



OPEN ACCESS

EDITED BY

Julia Hyland Bruno,
New Jersey Institute of Technology,
United States

REVIEWED BY

Lesley J. Rogers,
University of New England, Australia
Ana Amador,
University of Buenos Aires, Argentina

*CORRESPONDENCE

Nicolas Giret
✉ nicolas.giret@cnrs.fr

[†]These authors have contributed equally to this work

RECEIVED 01 November 2024

ACCEPTED 20 January 2025

PUBLISHED 06 February 2025

CITATION

Rolland M, Zai AT, Hahnloser RHR, Del Negro C and Giret N (2025) Visually-guided compensation of deafening-induced song deterioration.

Front. Psychol. 16:1521407.

doi: 10.3389/fpsyg.2025.1521407

COPYRIGHT

© 2025 Rolland, Zai, Hahnloser, Del Negro and Giret. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Visually-guided compensation of deafening-induced song deterioration

Manon Rolland^{1†}, Anja T. Zai^{2,3†}, Richard H. R. Hahnloser^{2,3}, Catherine Del Negro¹ and Nicolas Giret^{1*}

¹Institut des Neurosciences Paris Saclay, CNRS, Université Paris Saclay, Saclay, France, ²Institute of Neuroinformatics, ETH Zurich and UZH, Zurich, Switzerland, ³Neuroscience Center Zurich (ZNZ), University of Zurich and ETH Zurich, Zurich, Switzerland

Human language learning and maintenance depend primarily on auditory feedback but are also shaped by other sensory modalities. Individuals who become deaf after learning to speak (post-lingual deafness) experience a gradual decline in their language abilities. A similar process occurs in songbirds, where deafness leads to progressive song deterioration. However, songbirds can modify their songs using non-auditory cues, challenging the prevailing assumption that auditory feedback is essential for vocal control. In this study, we investigated whether deafened birds could use visual cues to prevent or limit song deterioration. We developed a new metric for assessing syllable deterioration called the spectrogram divergence score. We then trained deafened birds in a behavioral task where the spectrogram divergence score of a target syllable was computed in real-time, triggering a contingent visual stimulus based on the score. Birds exposed to the contingent visual stimulus—a brief light extinction—showed more stable song syllables than birds that received either no light extinction or randomly triggered light extinction. Notably, this effect was specific to the targeted syllable and did not influence other syllables. This study demonstrates that deafness-induced song deterioration in birds can be partially mitigated with visual cues.

KEYWORDS

vocal control, deafening, sensorimotor, songbird, sensory feedback, degradation

Introduction

Human language and speech learning rely on vocal learning processes that involve imitation and sensory-motor integration (Doupe and Kuhl, 1999). Infants learn to speak by imitating the vocal communication signals of the individuals around them. During development, an individual's vocalizations become more and more accurate, guided by the sensory, i.e., auditory feedback of their own voice (Doupe and Kuhl, 1999; Brainard and Doupe, 2002; Tyack, 2020; Zhang et al., 2023). Infants born deaf or with severe hearing impairments never acquire typical adult speech. In adults, post-lingual deafness—the complete loss of hearing abilities after acquisition of speech sounds—leads to a degradation of speech, including the loss of phonetic precision of vowels and consonants and the change in duration of syllables and sentences (Cowie et al., 1982; Leder et al., 1987a, 1987b, 1987c; Waldstein, 1990; Lane and Webster, 1991). Auditory feedback thus plays a critical role for adult speech maintenance. Sharing close similarities with speech learning in humans, learning and maintenance of the vocal signals produced by songbirds also critically rely on the processing of auditory feedback. Deaf juvenile songbird never manage to develop typical song of their own species (Konishi, 1964, 1965a, 1965b, 2004; Iyengar and Bottjer, 2002). Intact hearing abilities are thus required for a juvenile to learn to imitate the songs of adult conspecifics. Once

adult, songbirds keep on processing the auditory feedback that allows them to adjust their vocal behavior when facing environmental noise constraints (Okanoya and Yamaguchi, 1997; Leonardo and Konishi, 1999; Brainard and Doupe, 2000; Tumer and Brainard, 2007; Andalman and Fee, 2009; Sober and Brainard, 2009; Derryberry et al., 2020). Deafened adult songbirds exhibit a progressive and dramatic degradation of their vocal production (Nordeen and Nordeen, 1992, 2010; Woolley and Rubel, 1997; Lombardino and Nottebohm, 2000; Horita et al., 2008; Wittenbach et al., 2015).

Auditory feedback is thus critical for song maintenance, even in closed-ended learner species such as zebra finches (*Taeniopygia guttata*) producing highly stereotyped song syllables in adulthood (>90 days post-hatch, dph). Interestingly however, adult male zebra finches are able to adjust song syllables using non-auditory feedback signals, i.e., a visual signal such as a transient light extinction (Zai et al., 2020) or mild subcutaneous electric stimulations (McGregor et al., 2022). The behavioral training paradigm relies on the contingent delivery of the sensory signal depending on the pitch of a song syllable. Importantly, deaf birds modified the pitch of the selected song syllable in order to get more transient light extinction while hearing birds did the opposite (Zai et al., 2020). Furthermore, the magnitude of pitch change was much higher in deaf than in hearing birds. The intrinsic value of the transient light extinction could therefore be assumed to differ between deaf and hearing birds, being attractive for deaf birds while repulsive for hearing birds, respectively.

As deaf birds adapt the pitch of a selected song syllable using information provided by a visual signal (Zai et al., 2020), we wondered whether a similar behavioral protocol could be used to prevent or at least delay deafening-induced song degradation. Because deafening-induced song degradation includes a wide range of spectro-temporal changes in the structure of song syllables (Nordeen and Nordeen, 1992; Wang et al., 1999; Brainard and Doupe, 2001; Horita et al., 2008)—syllables uniformly and gradually get noisier (Horita et al., 2008; Hamaguchi et al., 2014)—we aimed to examine whether deafened birds could use non auditory signals to control the acoustic structure of a selected song syllable. To do so, we first developed a metric for measuring syllable acoustic stability and for estimating in real time the degree of vocal deterioration. We then modified the operant conditioning paradigm based on the pitch-contingent delivery of a transient light extinction (Zai et al., 2020) such that the visual signal was delivered depending on the syllable acoustic stability measure.

Method

Subject and groups

We used 40 adult male (2 birds were 90 days post-hatch (dph), all others were >100 dph) zebra finches (*Taeniopygia guttata*) raised in our breeding colonies in Orsay and Saclay (France) or in Zurich (Switzerland). All experiments have been approved by the French Ministry of Research and ethical committee “Paris-Sud et Centre (CEEA n°59, project 2017-12) or by the Cantonal Veterinary Office of the Canton of Zurich, Switzerland (license numbers 207/2013 and ZH077/17) and comply with the EU Council Directive 2010/63 of the European Parliament on the protection of animals used for scientific purposes.

Throughout the experiments, the birds were housed individually in sound-proof chambers under a 14-h day-10-h night photoperiod cycle, with access to food and water *ad libitum*. Birds resumed singing at a normal rate after 2–5 days in the experimental environment. Chambers were equipped with a wall-attached microphone. Sound signals were band-pass filtered, digitized at a sampling rate of 32 kHz and songs were detected online using the RecOOrder software (Herbst et al., 2023) written in LabView (National Instruments, Inc.).

Birds were divided into four groups. They included one group of hearing control birds ($n = 7$, mean age \pm std. = 217.4 ± 113.3 dph) and three groups of deafened birds: deaf LO ($n = 13$, mean age \pm std. = 122.4 ± 19.8 dph), deaf random LO ($n = 5$, mean age \pm std. = 127.0 ± 12.5 dph) and deaf no LO ($n = 15$, mean age \pm std. = 121.3 ± 17.6 dph) birds. Deaf LO birds were exposed to the syllable-contingent delivery of a transient light extinction while deaf random LO birds were exposed to a similar amount of transient light extinction than deaf LO birds, but the visual signal exposure did not depend on what the bird was singing. Deaf no LO and hearing birds were not exposed to transient light extinction during the entire course of the experiment.

Deafening procedure

All birds except hearing birds underwent the deafening procedure. To do so, we performed a bilateral cochlear ablation (Schwartzkopff, 1949; Zai et al., 2020) that results in the complete and irreversible suppression of auditory feedback. The birds were anesthetized by inhalation of a mixture of oxygen and isoflurane (induction: 2–3%; maintenance: 1–2%). Once the flexion reflex was no longer observable, birds were placed in the stereotaxic apparatus in the prone position and their beak placed at 90° to the horizontal axis. We removed the feathers between the two ears, at the lower part of the skull at the back of the head. Disinfectant (Vetedine) and local anesthetic (Lurocaine) were applied to the skin of the skull 10 min before incising it. The skin was incised over 5 millimeters at the level of the hyoid bone in the antero-posterior direction, in order to expose the neck muscles. The muscles were gently pushed down the skull to expose the cranial surface where the semicircular canals are visible through the skull. The craniotomy was performed with forceps just below the point where the posterior and external semicircular canals cross. Under a light microscope, the topography of the area was observed to determine the position of the dome forming the upper part of the bony canal containing the cochlea, above the bony crest, at a position anteromedial to the oval window. At this dome, a small window was opened with forceps to allow removal of the cochlea. The cochlea was removed from the cavity with a custom-made tungsten hook. After removal, the cochleae including the lagenas were photographed to verify their integrity. We bilaterally removed cochlea to induce total deafness. Birds resumed their singing activity (>400 daily song motifs) on average 1.6 days after the procedure.

