

---

**Research Article: New Research | Sensory and Motor Systems**

## **The Effects of Pitch Shifts on Delay-induced Changes in Vocal Sequencing in a Songbird**

Vocal Sequencing in a Songbird

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

*Department of Biology, Emory University, Atlanta, GA, USA*

DOI: 10.1523/ENEURO.0254-16.2017

Received: 28 August 2016

Revised: 16 December 2016

Accepted: 1 January 2017

Published: 11 January 2017

---

**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.

**Funding:** NSF  
1456912

**Funding:** NIH  
NS084844

**Funding:** Woodruff Foundation  
100002411  
n/a

**Conflict of Interest:** Authors report no conflict of interest.

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

**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](mailto:samuel.j.sober@emory.edu)

**Cite as:** eNeuro 2017; 10.1523/ENEURO.0254-16.2017

**Alerts:** Sign up at [eneuro.org/alerts](http://eneuro.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

Accepted manuscripts are peer-reviewed but have not been through the copyediting, formatting, or proofreading process.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Copyright © 2017 the authors

1 **The Effects of Pitch Shifts on Delay-induced Changes in Vocal Sequencing in a Songbird**

2  
3 Abbreviated title: Vocal Sequencing in a Songbird

4  
5 MacKenzie Wyatt, Emily A. Berthiaume, Conor W. Kelly, and Samuel J. Sober  
6 Department of Biology, Emory University, Atlanta, GA, USA  
7

8 Author contributions:

9 M.W., E.A.B., and S.J.S. designed research; M.W., E.A.B., and C.W.K. performed research; M.W.,  
10 E.A.B., and S.J.S. analyzed data and wrote the paper.  
11

12 Correspondence should be addressed to:

13 Samuel J. Sober  
14 Emory University Department of Biology  
15 1510 Clifton Rd. NE, Rm 2006  
16 Atlanta, GA 30322  
17 samuel.j.sober@emory.edu  
18

19 Number of tables: 0

20 Number of figures: 7

21 Number of multimedia: 0  
22

23 Number of words for Abstract: 246

24 Number of words for Significance Statement: 75

25 Number of words for Introduction: 787

26 Number of words for Discussion: 767  
27

28 Acknowledgements

29 The authors thank Lukas Hoffmann, David Nicholson, and Kyle Srivastava for helpful conversations.  
30

31 Present address for MacKenzie Wyatt:  
32 University of Kentucky College of Medicine  
33 UK Medical Center MN 150  
34 800 Rose St.  
35 Lexington, KY 40506  
36

37 Present address for Emily Berthiaume:  
38 David Geffen School of Medicine at UCLA  
39 10833 Le Conte Ave  
40 Los Angeles, CA 90095  
41

42 Present address for Conor Kelly:  
43 Medical College of Georgia  
44 1459 Laney Walker Blvd.  
45 Augusta, GA 30912  
46

47 Conflict of Interest: Authors report no conflict of interest.  
48

49 Funding sources: National Institutes of Health R01 NS084844 (S.J.S), National Science Foundation  
50 Grant 1456912 (S.J.S), Woodruff Foundation (M.W.).

51

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

70

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

77

78 **Introduction**

79 Learned vocal behaviors depend strongly on auditory feedback. In both birdsong and human speech,  
80 adults rely on auditory feedback to detect and correct errors in vocal production. This reliance on  
81 auditory information can be demonstrated by manipulating auditory feedback and measuring the  
82 effects on vocal output. Complete elimination of auditory feedback by and deafening in adulthood  
83 leads to dramatic vocal performance deficits (McGarr, 1983; Okanoya and Yamaguchi, 1997; Woolley  
84 and Rubel, 1997; Lombardino and Nottebohm, 2000). More subtle manipulations of auditory signals  
85 reveal the complex influence of sensory feedback on motor programming. Artificially delaying auditory  
86 feedback in human speakers can cause vocal sequencing errors, including unwanted repetitions of  
87 consonants and words, in normally fluent speakers (Fairbanks, 1955; Chase, 1958; Yates, 1963).  
88 Such results suggest that the sequencing errors observed in persons who stutter might result from  
89 disorders of auditory feedback processing (Buchel and Sommer, 2004; Hampton and Weber-Fox,  
90 2008). Intriguingly, artificially delaying auditory feedback is sometimes effective as a treatment for  
91 stuttering (Ryan and Van Kirk, 1974; Kalinowski and Stuart, 1996), further linking the dependence of  
92 vocal sequencing on the timing of auditory feedback and emphasizing the complex relationship  
93 between sensory feedback and speech production. Analogously, studies of birdsong have shown that  
94 perturbations of auditory feedback timing can degrade vocal production. Delayed playbacks of a  
95 bird's own syllable during singing leads to song degradation after chronic exposure in zebra finches  
96 (Leonardo and Konishi, 1999; Cynx and von Rad, 2001). In Bengalese finches, a species whose song  
97 contains "branch points" where vocal sequencing is probabilistic rather than fixed (Fig. 1), acute  
98 changes in vocal sequencing can result from delayed playbacks of a bird's own song syllable while  
99 singing (Sakata and Brainard, 2006). The similarities of these results across species suggest  
100 songbirds as a promising animal model for disorders of human speech production.

