Medline](#)
- Kalinowski J, Armson J, Roland-Mieszkowski M, Stuart A, Gracco VL (1993) Effects of alterations in auditory feedback and speech rate on stuttering frequency. Lang Speech 36: (Pt 1):1–16. [CrossRef](#)
- Katseff S, Houde J, Johnson K (2012) Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback? Lang Speech 55:295–308. [Medline](#)
- Kelly CW, Sober SJ (2014) A simple computational principle predicts vocal adaptation dynamics across age and error size. Front Integr Neurosci 8:75. [CrossRef Medline](#)
- Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. Nature 399:466–470. [CrossRef Medline](#)
- Liu H, Larson CR (2007) Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. J Acoust Soc Am 122:3671–3677. [CrossRef Medline](#)
- Lombardino AJ, Nottebohm F (2000) Age at deafening affects the stability of learned song in adult male zebra finches. J Neurosci 20:5054–5064. [Medline](#)
- MacDonald EN, Goldberg R, Munhall KG (2010) Compensations in response to real-time formant perturbations of different magnitudes. J Acoust Soc Am 127:1059–1068. [CrossRef Medline](#)
- Matheson AM, Sakata JT (2015) Relationship between the sequencing and timing of vocal motor elements in birdsong. PLoS One 10:e0143203. [CrossRef Medline](#)
- McGarr NS (1983) The intelligibility of deaf speech to experienced and inexperienced listeners. J Speech Hear Res 26:451–458. [Medline](#)
- Natke U, Grosser K, Kalveram K (2001) Fluency, fundamental frequency, and speech rate under frequency-shifted auditory feedback in stuttering and nonstuttering persons. J Fluency Disorders 26:227–241. [CrossRef](#)
- Okanoya K (2004) The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. Ann N Y Acad Sci 1016:724–735. [CrossRef Medline](#)
- 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. [Medline](#)
- Ryan BP, Van Kirk B (1974) The establishment, transfer, and maintenance of fluent speech in 50 stutterers using delayed auditory feedback and operant procedures. J Speech Hear Disord 39:3–10. [Medline](#)
- Sakata JT, Brainard MS (2006) Real-time contributions of auditory feedback to avian vocal motor control. J Neurosci 26:9619–9628. [CrossRef Medline](#)
- Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. Nat Neurosci 12:927–931. [CrossRef Medline](#)
- Sober SJ, Brainard MS (2012) Vocal learning is constrained by the statistics of sensorimotor experience. Proc Natl Acad Sci U S A 109:21099–21103. [CrossRef Medline](#)
- Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively maintained in a well learned motor skill. J Neurosci 32:15414–15425. [CrossRef Medline](#)
- Wittenbach JD, Bouchard KE, Brainard MS, Jin DZ (2015) An adapting auditory-motor feedback loop can contribute to generating vocal repetition. PLoS Comput Biol 11:e1004471. [CrossRef Medline](#)
- Wohlgemuth MJ, Sober SJ, Brainard MS (2010) Linked control of syllable sequence and phonology in birdsong. J Neurosci 30:12936–12949. [CrossRef Medline](#)
- 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. [Medline](#)
- Yates AJ (1963) Delayed auditory feedback. Psychol Bull 60:213–232. [Medline](#)