Visual substitution task of syllable similarity: LO protocol

We ran a custom-made LabView (National Instruments, Inc., code available on the Gitlab platform: <https://gitlab.switch.ch/>)

[hahnloser-songbird/published-code/SpectrogramDivergence](#)) program to provide visual substitution contingent on syllable similarity to deafened birds. We targeted a harmonic syllable, later called “target syllable,” using a two-layer neural network (perceptron) trained on a subset of manually clustered vocalizations (Yamahachi et al., 2020). We first computed the spectrogram of each rendition of the target syllable produced on the last baseline day before the start of the LO protocol. We applied a Hamming windowing with a sample size of 512 (32 kHz sampling rate) on the raw sound signal. The windowed signal was then transformed into a linear power sound spectrogram using the fast Fourier transform, computed on segments of 512 samples with an overlap of 128 samples (corresponding to 4 ms). The resulting spectrogram has a resolution of 4 ms per column and 63 Hz per row. To compute the reference spectrogram, we selected a window of the last 48 ms before the detection performed by the perceptron. We applied a high-pass filter on the sound signal to keep frequencies above 630 Hz (low frequencies mostly contain non-vocal sounds). For each detection, we thus obtained a matrix of 12 columns (4 ms/column), and 118 rows (63 Hz/row) covering a frequency bandwidth of 630 to 8,694 Hz. From each cell of the matrix, we can infer the time (which column), the frequency (which line) and the sound amplitude (cell value), the latter being normalized between zero and one corresponding to the minimum and maximum sound amplitude of the matrix in order to compensate for variability caused by the bird changing position toward the microphone. We computed the average reference spectrogram of the target syllable by computing the mean of all matrices obtained for each detection of the target syllable.

After deafening, our custom-made program computed a dissimilarity score between the normalized average reference spectrogram of a 48 ms (12 columns of 4 ms and 118 rows of 63 Hz) part of the target syllable and the normalized spectrogram (computed as before) of each rendition of the target syllable sang by the bird. This score, that we call “Spectrogram divergence score,” corresponds to the Euclidian distance computed on a fixed duration window of the target syllable, the score thus provides a measure of syllable dissimilarity. At most 12 milliseconds following syllable dissimilarity measurement, we provided the visual substitution that consisted in a transient light extinction (light-off, LO) with a duration in the range of 100 to 500 ms in the housing chamber of the bird either in contingency to the spectrogram divergence score, i.e., when the score was lower than a manually set threshold for the deaf LO birds or at times not-contingent with the score for the deaf random LO birds. Deaf control and hearing birds were never exposed to transient light extinction during the day. The task started when a bird was producing at least 400 song motifs per day on at least 3 consecutive days post-deafening, which on average was after 8.3 days. Birds were then involved in the task during 6 consecutive weeks.

Song analyses

Song data were processed offline using a custom program written with Matlab (v2023b). We performed a manual clean-up of the cluster containing the automatically detected part of the target syllables (window of 48 ms) in order to remove any false positives from the dataset of each recording day. We then computed the spectrogram divergence score as described before between the normalized spectrogram of each rendition of the target syllable part and the normalized baseline reference spectrogram.

We carried out a full clustering of all song syllables recorded from 1 day per week per bird, so we analyzed 7 days per bird at maximum, i.e., one baseline day and 6 days post-deafening. To do so, song syllables were split based on threshold crossing of the root mean square (RMS) sound waveform, where the threshold was adjusted from one bird to another but kept constant for a given bird for all days analyzed. Individual song syllables were manually sorted in different clusters according to their spectrographic similarity and position within the song motif. We then randomly extracted a sample of 200 renditions (maximum) of each syllable per week of experiment and per bird. From each song syllable rendition, including the entire target syllable, we computed the spectrogram divergence score as described above, but on the entire syllable (e.g., not as previously on a window of 48 ms), using our custom-made algorithms. We also used the Sound Analysis Pro program (Tchernichovski et al., 2000) to extract syllable duration, mean entropy and entropy variance. Mean entropy and entropy variance were computed as mean over the entire syllable. Note that because of a loss of data (hard drive failure) of the hearing birds, for this group of birds we were only able to compute the spectrogram divergence scores of the 48-ms target syllable. For graphical representations at the population level, data of each feature were normalized using z-score values z:

$$z = \frac{x - \mu_{baseline}}{\sigma_{baseline}}$$

where x is the value of the feature of interest for a specific syllable rendition, μ and σ are the mean and the standard deviation values, respectively, of the same feature computed on all syllable renditions on the baseline day.

Statistics

To evaluate whether the visual substitution task allowed deafened birds to maintain the target song syllable, we first computed the spectrogram divergence score of the 48-ms subpart of the target syllable and applied a linear mixed-effect model (LME) defined as:

$$y = lme(\text{fixed}=\text{SDE} \sim \text{Group} * \text{Day}, \text{random}=\sim 1 | \text{Bird})$$

Where the spectrogram divergence score (SDE) is the fixed variable, the factors *Group* and *Day*, respectively, account for the experimental groups (deaf LO, deaf no LO, deaf random LO and hearing birds) and days of experiment (from -5 to 35 days relative to the onset of the LO protocol). Data are nested for each bird.

In deafened birds, to determine whether the task impacted other acoustic features and other song syllables, for 1 day per week of experiment per bird, we extracted 200 renditions of each song syllable. For each acoustic feature (spectrogram divergence score, duration, mean entropy and entropy variance), we computed linear mixed-effect models on the raw measurements defined as:

$$y = lme(\text{fixed}=\text{Feature} \sim \text{Group} * \text{Week}, \text{random}=\sim 1 | \text{Bird})$$

Where *Feature* corresponds to one of the acoustic features measured, and the factors *Group* and *Week*, respectively, accounts for the experimental groups (deaf LO, deaf no LO and deaf random LO) and weeks of experiment (from -1 for pre-deafening, to 6 weeks post onset of the task). Data are nested for each bird.

For each LME, we computed an ANOVA on the output of the model followed, when appropriate, by *post-hoc* Tukey tests between significant factors.

All statistical tests were done using custom-written scripts in Matlab (v2023b) or R v4.4.1.

Results

The spectrogram divergence score: a new measure of song stability

To examine whether deafened zebra finches are able to exploit a visual signal to compensate the post-deafening song degradation, we used a method based on the comparison of spectrograms, the spectrogram of the currently produced target syllable and a ‘reference spectrogram’ of this syllable. The reliable reference spectrogram was computed from all syllable renditions ($>1,000$) of the target syllable produced on a baseline day.

At first, we ensured that this method allowed us to evaluate to what extent a syllable structure can be altered. To this end, we created samples of birdsongs that included various amount of acoustic distortion. They were built from songs of 19 hearing adult male zebra finches. We randomly extracted 200 renditions of an individual’s song motif produced in a single day and added a white noise (WN) of an intensity between 0 to 100% (step: 10%) such that the maximum sound amplitude never exceeded the one in the original song (Figure 1B). Spectrogram divergence scores were calculated from the comparison between the reference syllable spectrogram and the noisy syllable spectrograms. Score values appear to depend on the percentage of white noise intensity, increasing as intensity increases up to 40%, with little overlap between distributions of scores obtained for a given percentage and the slightly higher percentage, e.g., between 10 and 20% of WN in the signal. Beyond 50% of WN intensity, score values reached a plateau at a value of ~ 0.2 (Figures 1Ci–Di). We also performed comparisons from entire song motifs, e.g., song motif 1 with 0% of WN vs. song motif 1 with 100% of WN, to measure the impact of white noise overlay. Spectrogram divergence scores for the entire motif also asymptotically increased as the WN intensity increased (Figures 1Cii–Dii). Importantly, they were similar to score values based on the analysis of the target syllable structure indicating that the degree of alteration of a selected target syllable may provide information of the distortion at the level of the entire song.

Other possible measures of the degree of acoustic similarity between song motifs are the dissimilarity and the entropy scores provided by Sound Analysis Pro (SAP, Tchernichovski et al., 2000). The entropy score provides a measure of the randomness of the difference between acoustic signals. Computing the two scores from the samples of song motifs also resulted in quite similar positive and horizontal asymptotic curves (Figures 1Ciii–Diii, Civ–Div). We can however note the great variability in SAP dissimilarity scores obtained by song motifs masked by WN at a given intensity, especially when the WN percentage is 20 or 30% and, consequently, the overlap between

score distributions of song motifs masked by different intensities of WN. Also, the percentage of errors of computation using SAP increased with the level of WN added to the song motif while it remained zero using the spectrogram divergence scores (bar plots in green in Figures 1Ciii–Diii). In spite of these differences, analyses based on spectrogram divergence score or SAP measures provided a similar picture of changes in song structure due to the addition of various levels of WN. This led us to use the spectrogram divergence score to measure the degree of dissimilarity between syllables or song motifs and to evaluate the degree of deafening-induced song degradation. Moreover, the spectrogram divergence score relies on a very rapid computation (<4 ms), allowing us to deliver immediate transient light-off extinction in a syllable-contingent manner.