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

121 We used Bengalese finches to investigate whether delay-induced changes in vocal production  
122 were influenced by alterations in the pitch of auditory feedback. Although the effects of auditory  
123 feedback delays on song have been tested previously using song-triggered playbacks of previously-  
124 recorded songs, technical challenges have prevented the use of continuous delayed feedback. We  
125 overcame this obstacle using miniaturized headphones (see Methods), which provided continuous,

126 delayed feedback in real time (Hoffmann et al., 2012). Experimental conditions included a null  
127 condition (no delay or pitch manipulation), delayed auditory feedback (DAF) without any pitch shift,  
128 and a condition in which auditory feedback was both delayed and pitch shifted (DAF+PS). We  
129 predicted that, in agreement with prior findings using playbacks of short segments of song (Sakata  
130 and Brainard, 2006), delayed feedback would induce changes in the syllable transition probabilities.  
131 In particular, we hypothesized that, as reported by Sakata and Brainard (2006), delayed feedback  
132 would result in the most common transition (the “primary” transition) becoming less prevalent, and the  
133 non-primary transition becoming more common. We further hypothesized that the large pitch shift in  
134 the DAF+PS condition would reduce the magnitude of the changes observed in the DAF condition.

### 135 **Materials and Methods**

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

#### 142 *Experimental procedure.*

143 Miniature, lightweight headphones were custom-built out of lightweight carbon fiber and custom-fit to  
144 each bird’s head (Hoffmann et al., 2012). A condenser microphone in the bird’s cage (Fig. 2a)  
145 captured the birds acoustic output, which was routed to online sound-processing hardware (Eventide  
146 H7600), which provided perturbations of the pitch and/or timing of auditory feedback in real time. This  
147 manipulated feedback was then relayed to miniaturized speakers (EH-7157-000, Knowles) inside the  
148 headphones. In addition to the speakers, the headphones apparatus included a miniaturized  
149 microphone (EM-3046, Knowles) placed between the speaker and the opening of one of the ear  
150 canals. This microphone allowed us to monitor the performance of the headphones apparatus. The

151 miniaturized microphone was used to calibrate the system such that the acoustic signal played  
152 through the headphones speakers was  $\sim 2$  log units greater than auditory feedback leaking through  
153 the carbon fiber frame. The headphones therefore shielded the birds' airborne vocalizations, allowing  
154 the altered feedback to replace the natural version. As described in detail below, auditory feedback  
155 conditions included a null condition in which no manipulation was introduced, a "delayed auditory  
156 feedback" (DAF) condition in which auditory feedback was delayed by 175 msec, or a "delay + pitch  
157 shift" (DAF+PS) condition in which both a 175 msec delay and an upward or downward pitch shift  
158 were applied simultaneously. As detailed previously, the sound processing hardware relayed the  
159 online acoustic signal to the headphones with a minimal delay (i.e. when auditory feedback was not  
160 being intentionally delayed) of  $\sim 10$  msec, a delay which does not in itself evoke any measurable  
161 changes in vocal behavior (Sober and Brainard, 2009; Hoffmann et al., 2012; Kelly and Sober, 2014).  
162 Note that subjects might receive some unmanipulated auditory feedback via bone conduction,  
163 although such a scenario is very unlikely to account for our results (see Discussion).