Impact of the light-off exposure on the spectrogram divergence score in deafened birds

A previous experiment (Zai et al., 2020) revealed that deaf birds can exploit a visual signal to adapt the pitch of a song syllable. Here, we wondered whether deaf birds could take one step further in using this visual signal to offset the deafening-induced degradation of a given song syllable. We recorded the songs of 40 birds over a minimum of 7 consecutive weeks. A total of 33 birds were deafened after about 1 week of baseline song recording (Figure 2A). Birds were trained according to an operant conditioning paradigm based on a computer program that detects a specific target song syllable, usually a harmonic-like syllable (Canopoli et al., 2014; Zai et al., 2020). Deafened birds were separated in three independent groups: deaf LO ($n = 13$), deaf no LO ($n = 15$), and deaf random LO ($n = 5$). Deaf LO birds were trained by briefly switching off the light in the sound-isolation chamber whenever the spectrogram divergence score of the targeted syllable was below a threshold. The threshold for triggering the light extinction was adjusted on a daily basis to the median of the previous day. Deaf random LO birds were exposed to a similar amount of transient light extinction than deaf LO birds, but the visual signal did not depend on the syllable spectrogram divergence score. Deaf no LO and hearing birds were not exposed to transient light extinction during the entire course of the experiment.

At first, we visually inspected spectrograms from songs recorded before and after deafening. As previously described, we observed a loss of song motif stereotypy occurring progressively over the six consecutive weeks of song recording. Figure 2 shows spectrograms of song examples from three deaf birds, one per experimental group, before deafening and during the 6 weeks of the behavioral task. Importantly, these three example birds (a deaf no LO bird: 11C, a deaf LO bird: 11E and a deaf random LO bird: 11R) belonged to the same clutch (clutch 11), so they were the same age and exposed to the same father’s song as juveniles. Deafening these birds resulted in a reduction in the degree of similarity between their songs, as indicated by the SAP similarity score computed from a subset of 200 songs recorded before and 6 weeks after the deafening procedure (see Table 1 for details). We performed a UMAP analysis on the song recording datasets, selecting 1 day of the pre-deafening baseline recording period and 1 day of the sixth post-deafening week. UMAP graphs offer the advantage to represent high-dimensional data in just a few dimensions, e.g., two in Figure 2D. Qualitatively, most of the clusters from songs

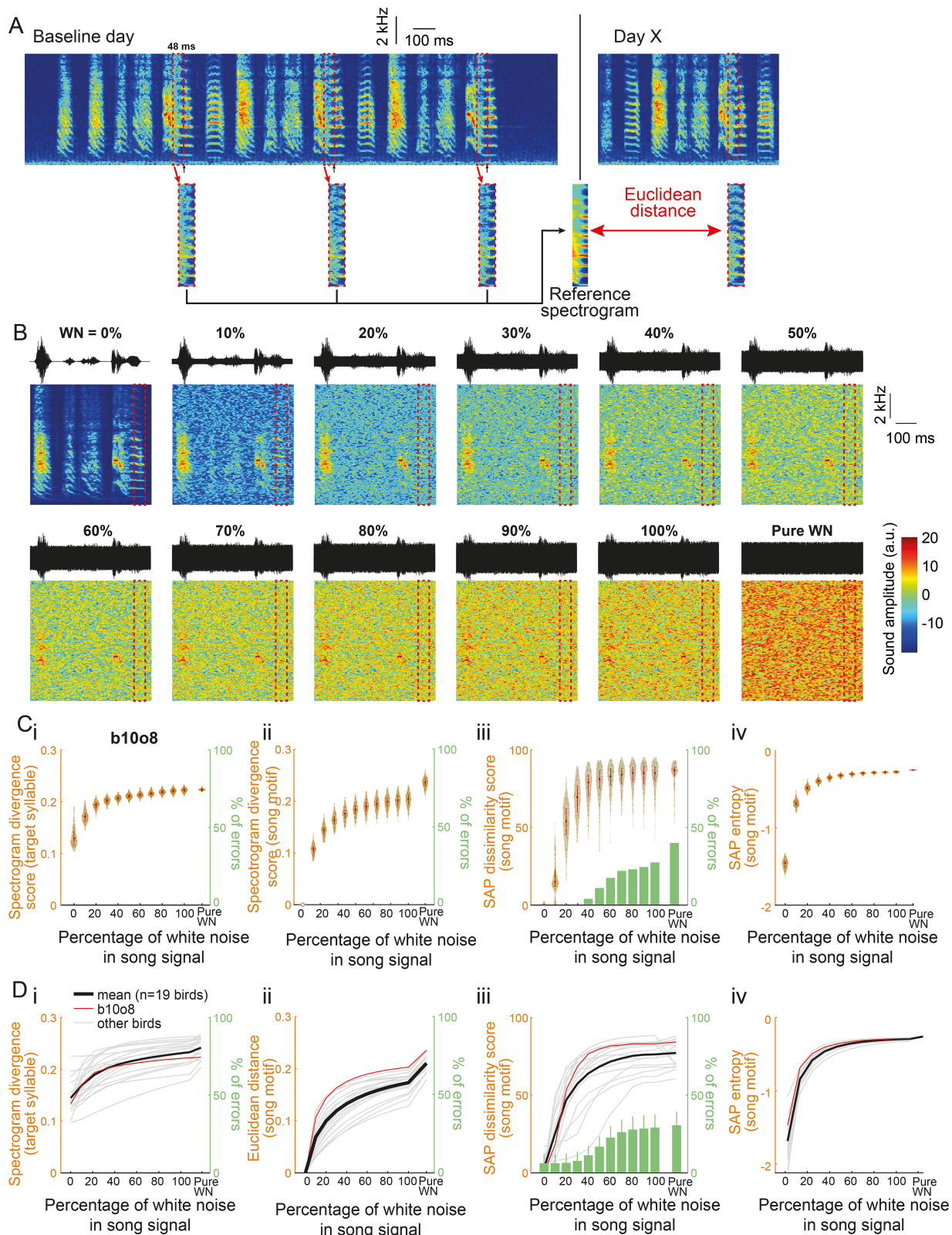


FIGURE 1

A new feature to assess sound degradation. (A) Spectrogram of three consecutive motifs produced by a single bird on a baseline day. An algorithm is trained to detect a song syllable (black arrows) and to extract its spectrogram from the 48 ms window before the detection point (dashed red box). The median of all 48-ms syllable excerpts provided the "Reference spectrogram." On the following days, the Euclidean distance was computed between the spectrograms of the targeted syllable and the Reference spectrogram to obtain the Spectrogram divergence score. (B) Waveforms (top) and spectrograms (bottom) of the same song motif from a single bird (labeled b10o8) with increasing levels of white noise (WN) added to the signal and of pure white noise. The dashed red box highlights this bird's 48-ms syllable excerpt used to assess the spectrogram divergence score. (C) Violin plot

(Continued)

FIGURE 1 (Continued)

showing the spectrogram divergence score (left y axis) computed between 200 randomly selected renditions of the 48-ms syllable excerpt for bird b10o8 with increasing levels of white noise in the signal (see B for spectrogram representation) and its reference spectrogram computed in the absence of white noise. White dots on the violin plots correspond to the mean spectrogram divergence scores. Note that, whatever the level of WN, there were no errors in the computation of the spectrogram divergence score (right y axis). (ii) Euclidean distances between 200 randomly selected song motifs with the same motifs that include increasing levels of white noise in the signal or with pure white noise (left y axis). There is again no error in the computation of the score (right y axis). (iii) Violin plots showing the dissimilarity score computed using Sound Analysis Pro (SAP; formula: $100 - \text{similarity score}$) between 200 randomly selected song motifs with the same motifs that include increasing levels of white noise in the signal or with pure white noise (left y axis). The percentage of errors of computation (when SAP failed in computing the similarity score between two songs) increases with the level of white noise added to the signal (green bar plot, right y axis). (iv) Violin plot of the entropy between 200 randomly selected song motifs with the same motifs that include increasing levels of white noise in the signal or with pure white noise. (D) Spectrogram divergence scores (i), Euclidean distance (ii), SAP dissimilarity score (iii) and SAP entropy (iv) computed as in (C) for 19 birds. Mean is shown with a thick black line and individual traces are in grey. Example bird from (A–C) (b10o8) is shown in red.

produced by the birds before and 6 weeks after the deafening procedure showed a low degree of overlap, so reflecting the overall modifications of the individual vocal repertoire. In comparison, the UMAP analysis performed in an example hearing bird shows that the vocal repertoire remained more stable than in deafened birds over a similar period of time (Figure 2D).

Although visually informative, the UMAP analysis told nothing about changes in song structure. We first assessed whether the exposure to the behavioral training conditions differentially impacted the spectrogram divergence scores of the sibling of three birds (Figure 3A). The spectrogram divergence score increased over days for the example deaf no LO bird (11C bird: +23% between the baseline period and days 25 to 30 post start of the LO protocol) while it remains rather stable for the two examples of deaf LO (11E bird: +14.5%) and deaf random LO (11R bird: +12.3%) birds, even though the number of renditions of the target syllable sang by the example deaf LO bird decreased more rapidly than for his control sibling. At the population level, the spectrogram divergence scores of the 48-ms subpart of the target syllable remained stable for hearing birds while it regularly increased for deafened birds of the three groups (Figure 3B). Note that even if the LO protocol lasted 6 consecutive weeks, we restricted the analysis at the population level to 4 consecutive weeks in order to include all the birds in the analysis.

Indeed, over time, the target syllable was less and less detected in deafened birds. The difference between hearing and deafened birds increased across days revealing a cumulative effect and thus a gradual degradation of the overall accuracy of the 48-ms target song syllable in all deafened birds. We computed the spectrogram divergence scores on all detected renditions of the 48-ms target syllable and normalized (Z-scored) the data to the last day before the onset of the LO protocol. A linear mixed effect model (LME) on the spectrogram divergence scores calculated over 35 consecutive days, including 5 days of baseline, with the birds as a nesting factor, revealed significant changes over days ($F_{34,1,066} = 22.18$, $p < 0.001$) and significant differences between groups ($F_{3,36} = 3.30$, $p < 0.04$). A post-hoc analysis (Tukey) on the group effect revealed greater changes in the syllable structure in deafened birds which were not exposed to the LO procedure compared to hearing birds (Deaf no LO vs. hearing birds; $t\text{-test}_{36} = 3.08$, $p < 0.02$), and compared to deafened birds exposed to the LO procedure (Deaf LO vs. Deaf no-LO birds; $t\text{-test}_{36} = 2.75$, $p < 0.05$). Comparisons based on scores computed on the target syllable produced by birds of the random LO group, which included only five birds, did not reveal any difference with the other groups. These results thus highlight a partial maintenance, or at least a slower

deafening-related degradation of the 48-ms target syllables due to the exposure to the visual signal.

Entropy and entropy variance were previously used to evaluate post-deafening song degradation (Horita et al., 2008). In order to assess whether the effect of the LO exposure on syllable structure in deaf birds was also observed using entropy and entropy variance as measures, we randomly extracted 200 renditions of the target syllable (48 ms) per week and per deaf bird (Figures 3C,D). A LME computed on these two measures revealed significant changes of both entropy and entropy variance over weeks (entropy: $F_{4,115} = 9.99$, $p < 0.001$; entropy variance: $F_{4,115} = 10.83$, $p < 0.001$). The target syllable entropy did not depend on the use of light extinction (no group effect: $F_{2,30} = 0.09$, $p = 0.91$; no group*week interaction: $F_{8,115} = 0.51$, $p = 0.85$). In contrast, the entropy variance was affected by the behavioral protocol (group*week interaction: $F_{8,115} = 2.10$, $p = 0.04$), with a significant difference between birds contingently exposed to the light extinction compared to those not exposed to the visual signal (Deaf LO vs. Deaf no LO; $t\text{-test}_{36} = 2.64$, $p = 0.03$), providing an additional evidence that exposure to the visual signal affected the deafening-induced degradation of the target syllable, at least the 48 ms period used as reference to deliver the visual signal. The entropy variance of the syllable target of the five birds of the group randomly exposed to the LO procedure did not differ from that of the two other groups of deaf birds.

In a previous study, the exposure to a pitch-contingent LO protocol was found to affect the singing activity of deaf birds (Zai et al., 2020). The pitch-contingent delivery of a transient light extinction led deafened birds to show a high motivation to sing (Zai et al., 2020). We examined here whether the singing rate varied depending on the contingency of the LO protocol. We estimated the number of motifs sung by individuals on a daily basis from the number of renditions of the target syllable, assuming that the birds always included the target song syllable within each song motif. As shown by the three examples of deaf birds (blue line in Figure 3A), birds sang less and less over the 6 weeks of song recording (Figure 3E). A LME computed on the number of target syllables automatically detected, with the birds as a nesting factor, revealed significant group effect ($F_{3,36} = 5.29$, $p < 0.005$), day effect ($F_{34,1,066} = 12.80$, $p < 0.001$) and day*group interaction ($F_{102,1,066} = 1.77$, $p < 0.001$). The number of detected target syllables showed a gradual decrease over days in the three groups of deafened birds but not in the group of hearing birds (post-hoc Tukey test; deaf LO vs. hearing birds: $t\text{-test}_{36} = 3.51$, $p < 0.007$; deaf no LO vs. hearing birds: $t\text{-test}_{36} = 4.09$, $p < 0.002$; deaf random LO vs. hearing birds: $t\text{-test}_{36} = 3.02$, $p < 0.03$) but with no significant differences between groups of the deafened birds. The

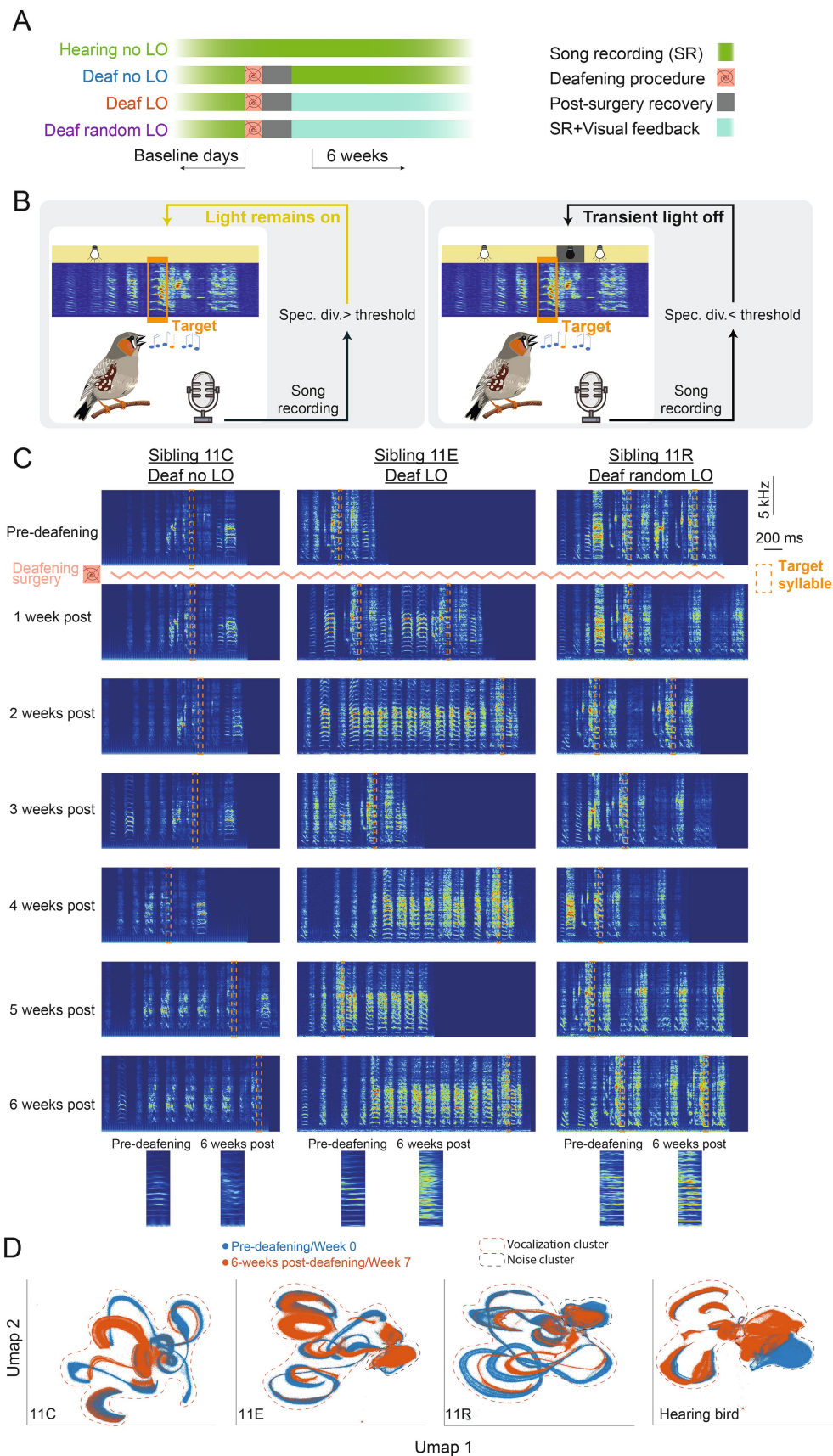


FIGURE 2

A behavioral task to counteract deafening induced song degradation. (A) Birds were divided into four groups. All birds were isolated in a sound-proof chamber and their song was recorded on a few consecutive baseline days. Except the hearing birds ($n = 7$ birds), all the other birds were deafened

(Continued)

FIGURE 2 (Continued)

following the procedure described in (Zai et al., 2020). After the deafening procedure, birds recovered for a few days before the onset of the behavioral paradigm. (B) During the task, songs were recorded online and a target syllable specific for each bird was automatically detected. For deaf LO birds ($n = 13$ birds), whenever the spectrogram divergence score was lower than a certain threshold, the housing light in the sound-proof chamber was transiently switched off for 200 ms, the light remained on otherwise. For deaf no LO birds ($n = 15$ birds), the light was never switched off while for deaf random LO birds ($n = 5$ birds), the light was transiently but randomly switched off when the bird produced the target syllable. (C) Spectrograms of song motifs produced by three birds from the same clutch before and 1 to 6 weeks post-deafening and onset of the LO protocol. The target syllable for each bird is shown with a dashed orange box and zoomed in versions produced pre-deafening and 6-weeks post-deafening are highlighted. (D) UMAP projections of all sounds recorded (i.e., birds vocalizations and cage noise) in the sound proof chamber for each example deafened bird and for a hearing control bird before deafening (blue) and 6 weeks post-deafening and onset of LO protocol for deafened birds, or 7 weeks later for the hearing bird. Clusters that include bird vocalizations (songs and calls) and noise are surrounded by dashed red and blue lines, respectively.