164 The experimental design is outlined in Figure 2b. First, once birds habituated to the  
165 headphones, all birds sang during a period with zero pitch shift or delay for five days. After this null  
166 period, in the example shown at top in Figure 2b, the birds' auditory output was altered with delayed  
167 auditory feedback ("DAF block") for five days. The birds heard their natural vocalizations at 175 msec  
168 delay relative to output. We selected this delay magnitude after preliminary studies in two birds (not  
169 used in the present study) suggested that this delay consistently evokes changes in vocal sequence;  
170 however different delay values were not tested systematically. After the altered auditory feedback  
171 block, birds were subjected to second five-day null period of singing with zero pitch shift and no  
172 introduced delay. The birds were then subjected to a delayed auditory feedback and pitch shift block  
173 ("DAF+PS") lasting five days. During the DAF+PS block, delayed feedback at 175 msec was  
174 concurrently pitch shifted up or down by three semitones. Both the sign of the pitch shift (i.e. upward



175 or downward) in the DAF+PS condition and the order of the DAF/DAF+PS blocks were varied across  
176 birds to counterbalance for any learning order effects.

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

189 *Measuring song syntax features.*

190 As in previous uses of the headphones paradigm (Sober and Brainard, 2009), we analyzed songs  
191 produced during a fixed time window (here, 8 am - 12 noon). In cases where birds produced more  
192 than sixty bouts of song during this interval, we used only sixty bouts (spaced evenly across the  
193 interval) in the analysis. Syllable onsets and offsets were determined using an amplitude threshold,  
194 and song syllables were assigned arbitrary labels (e.g. a-f in Fig. 1) by visual inspection. Note that the  
195 use of the same letters for labeling syllables across different birds does not indicate acoustic  
196 similarities between the birds' syllables.

197 We examined syllable sequencing in two contexts: branch points and repeated syllables. At a  
198 branch point, a single syllable can be followed by multiple different syllables. Such sequence  
199 variability is a hallmark of Bengalese finch song (Okanoya, 2004; Wohlgemuth et al., 2010; Matheson



200 and Sakata, 2015), and branch point probabilities are actively maintained during vocal learning  
201 (Warren et al., 2012). At each branch point we quantified the probability of each transition (e.g. Fig.  
202 1b). We used a z-test for proportions to compare probabilities from altered auditory feedback  
203 conditions to the null condition immediately preceding it. For a group analysis of the effects of DAF  
204 and DAF+PS on branch point probabilities across birds, we used a one-sided Wilcoxon signed rank  
205 test to evaluate our hypothesis that the effects of delayed auditory feedback would be reduced if the  
206 delay were performed in the presence of a pitch shift. We performed this statistical test only on  
207 changes in the probability of the primary (most common) transition, for two reasons. First and most  
208 importantly, the probabilities of primary and non-primary transitions at a single branch point are not  
209 independent. For example, if there are only two transitions and one increases by 10%, then the other  
210 must decrease by the same amount, so it would be incorrect to consider changes in two transitions at  
211 a single branch points as separate measurements. Second, we focused on the primary transition to  
212 evaluate our hypothesis that delayed auditory feedback (in both the DAF and DAF+PS condition)  
213 would lead to a reduction in the probability of the primary transition, as observed previously in a  
214 similar experiment (Sakata and Brainard, 2006). In total, the four birds used in our studies yielded a  
215 total of 9 branch points (1-4 per bird), consisting of four cases in which one syllable could be followed  
216 by one of two different syllables, and five cases in which one syllable could be followed by one of  
217 three different syllables.

218 Bengalese finches also commonly produce “repeated” syllables (e.g. syllable “c” in Fig. 3a),  
219 which are produced multiple times in succession. We quantified the distribution of repeat numbers for  
220 each repeated syllable in each tested auditory feedback condition (for example, the excerpt of song  
221 shown in Fig. 3a contains a case in which syllable “c” is repeated four times). We used a Kolmogorov-  
222 Smirnov test to determine whether the repeat distributions of individual syllables differed significantly  
223 across feedback condition, comparing the repeat distribution in the DAF or DAF+PS condition with  
224 that of the null period immediately preceding it. As in the analysis of branch point probabilities, we

225 used a Wilcoxon signed-rank test to determine whether the change in mean repeat number from null  
226 to DAF was significantly larger than the changes induced by the DAF+PS. In all statistical tests of  
227 both branch points and repeated syllables, we used data from only the last three days of each  
228 auditory feedback conditions. The four birds examined yielded a total of 20 repeated syllables (3-9  
229 per bird). In all analyses described above, for each branch point or repeated syllable we combined  
230 data across the last three days of the five-day feedback epoch (null, DAF, or DAF+PS) when  
231 computing the effects of each feedback condition.