TABLE 1 Sound Analysis Pro similarity score (Tchernichovski et al., 2000) of 3 siblings from the same clutch pre-deafening vs. 6-weeks post-deafening.

Birds	SAP similarity score (mean %±SD) pre-deafening	SAP similarity score (mean %±SD) 6-weeks post-deafening
11C (deaf no LO) vs. 11E (deaf LO)	33.2±11.1	29.3±6.9
11C vs. 11R (deaf random LO)	66.3±18.7	42.0±9.8
11R vs. 11E	60.9±9.6	48.5±9.8

singing activity showed thus a similar gradual decrease in the three groups of deafened birds, so independently of the contingency of the LO protocol with the spectrogram divergence score (Figure 3E).

The spectrogram divergence score was computed on a 48-ms window of the target syllable. The syllable duration was longer leaving unclear whether the slower degradation of the target syllable structure exhibited by deaf LO birds was limited to the 48 ms window or whether it extended to the entire target syllable, and beyond, to the entire song. We computed offline the spectrogram divergence scores from the entire target syllable and all other song syllables (“non-target syllables”). We established a reference syllable spectrogram for each syllable type found in the individual’s song motif, from 200 randomly selected renditions of each syllable during a baseline day. We also computed spectrogram divergence scores on randomly selected 200 syllable renditions from 1 day per week over the 4 weeks of the post-deafening period. As shown in Figure 4A, spectrogram divergence scores of the entire target syllables as well as other syllables increased over weeks (week effect; target syllable: $F_{4,120} = 77.94, p < 0.001$; non-target syllables: $F_{4,509} = 120.70, p < 0.001$).

When we compared the extent of changes in spectrogram divergence scores over time between the three groups of deafened birds, the target syllable as well as the non-target syllables did not reveal any effect of the behavioral protocols (target syllable, group effect: $F_{2,30} = 0.95, p = 0.40$; group*week interaction: $F_{8,120} = 1.55, p = 0.15$; non-target syllable, group effect: $F_{2,29} = 1.14, p = 0.33$; group*week interaction: $F_{8,509} = 0.42, p = 0.91$). Spectrogram divergence scores therefore suggest that the impact of the exposure to the light off procedure did not extend beyond the 48 ms period used to compute the threshold value for triggering the visual signal.

We also analyzed whether the duration, the entropy, and the entropy variance of the syllables changed over weeks, depending on the behavioral protocol. We used the same dataset of 200 randomly selected renditions of each syllable type per week and per bird. The duration of the target syllable was rather stable over weeks, at least the first 3 weeks ($F_{4,117} = 2.01, p < 0.10$; Figure 4B), while the duration of non-target syllables decreased ($F_{4,521} = 37.83, p < 0.001$) with no difference in the course of changes between the three groups of deafened birds (target syllable: $F_{2,30} = 0.11, p = 0.89$; non-target

syllables: $F_{2,29} = 0.62, p = 0.54$). The entropy of target and non-target syllables increased over weeks (target syllable: $F_{4,117} = 15.10, p < 0.001$; non-target syllables: $F_{4,521} = 11.70, p < 0.001$; Figure 4C) while entropy variance regularly decreased (target syllable: $F_{4,117} = 5.42, p < 0.001$; non-target syllables: $F_{4,521} = 17.48, p < 0.001$; Figure 4D), indicating that song syllables gradually became uniformly noisier. The course of changes in the entropy of the target syllable did not reveal any difference between the three groups of deafened birds (group effect: $F_{2,30} = 1.21, p = 0.31$; group*week interaction: $F_{8,117} = 1.82, p = 0.32$). In contrast, the course of the entropy of other song syllables varied depending on the behavioral protocol used (group effect: $F_{2,29} = 4.18, p < 0.03$). The entropy differed significantly between deaf no LO and deaf random LO birds (post-hoc Tukey test, deaf no LO vs. random LO birds: $t_{29} = 2.83, p < 0.03$), with no difference between deaf LO birds and the two other groups.

Only the entropy revealed a difference depending on the behavioral protocol used. The entropy variance of both the entire target syllable and the set of other syllables decreased over weeks, with no difference in the time course between groups (target syllable; group effect: $F_{2,30} = 0.63, p = 0.54$; group*week interaction: $F_{8,117} = 0.45, p = 0.88$; other syllables; group effect: $F_{2,29} = 0.18, p = 0.83$; group*week interaction: $F_{8,521} = 0.49, p = 0.86$).

Discussion

In both humans and songbirds, loss of auditory feedback in adulthood leads to progressive loss of precise vocal control (Waldstein, 1990; Lane and Webster, 1991; Nordeen and Nordeen, 1992; Lombardino and Nottebohm, 2000; Horita et al., 2008; Tschida and Mooney, 2012). This observation led to the hypothesis that auditory feedback is necessary for the maintenance of speech production. However, previous studies showed that both deafened and hearing birds can adjust the pitch of a selected song syllable using non-auditory feedback signals, including a light extinction or a cutaneous stimulation contingent to the pitch of the target syllable (Zai et al., 2020; McGregor et al., 2022). Here, we provide evidence that deafened birds can also use light-off signals to slowdown deafness-induced song

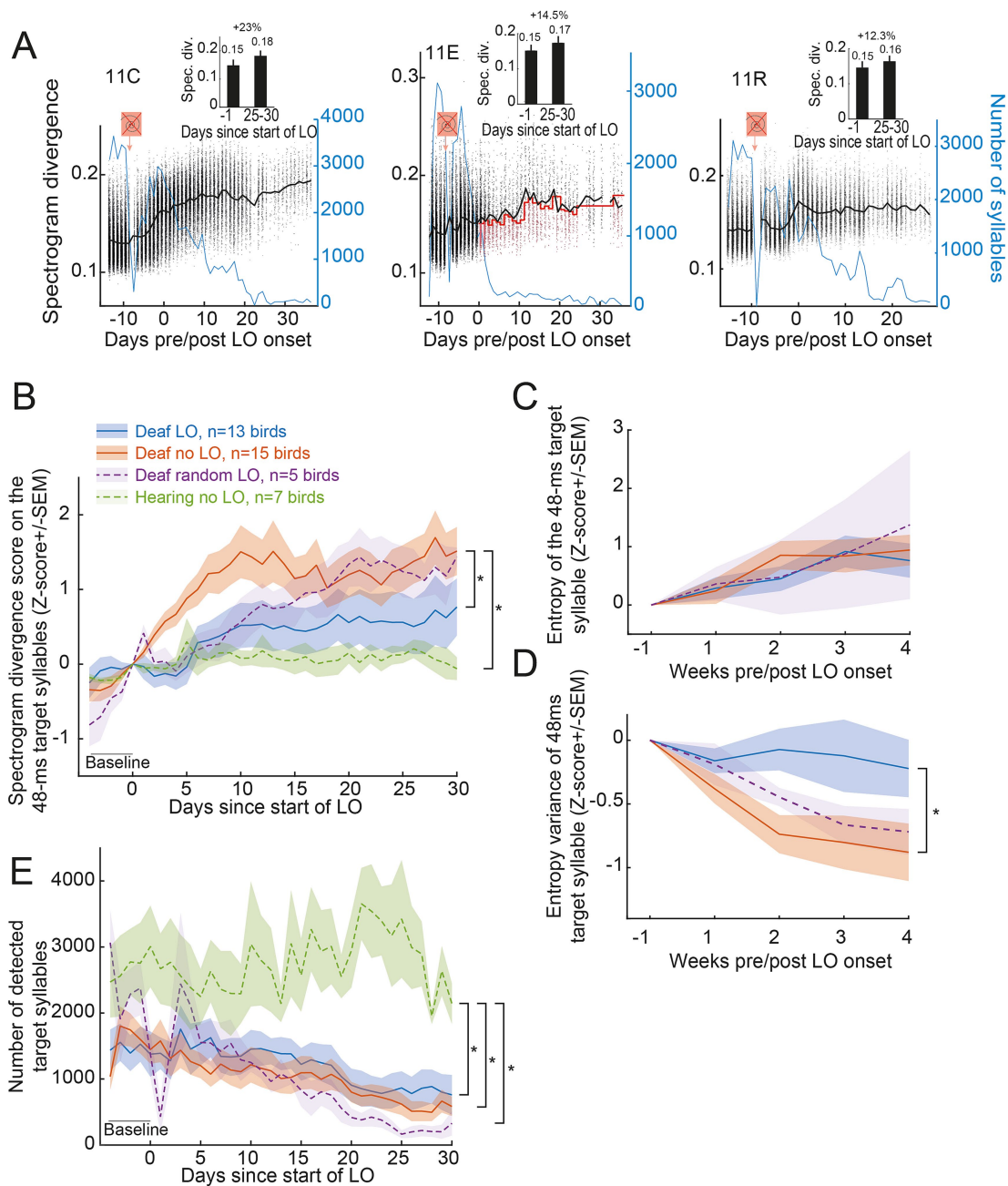


FIGURE 3

Transient light extinction contingent on the spectrogram divergence score slows syllable degradation in deaf birds. (A) Variations in spectrogram divergence scores (black dots, left y axis) and the number of detected target syllables (blue line, right y axis) over days for three example deafened birds (same birds as in Figure 2): the deafened 11C no LO (left), the deafened 11E LO (middle) and the deafened 11R random LO bird (right). Black line indicates the mean value of the spectrogram divergence score with the dots representing the raw values calculated for each detected target syllable. Red line for the deafened 11E LO bird indicates the threshold value used for triggering the transient light extinction. The symbol in orange indicates when the deafening procedure occurred. The LO exposure started on day zero, with the same day used for both the 11C and the 11E birds. Inserts show the mean (\pm SEM) spectrogram divergence score over the last day before the onset of the LO protocol and days 25–30 after the onset of the LO protocol. Shown above is the percent of change. (B) Normalized spectrogram divergence scores (mean Z-score \pm SEM) computed on all detected renditions of the 48-ms target syllable for all birds before and during the LO protocol. (C,D) Entropy (C) and entropy variance (D) computed on a subset of 200 randomly selected renditions per week per bird of the 48-ms target syllable. (E) The number of target syllables automatically detected decreased for all deafened birds. *, significant group effect, $p < 0.05$ (see text for details).

degradation of a song syllable. A deterioration in the song structure was exhibited by all deafened birds. The deterioration of the structure of a given syllable, measured thanks to the spectrogram divergence score, however, appeared to be reduced in deafened birds exposed to the light off procedure in comparison to deafened birds that did not experienced visual signals. The present study therefore suggest that

visual information could, to some extent, be used by deafened birds to control the structure of a given song syllable.

To our knowledge, compensating for deafening-induced song degradation after deafness had never been attempted. However, deafening-induced song degradation includes a wide range of spectro-temporal changes in the structure of song syllables (Nordeen and

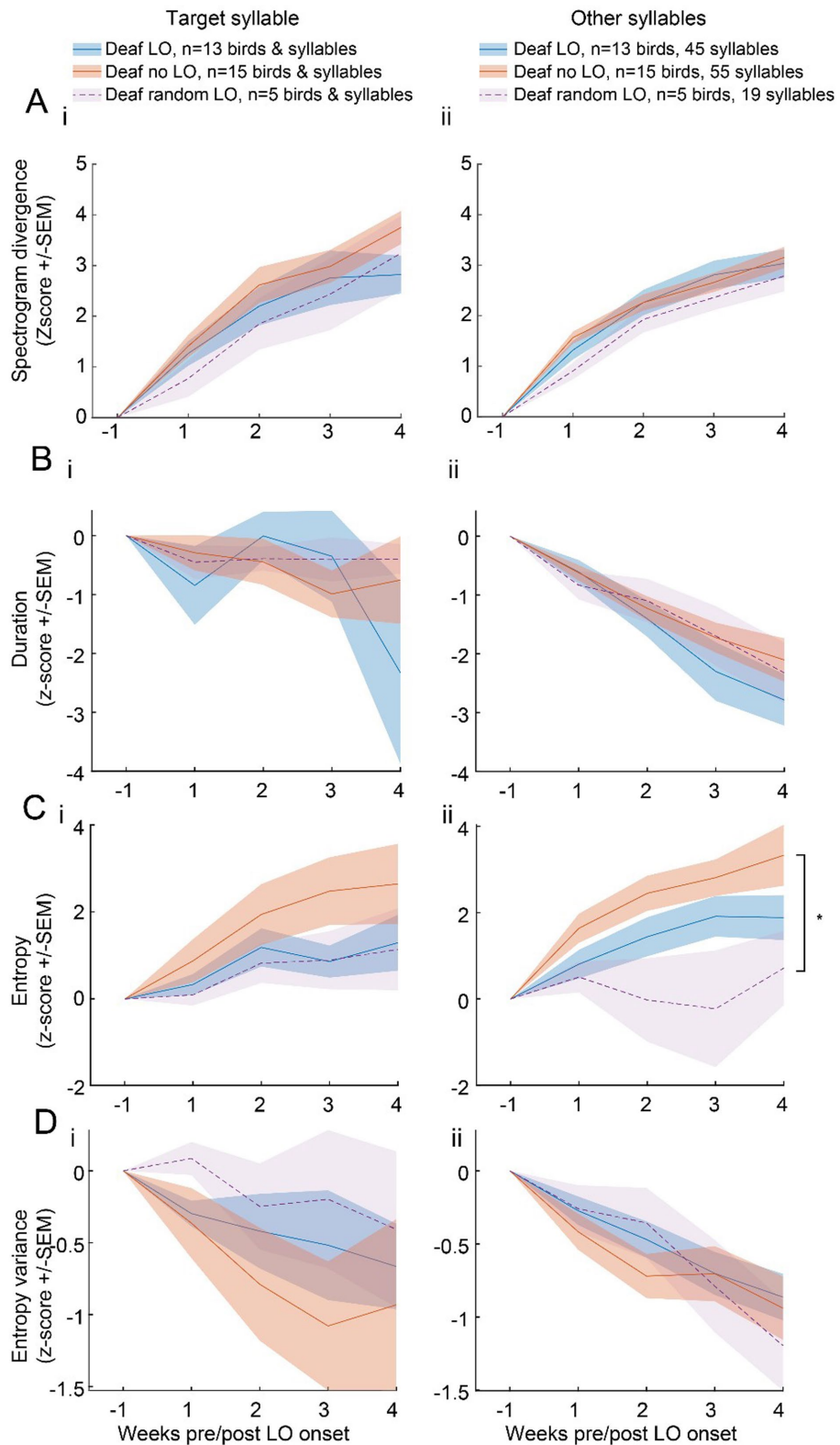


FIGURE 4

Various acoustic features show no clear evidence of an impact of a LO exposure on the structure of the entire target syllables or of other song syllables. Acoustic measures were computed on subsets of 200 randomly selected target (left panel) and non-target (right panel) syllables per bird and per week. Spectrogram divergence score (A), syllable duration (B), entropy (C) and entropy variance (D) of the entire target (i) or non-target (ii) syllables over the 6 weeks after the onset of the LO protocol for deaf LO birds. Spectrogram divergence scores evolved within a similar range for the three groups of deafened birds, deaf LO (blue), deaf no LO (red) and deaf random LO (purple). *, significant group effect, $p < 0.05$ (see text for details).

Nordeen, 1992; Wang et al., 1999; Brainard and Doupe, 2001; Horita et al., 2008). Also, prior to the present study, no acoustic measurement of the entire syllable structure had been carried out in real-time to be useful in a behavioral task. In previous studies, song syllable degradation was quantified either visually by the observation of spectrograms, or by the quantification of some acoustic features, including mean entropy and entropy variance (Horita et al., 2008; Pytte et al., 2012; Tschida and Mooney, 2012). These measures were, however, performed offline. Therefore, training birds in a behavioral paradigm to enable them to maintain potentially the overall structure of one of their song syllables by specifically targeting it, required to be innovative. In order to quantify deafness-induced song degradation more globally, i.e., on the basis of multiple acoustic features, and very quickly, we developed a new method based on the computation of the spectrogram divergence score. This measure can be seen as rapidly reflecting the overall stability of a song syllable and post-deafening song changes.

The spectrogram divergence score values remained stable in hearing birds over weeks. Deafening induced striking changes in song and syllable structure and song spectrograms gradually became less similar to the song produced before. Consistently, spectrogram divergence scores in deafened birds that were not exposed to the light off extinction protocol gradually increased over weeks, in a range of values that significantly differed from score values computed from spectrograms of hearing birds. Also, this measure allowed us to capture fine changes in syllable structure since spectrogram divergence score values revealed a deafening-induced impact from the first weeks. The spectrogram divergence score that has the advantage of reflecting the global structure of vocal signals in a way that is fast enough to be used in a real-time behavioral task, therefore proved to be a reliable and effective measure of song degradation. However, the measure could still be improved since it seemed to be harder for birds to detect it than pitch, given that deafened birds did not increase their singing rate upon LO exposure contingent on the spectrogram divergence score, contrary to deafened birds exposed to LO contingent on song syllable pitch (Zai et al., 2020). This method could considerably contribute to quantify short and long-term degradation of syllable structure in future studies and could therefore have relevance to the field of vocal plasticity.