## 234 **Results**

235 As hypothesized, delaying auditory feedback often induced changes in syllable sequencing. Figure 3  
236 shows data from one branch point. In the first null period, syllable “b” was followed by syllable “c”  
237 more than 95% of the time (green trace, Fig. 3b), and was therefore the primary transition at that  
238 branch point (see Methods). Syllable “b” was followed by syllable “a” less than 5% of the time (orange  
239 trace, Fig. 3b). During the DAF condition (blue shaded region, Fig. 3b) transition probabilities  
240 gradually shifted, with the b→c transition becoming less common and the b→a transition becoming  
241 more common. Figure 4a summarizes the effects of DAF on transition probabilities across all branch  
242 points examined. Our dataset contained 9 branch points with a mean of 314 (range 71-739) iterations  
243 of each branching sequence. In 7 of 9 cases, delayed feedback led to a significant reduction in the  
244 probability of the primary transition (Fig. 4a, filled symbols,  $p < 0.05$ , z-test for proportions). When  
245 considered as a group, transition probabilities decreased significantly as a result of DAF being  
246 applied ( $p < 0.01$ , one-sided Wilcoxon signed rank test).

247 In the example shown in Figure 3, the DAF+PS condition (pink shaded region) had a similar,  
248 but smaller, effect on syllable sequencing than DAF, with the b→c transition becoming slightly less  
249 prevalent in the DAF+PS epoch compared to the preceding null period. Figure 4b shows the effects of

250 DAF+PS across transition points. Similar to the DAF condition (Fig. 4a), DAF+PS induced significant  
251 changes in most cases (Fig 4b, filled symbols,  $p < 0.05$ , z-test for proportions) and as a group  
252 exhibited a significant reduction in the probability of the primary transition ( $p < 0.01$ , one-sided  
253 Wilcoxon signed rank test).

254 We then asked whether, consistent with our hypothesis, the changes induced by DAF+PS  
255 were smaller than those induced by DAF alone. Figure 4c compares the change in transition  
256 probability induced by DAF (“ $\Delta p$ , delay”) with that induced by DAF+PS (“ $\Delta p$ , delay + shift”). As  
257 hypothesized, the effects of DAF+PS were significantly smaller ( $p < 0.05$ , one-sided Wilcoxon signed  
258 rank test).

259 Notably, although overall DAF+PS produced significantly smaller changes in transition  
260 probability than DAF, in one case (triangle symbols in Fig. 4), much larger changes were observed in  
261 the DAF+PS condition. Data from this branch point is shown in Figure 5. Therefore, it is important to  
262 emphasize that although the group analysis demonstrated smaller changes once pitch shifts were  
263 added to delays, the opposite was seen in one individual case.

264 We also examined the effect of auditory feedback manipulations on the distribution of repeated  
265 syllables (Fig. 6a shows an example containing four different repeated syllables). Figure 6b shows an  
266 example from our dataset in which DAF induces a significant change in the distribution of repeats of  
267 syllable “g” ( $p < 0.05$ , Kolmogorov-Smirnov test). As shown in Figure 7a, DAF frequently led to  
268 significant changes in repeat distribution (filled symbols), although there was no significant bias  
269 towards increases or decreases in mean repeat number ( $p = 0.13$ , two-sided Wilcoxon signed rank  
270 test). The DAF+PS condition (Fig. 7b) also induced significant changes in the repeat distribution in  
271 many cases. Interestingly, in the majority (16/20) of these cases, DAF+PS reduced the mean number  
272 of repeats ( $p < 0.01$ , two-sided Wilcoxon signed rank test). Figure 7c and d show the same data as  
273 Figure 7a and b, respectively, but represented as a change in repeat number between the null and  
274 DAF/DAF+PS conditions.