On the basis of both the spectrogram divergence score and entropy variance, deafened birds exposed to the transient light extinction contingently to the spectrogram divergence score retained to a certain extent the overall structure of the target syllable. Both the use of a visual signal and the contingency of its delivery might be crucial in enabling the birds to control, at least in part, the structure of the target syllable. Deafened birds that were not exposed to light extinction exhibited a faster and a more severe degradation of the target syllable structure. The amount of degradation of the target syllable appeared to be moderate in birds randomly exposed to the visual signal. It remains, however, to be examined to what extent the contingency of the visual cue delivery guides the control of the target syllable structure. No clear difference in the severity of the degradation allows distinguishing birds contingently exposed to light extinction from birds exposed randomly. The present study, nevertheless, provides evidence of a weaker impact of the deafness when a visual cue was delivered. This suggests that deaf birds are able to exploit a transient light extinction signal not only to adjust the pitch of a target syllable (Zai et al., 2020), but also to maintain at least partially its spectro-temporal structure. Consequently, even though the auditory feedback is critical to provide evaluation of what has been vocally

produced, information of another sensory modality can influence the song production.

The impact reported here on syllable structure maintenance appears not to extend beyond the 48-ms period used to calculate the dissimilarity score between the reference spectrogram and the spectrogram of the new rendition of the syllable target. No difference was found in spectrogram divergence scores computed from either the entire target syllable or other non-targeted syllables between deaf birds exposed or not to the light-extinction. Syllable duration also did not exhibit any difference. Only entropy measures indicate a possible effect of the behavioral conditions. Indeed, entropy scores of the non-target syllables remained more stable for deafened random LO than for deaf no LO birds. This result may suggest that perhaps the reliable temporal occurrence of the random LO causes birds to pay attention to their song syllables, and the lack of contingency with the target syllable leads them to focus on non-target syllables. This unexpected result should be taken with caution because of the limited number of syllables from the few deaf random LO birds (syllables from only 5 compared to 15 birds). The limited impact of the behavioral paradigm on the structure of the entire target syllable and on the other song syllables are consistent with previous studies investigating the impact of a pitch-shifting paradigm on the acoustic structure of the syllable target (Tumer and Brainard, 2007; Canopoli et al., 2014; Pehlevan et al., 2018; Zai et al., 2020). Even if birds, including deafened birds, are able to modify the fundamental frequency value of the target syllable in a time-period that extends beyond the time window used to calculate the pitch, changes only occur in a very narrow time-period around the pitch calculation time window (Canopoli et al., 2014; Zai et al., 2020). Other song syllables do not show any changes in their fundamental frequency or other spectro-temporal features (Tumer and Brainard, 2007). Additionally, the behavioral training task used in the present study, because it was based on the overall acoustic structure of the syllable and not on a single acoustic parameter, can be considered as probably more complex or difficult for the birds than the pitch-shifting behavioral task. The fundamental frequency may only be controlled by one set of muscles surrounding the syrinx (Larsen and Goller, 2002; Goller and Riede, 2013). To maintain all the acoustic parameters of a given syllable, it is likely that a greater number of muscles need to be controlled. It is also worth noting that the sensory feedback that was provided during the task is very simple since it only consisted in a transient extinction of the housing light tuned to a subpart of a syllable. It would be interesting to investigate whether birds would be able to use a richer visual feedback, for example with various colors tuned to several acoustic features of the song syllables.

Remaining to be investigated are the changes in neural networks and to identify specific neural plasticity mechanisms underlying the integration of visual signals used to replace auditory information in deaf birds. Given brain reorganizations that occur as a result of deafness (Pytte et al., 2010; Tschida and Mooney, 2012; Zhou et al., 2017), the behavioral training task we used here, could therefore have impacted, at least in part, these reorganizations. It would be interesting in order to study multisensory interactions that can take part in vocal production, to investigate how visual signals contingent on one acoustic feature, e.g., the fundamental frequency, of a song syllable affect brain circuitry. In 1985, a study showed that neurons in the sensorimotor HVC nucleus may exhibit changes in their activity in response to visual stimuli, suggesting that the song system receives visual information (Bischof and Engelage, 1985). We have shown previously that a lesion

of the songbird basal ganglia (Area X) abolishes the ability of deafened birds to adjust their song syllable pitch using a transient light extinction (Zai et al., 2020). In order to exploit transient light stimulus, birds probably use information that enables them to control motor output so that they can modify it. Information may be proprioceptive, but could also be based on an internal motor copy of the song command from motor centers to Area X (i.e., efferent copy, Giret et al., 2014). This copy, in conjunction to the release of reinforcing signals (e.g., dopamine) into Area X during the contingent sensory feedback (Gadagkar et al., 2016; Roeser et al., 2023), could provide, in parallel, information about the motor command in sensory regions. A future challenging question is how visual information in deafened birds affects song system processing to guide the vocal motor control and counteract the deafening-induced syllable degradation in absence of auditory feedback (Rolland et al., 2022). Tract tracing studies identified both direct and indirect projections from subpallial and pallial visual areas to auditory areas and to song related nuclei (Wild and Gaede, 2016; Stacho et al., 2020). Putative candidate visual areas for the integration of visual feedback include the optic tectum through its projection to an auditory midbrain area (mesencephalic lateralis nucleus, pars dorsalis, MLd), the rostral uvulaeformis nucleus (Uva) which indirectly projects to the HVC, or the Wulst which indirectly projects to the Area X (Wild, 1994; Shimizu and Bowers, 1999; Person et al., 2008; Mandelblat-Cerf et al., 2014; Woolley, 2019). Future works should investigate through which pathway information provided by non-auditory (visual) feedbacks reaches the song system nuclei and to what extent the vocal behavior of songbirds relies on visual information processing by these nuclei.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The animal study was approved by French Ministry of Research; Ethical committee Paris-Sud et Centre; Cantonal Veterinary Office of the Canton of Zurich, Switzerland. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

MR: Data curation, Formal analysis, Investigation, Writing – review & editing. ATZ: Conceptualization, Data curation, Formal

analysis, Investigation, Methodology, Software, Writing – review & editing. RHRH: Funding acquisition, Project administration, Software, Supervision, Writing – review & editing. CDN: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. NG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Centre National de la Recherche Scientifique and the University of Paris Saclay. M.R. was supported by the Ministère de la Recherche, de l'Enseignement supérieur et de l'Innovation.

Acknowledgments

We thank Sophie Cavé-Lopez for help in conducting the experiments. We thank Mélanie Dumont and Caroline Rousseau for taking care of the songbird facility.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Andalman, A. S., and Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl. Acad. Sci. USA* 106, 12518–12523. doi: 10.1073/pnas.0903214106
- Bischof, H.-J., and Engelage, J. (1985). Flash evoked responses in a song control nucleus of the zebra finch (*Taeniopygia guttata castanotis*). *Brain Res.* 326, 370–374. doi: 10.1016/0006-8993(85)90048-4
- Brainard, M. S., and Doupe, A. J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404, 762–766. doi: 10.1038/35008083
- Brainard, M. S., and Doupe, A. J. (2001). Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J. Neurosci.* 21, 2501–2517. doi: 10.1523/JNEUROSCI.21-07-02501.2001
- Brainard, M. S., and Doupe, A. J. (2002). What songbirds teach us about learning. *Nature* 417, 351–358. doi: 10.1038/417351a
- Canopoli, A., Herbst, J. A., and Hahnloser, R. H. R. (2014). A higher sensory brain region is involved in reversing reinforcement-induced vocal changes in a songbird. *J. Neurosci.* 34, 7018–7026. doi: 10.1523/JNEUROSCI.0266-14.2014