275 We next evaluated our hypothesis that the DAF+PS condition would evoke smaller changes in  
276 repeat number than the DAF condition. Comparing these changes (Fig. 7e) did not reveal any  
277 significant difference between the two conditions ( $p=0.65$ , two-sided Wilcoxon signed rank test). We  
278 further asked whether any differences existed between the data from the two conditions shown in Fig.  
279 7e by performing a two-sample Kolmogorov-Smirnov test, which similarly failed to detect any  
280 significant difference ( $p=0.77$ ). Therefore, although in individual cases both branch point probabilities  
281 (Fig. 4a,b) and repeat number (Fig. 7a,b) were often significantly modulated by DAF and/or DAF+PS,  
282 the effects of these two alterations of auditory feedback differed significantly only for branch point  
283 probabilities (Fig. 4c) but not for the distribution of repeated syllables (Fig. 7e).

## 284

### 285 **Discussion**

286 Manipulation of auditory feedback induced robust sequence changes in the song of adult Bengalese  
287 finches. As hypothesized, both DAF and DAF+PS induced changes in transition probabilities and  
288 repeat length distributions in a substantial number of individual cases (filled symbols, Fig. 4a and  
289 7a,b). At branch points, both feedback manipulations induced a reduction in the probability of the  
290 primary transition (Fig. 4a,b). In contrast, whereas DAF did not significantly bias changes in mean  
291 repeat number upwards or downwards (Fig. 7a,c), DAF+PS caused a reduction in mean repeat  
292 number in a significant fraction of cases (Fig. 7b,d). Together, these results demonstrate that  
293 continuous, real-time manipulation of auditory feedback can strongly modulate vocal performance  
294 and that in some contexts, the addition of a pitch shift can significantly reduce the vocal changes  
295 induced by auditory feedback delays.

296 A number of studies have examined the consequences of subjecting birds to song-triggered  
297 playbacks of (previously recorded) samples of the bird's own song, a manipulation that approximates  
298 delayed auditory feedback. In contrast, our technique employs miniaturized headphones to  
299 continuously stream manipulated auditory feedback. It is possible that this methodological difference

300 accounts for some apparent discrepancies between our results in the DAF condition and prior  
301 findings. Notably, Sakata and Brainard (2006) used playbacks of 1-3 song syllables at specific times  
302 during Bengalese finch songs. Similar to our findings, they found that playbacks targeted to branch  
303 points reduced the probability of the primary transition (Sakata and Brainard, 2006). In contrast to our  
304 findings, however, this earlier paper noted that the effects of feedback manipulation were observed  
305 very soon after the manipulation was introduced and did not increase with continued exposure,  
306 whereas in many of our experiments (Fig. 3b, 5b) the magnitude of DAF effects on sequence grew  
307 steadily over the first few days of exposure. Although the variation in behavioral effects might reflect  
308 the different methods of altering auditory feedback, additional studies would be required to isolate the  
309 effects of continuous, real-time feedback (our study) versus intermittent, pre-recorded feedback  
310 (Sakata and Brainard, 2006) from other methodological differences between the two studies,  
311 including the total time of exposure to altered feedback and the magnitude of the feedback delay.

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

321         Our findings highlight the importance of the characteristics of auditory feedback on vocal  
322 behavior. As shown in Figure 4c, DAF elicited significantly larger changes in branch point transition  
323 probability than did DAF+PS, as hypothesized. This finding is significant for two reasons. First, it  
324 parallels similar findings in persons who stutter. The vocal sequencing errors that typify stuttering can

325 be reduced by the application of an online pitch shift (Kalinowski et al., 1993; Natke et al., 2001;  
326 Buchel and Sommer, 2004). Our analogous finding in songbirds (i.e. that delay-induced sequencing  
327 changes can be partly reversed by pitch shifts) suggests that songbirds might be used as an animal  
328 model of how temporal and acoustic properties of auditory feedback might be manipulated to  
329 enhance the fluency of human speech. Second, our findings suggest that pitch shifts reduce  
330 songbirds' reliance on auditory feedback when sequencing vocal behavior. A prior study employing  
331 pitch shifts, but not delays, found that while smaller ( $\pm 0.5$  or 1.0 semitone) pitch shifts evoke  
332 compensatory changes in vocal pitch,  $\pm 3.0$  semitone pitch shifts did not evoke robust changes in  
333 vocal acoustics (Sober and Brainard, 2012). The present findings suggest that pitch-shifted auditory  
334 feedback is similarly disregarded when animals program upcoming vocal sequences.

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

#### 342 **References**

- 343 Buchel C, Sommer M (2004) What causes stuttering? PLoS Biol 2:E46.
- 344 Burnett TA, Freedland MB, Larson CR, Hain TC (1998) Voice F0 responses to manipulations in pitch  
345 feedback. J Acoust Soc Am 103:3153-3161.
- 346 Chase RA (1958) Effect of delayed auditory feedback on the repetition of speech sounds. J Speech  
347 Hear Disord 23:583-590.
- 348 Cynx J, von Rad U (2001) Immediate and transitory effects of delayed auditory feedback on  
349 bird song production. Anim Behav 62:305-312.

- 350 Fairbanks G (1955) Selective vocal effects of delayed auditory feedback. *J Speech Hear Disord*  
351 20:333-346.
- 352 Hampton A, Weber-Fox C (2008) Non-linguistic auditory processing in stuttering: evidence from  
353 behavior and event-related brain potentials. *J Fluency Disord* 33:253-273.
- 354 Hoffmann LA, Sober SJ (2014) Vocal generalization depends on gesture identity and sequence. *J*  
355 *Neurosci* 34:5564-5574.
- 356 Hoffmann LA, Kelly CW, Nicholson DA, Sober SJ (2012) A lightweight, headphones-based system for  
357 manipulating auditory feedback in songbirds. *J Vis Exp*:e50027.
- 358 Jones JA, Munhall KG (2000) Perceptual calibration of F0 production: evidence from feedback  
359 perturbation. *J Acoust Soc Am* 108:1246-1251.
- 360 Kalinowski J, Stuart A (1996) Stuttering amelioration at various auditory feedback delays and speech  
361 rates. *Eur J Disord Commun* 31:259-269.
- 362 Kalinowski J, Armson J, Roland-Mieszkowski M, Stuart A, Gracco VL (1993) Effects of alterations in  
363 auditory feedback and speech rate on stuttering frequency. *Lang Speech* 36 ( Pt 1):1-16.
- 364 Katseff S, Houde J, Johnson K (2012) Partial compensation for altered auditory feedback: a tradeoff  
365 with somatosensory feedback? *Lang Speech* 55:295-308.
- 366 Kelly CW, Sober SJ (2014) A simple computational principle predicts vocal adaptation dynamics  
367 across age and error size. *Front Integr Neurosci* 8:75.
- 368 Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory  
369 feedback. *Nature* 399:466-470.
- 370 Liu H, Larson CR (2007) Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex.  
371 *J Acoust Soc Am* 122:3671-3677.
- 372 Lombardino AJ, Nottebohm F (2000) Age at deafening affects the stability of learned song in adult  
373 male zebra finches. *J Neurosci* 20:5054-5064.



- 374 MacDonald EN, Goldberg R, Munhall KG (2010) Compensations in response to real-time formant  
375 perturbations of different magnitudes. *J Acoust Soc Am* 127:1059-1068.
- 376 Matheson AM, Sakata JT (2015) Relationship between the Sequencing and Timing of Vocal Motor  
377 Elements in Birdsong. *PLoS One* 10:e0143203.
- 378 McGarr NS (1983) The intelligibility of deaf speech to experienced and inexperienced listeners. *J*  
379 *Speech Hear Res* 26:451-458.
- 380 Natke U, Grosser K, Kalveram K (2001) Fluency, fundamental frequency, and speech rate under  
381 frequency-shifted auditory feedback in stuttering and nonstuttering persons. *J Fluency*  
382 *Disorders* 26:227-241.
- 383 Okanoya K (2004) The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax.  
384 *Ann N Y Acad Sci* 1016:724-735.
- 385 Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (*Lonchura striata var. domestica*) require  
386 real-time auditory feedback to produce normal song syntax. *Journal of Neurobiology* 33:343-  
387 356.
- 388 Ryan BP, Van Kirk B (1974) The establishment, transfer, and maintenance of fluent speech in 50  
389 stutterers using delayed auditory feedback and operant procedures. *J Speech Hear Disord*  
390 39:3-10.
- 391 Sakata JT, Brainard MS (2006) Real-time contributions of auditory feedback to avian vocal motor  
392 control. *J Neurosci* 26:9619-9628.
- 393 Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. *Nat Neurosci*  
394 12:927-931.
- 395 Sober SJ, Brainard MS (2012) Vocal learning is constrained by the statistics of sensorimotor  
396 experience. *Proc Natl Acad Sci U S A* 109:21099-21103.
- 397 Warren TL, Charlesworth JD, Tumer EC, Brainard MS (2012) Variable sequencing is actively  
398 maintained in a well learned motor skill. *J Neurosci* 32:15414-15425.

- 399 Wittenbach JD, Bouchard KE, Brainard MS, Jin DZ (2015) An Adapting Auditory-motor Feedback  
400 Loop Can Contribute to Generating Vocal Repetition. *PLoS Comput Biol* 11:e1004471.
- 401 Wohlgemuth MJ, Sober SJ, Brainard MS (2010) Linked control of syllable sequence and phonology in  
402 birdsong. *J Neurosci* 30:12936-12949.
- 403 Woolley SM, Rubel EW (1997) Bengalese finches *Lonchura striata domestica* depend upon auditory  
404 feedback for the maintenance of adult song. *J Neurosci* 17:6380-6390.
- 405 Yates AJ (1963) Delayed auditory feedback. *Psychol Bull* 60:213-232.
- 406
- 407

408

409

### **Figure legends**

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

416

417 **Figure 2: Experimental design.** a. Auditory feedback manipulated using miniaturized headphones.  
418 A microphone transmits a bird’s vocalizations to online sound processing hardware, which are used  
419 to introduce a 175-msec delay (with or without a  $\pm 3$  semitone pitch shift). b. Schedule of experimental  
420 conditions. Each bird was exposed to the delayed feedback alone (DAF) and delayed feedback plus a  
421 pitch shift (DAF+PS). The order of these conditions, as well as the direction of the pitch shift, was  
422 randomized across subjects. Prior to each DAF or DAF+PS epoch, birds sang in a “null” epoch free of  
423 delays or pitch shifts. The four experimental epochs lasted five days each.

424

425 **Figure 3: Effects of auditory feedback manipulations on branch point probabilities.** (a) One  
426 branch point (in which syllable “b” can be followed by either syllable “a” or “c”), out of 9 total branch  
427 points in our dataset. Spectrogram plotting conventions as in Figure 1a. (b) In this experiment, after a  
428 null period of unmanipulated auditory feedback the bird experienced delayed auditory feedback  
429 (DAF), followed by another period of unmanipulated feedback, followed by a combined delay and  
430 pitch shift (DAF+PS). Green and orange traces show the probability of the  $b \rightarrow c$  and  $b \rightarrow a$  transitions,  
431 respectively.

432

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

444  
445 **Figure 5: Additional example of effects of feedback manipulations on branch point**  
446 **probabilities.** Shown is data from one branch point (out of 9 total branch points in our dataset).  
447 Plotting conventions as in Figure 3. Data are from the branch point also shown in Figure 1.

448  
449 **Figure 6: Effects of auditory feedback manipulations on repeat lengths.** (a) Spectrogram shows  
450 excerpt of a song that contains three repeated syllables (“e”, “f”, and “g”). In this excerpt, syllable “g”  
451 is repeated 10 times. Spectrogram plotting conventions as in Figure 1a. (b) Distribution of repeat  
452 lengths of syllable “g” in the null condition (black solid line) and DAF condition (blue solid line).  
453 Dashed lines show the mean of each distribution.

454  
455 **Figure 7: Effects of auditory feedback manipulations on repeat lengths (group data).** (a) Mean  
456 repeat length in the null versus DAF conditions. Filled circles indicate that the difference between the  
457 repeat distribution in the null and DAF condition was statistically significant ( $p < 0.05$ , two-sample

458 Kolmogorov-Smirnov test). Across all cases, mean repeat numbers did not differ significantly as a  
459 result of DAF ( $p=0.13$ , one-sided Wilcoxon signed rank test) (b) Mean repeat lengths in the null  
460 versus DAF+PS condition. Other plotting conventions as in (a). Panels (c) and (d) show the same  
461 data as panels (a) and (b), respectively, displayed at the absolute difference in mean repeat number.  
462 Red dot in (a) and (c) corresponds to the data shown in Figure 6. (e) Comparison of the change in  
463 mean repeat number induced by the DAF (“delay”) and DAF+PS (“delay + shift”) conditions. No  
464 significant difference was detected between the changes in repeat number in the DAF and DAF+PS  
465 conditions ( $p=0.65$ , one-sided Wilcoxon signed rank test).