- Cowie, R., Douglas-Cowie, E., and Kerr, A. G. (1982). A study of speech deterioration in post-lingually deafened adults. *J. Laryngol. Otol.* 96, 101–112. doi: 10.1017/S002221510009229X
- Derryberry, E. P., Phillips, J. N., Derryberry, G. E., Blum, M. J., and Luther, D. (2020). Singing in a silent spring: birds respond to a half-century soundscape reversion during the COVID-19 shutdown. *Science* 370, 575–579. doi: 10.1126/science.abd5777
- Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common theme and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567
- Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., and Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science* 354, 1278–1282. doi: 10.1126/science.aah6837
- Giret, N., Kornfeld, J., Ganguli, S., and Hahnloser, R. H. R. (2014). Evidence for a causal inverse model in an avian cortico-basal ganglia circuit. *Proc. Natl. Acad. Sci. USA* 111, 6063–6068. doi: 10.1073/pnas.1317087111
- Goller, F., and Riede, T. (2013). Integrative physiology of fundamental frequency control in birds. *J. Physiol. Paris* 107, 230–242. doi: 10.1016/j.jphysparis.2012.11.001
- Hamaguchi, K., Tschida, K. A., Yoon, I., Donald, B. R., and Mooney, R. (2014). Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. *eLife* 3:e01833. doi: 10.7554/eLife.01833
- Herbst, J. A., Kotowicz, A., Wang, C. Z.-H., Rychen, J., Canopoli, A., Narula, G., et al. (2023). RecOOrder – a modular recording and control environment in LabVIEW. Available at: <https://www.research-collection.ethz.ch/handle/20.500.11850/617741> (Accessed May 24, 2024).
- Horita, H., Wada, K., and Jarvis, E. D. (2008). Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *Eur. J. Neurosci.* 28, 2519–2532. doi: 10.1111/j.1460-9568.2008.06535.x
- Iyengar, S., and Bottjer, S. W. (2002). The role of auditory experience in the formation of neural circuits underlying vocal learning in zebra finches. *J. Neurosci.* 22, 946–958. doi: 10.1523/JNEUROSCI.22-03-00946.2002
- Konishi, M. (1964). Effects of deafening on song development in two species of juncos. *Condor* 66, 85–102. doi: 10.2307/1365388
- Konishi, M. (1965a). Effects of deafening on song development in American robins and black-headed grosbeaks. *Z. Tierpsychol.* 22, 584–599.
- Konishi, M. (1965b). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* 22, 770–783.
- Konishi, M. (2004). The role of auditory feedback in birdsong. *Ann. N. Y. Acad. Sci.* 1016, 463–475. doi: 10.1196/annals.1298.010
- Lane, H., and Webster, J. W. (1991). Speech deterioration in postlingually deafened adults. *J. Acoust. Soc. Am.* 89, 859–866. doi: 10.1121/1.1894647
- Larsen, O. N., and Goller, F. (2002). Direct observation of syringeal muscle function in songbirds and a parrot. *J. Exp. Biol.* 205, 25–35. doi: 10.1242/jeb.205.1.25
- Leder, S. B., Spitzer, J. B., and Kirchner, J. C. (1987a). Speaking fundamental frequency of postlingually profoundly deaf adult men. *Ann. Otol. Rhinol. Laryngol.* 96, 322–324. doi: 10.1177/000348948709600316
- Leder, S. B., Spitzer, J. B., Kirchner, J. C., Flevaris-Phillips, C., Milner, P., and Richardson, F. (1987b). Speaking rate of adventitiously deaf male cochlear implant candidates. *J. Acoust. Soc. Am.* 82, 843–846. doi: 10.1121/1.395283
- Leder, S. B., Spitzer, J. B., Milner, P., Flevaris-Phillips, C., Kirchner, J. C., and Richardson, F. (1987c). Voice intensity of prospective cochlear implant candidates and normal hearing adult males. *Laryngoscope* 97, 224–227. doi: 10.1288/00005537-198702000-00017
- Leonardo, A., and Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* 399, 466–470. doi: 10.1038/20933
- Lombardino, A. J., and Nottebohm, F. (2000). Age at deafening affects the stability of learned song in adult male zebra finches. *J. Neurosci.* 20, 5054–5064. doi: 10.1523/JNEUROSCI.20-13-05054.2000
- Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M. S. (2014). A role for descending auditory cortical projections in songbird vocal learning. *eLife* 3:e02152. doi: 10.7554/eLife.02152
- McGregor, J. N., Grassler, A. L., Jaffe, P. I., Jacob, A. L., Brainard, M. S., and Sober, S. J. (2022). Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain. *eLife* 11:e75691. doi: 10.7554/eLife.75691
- Nordeen, K. W., and Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. Neural Biol.* 57, 58–66. doi: 10.1016/0163-1047(92)90757-U
- Nordeen, K. W., and Nordeen, E. J. (2010). Deafening-induced vocal deterioration in adult songbirds is reversed by disrupting a basal ganglia-forebrain circuit. *J. Neurosci.* 30, 7392–7400. doi: 10.1523/JNEUROSCI.6181-09.2010
- Okanoya, K., and 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. doi: 10.1002/(SICI)1097-4695(199710)33:4<343::AID-NEU1>3.0.CO;2-A
- Pehlevan, C., Ali, F., and Ölveczky, B. P. (2018). Flexibility in motor timing constrains the topology and dynamics of pattern generator circuits. *Nat. Commun.* 9:977. doi: 10.1038/s41467-018-03261-5
- Person, A. L., Gale, S. D., Farries, M. A., and Perkel, D. J. (2008). Organization of the songbird basal ganglia, including area X. *J. Comp. Neurol.* 508, 840–866. doi: 10.1002/cne.21699
- Pytte, C. L., George, S., Korman, S., David, E., Bogdan, D., and Kirn, J. R. (2012). Adult neurogenesis is associated with the maintenance of a stereotyped, learned motor behavior. *J. Neurosci.* 32, 7052–7057. doi: 10.1523/JNEUROSCI.5385-11.2012
- Pytte, C. L., Parent, C., Wildstein, S., Varghese, C., and Oberlander, S. (2010). Deafening decreases neuronal incorporation in the zebra finch caudomedial nidopallium (NCM). *Behav. Brain Res.* 211, 141–147. doi: 10.1016/j.bbr.2010.03.029
- Roeser, A., Gadagkar, V., Das, A., Puzerey, P. A., Kardon, B., and Goldberg, J. H. (2023). Dopaminergic error signals return to social feedback during courtship. *Nature* 623, 375–380. doi: 10.1038/s41586-023-06580-w
- Rolland, M., Del Negro, C., and Giret, N. (2022). Multisensory processes in birds: from single neurons to the influence of social interactions and sensory loss. *Neurosci. Biobehav. Rev.* 143:104942. doi: 10.1016/j.neubiorev.2022.104942
- Schwartzkopf, J. (1949). Über sitz und leistung von gehör und vibrationssinn bei vögeln. *Zeitschr. f. vergl. Physiol.* 31, 527–608. doi: 10.1007/BF00348361
- Shimizu, T., and Bowers, A. N. (1999). Visual circuits of the avian telencephalon: evolutionary implications. *Behav. Brain Res.* 98, 183–191. doi: 10.1016/S0166-4328(98)00083-7
- Sober, S. J., and Brainard, M. S. (2009). Adult birdsong is actively maintained by error correction. *Nat. Neurosci.* 12, 927–931. doi: 10.1038/nn.2336
- Stacho, M., Herold, C., Rook, N., Wagner, H., Axer, M., Amunts, K., et al. (2020). A cortex-like canonical circuit in the avian forebrain. *Science* 369:abc5534. doi: 10.1126/science.abc5534
- Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., and Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* 59, 1167–1176. doi: 10.1006/anbe.1999.1416
- Tschida, K. A., and Mooney, R. (2012). Deafening drives cell-type-specific changes to dendritic spines in a sensorimotor nucleus important to learned vocalizations. *Neuron* 73, 1028–1039. doi: 10.1016/j.neuron.2011.12.038
- Tumer, E. C., and Brainard, M. S. (2007). Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature* 450, 1240–1244. doi: 10.1038/nature06390
- Tyack, P. L. (2020). A taxonomy for vocal learning. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 375:20180406. doi: 10.1098/rstb.2018.0406
- Wildstein, R. S. (1990). Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J. Acoust. Soc. Am.* 88, 2099–2114. doi: 10.1121/1.400107
- Wang, N., Aviram, R., and Kirn, J. R. (1999). Deafening alters neuron turnover within the telencephalic motor pathway for song control in adult zebra finches. *J. Neurosci.* 19, 10554–10561. doi: 10.1523/JNEUROSCI.19-23-10554.1999
- Wild, J. M. (1994). Visual and somatosensory inputs to the avian song system via nucleus uvaefornis (Uva) and a comparison with the projections of a similar thalamic nucleus in a nonsongbird, *Columbia livia*. *J. Comp. Neurol.* 349, 512–535. doi: 10.1002/cne.903490403
- Wild, J. M., and Gaede, A. H. (2016). Second tectofugal pathway in a songbird (*Taeniopygia guttata*) revisited: Tectal and lateral pontine projections to the posterior thalamus, thence to the intermediate nidopallium. *J. Comp. Neurol.* 524, 963–985. doi: 10.1002/cne.23886
- Wittenbach, J. D., Bouchard, K. E., Brainard, M. S., and Jin, D. Z. (2015). An adapting auditory-motor feedback loop can contribute to generating vocal repetition. *PLoS Comput. Biol.* 11:e1004471. doi: 10.1371/journal.pcbi.1004471
- Woolley, S. C. (2019). Dopaminergic regulation of vocal-motor plasticity and performance. *Curr. Opin. Neurobiol.* 54, 127–133. doi: 10.1016/j.conb.2018.10.008
- Woolley, S. M. N., and Rubel, E. W. (1997). Bengalese finches *Lonchura striata* domestica depend upon auditory feedback for the maintenance of adult song. *J. Neurosci.* 17, 6380–6390. doi: 10.1523/JNEUROSCI.17-16-06380.1997
- Yamahachi, H., Zai, A. T., Tachibana, R. O., Stepien, A. E., Rodrigues, D. I., Cavé-Lopez, S., et al. (2020). Undirected singing rate as a non-invasive tool for welfare monitoring in isolated male zebra finches. *PLoS One* 15:e0236333. doi: 10.1371/journal.pone.0236333
- Zai, A. T., Cavé-Lopez, S., Rolland, M., Giret, N., and Hahnloser, R. H. R. (2020). Sensory substitution reveals a manipulation bias. *Nat. Commun.* 11:5940. doi: 10.1038/s41467-020-19686-w
- Zhang, Y., Zhou, L., Zuo, J., Wang, S., and Meng, W. (2023). Analogies of human speech and bird song: from vocal learning behavior to its neural basis. *Front. Psychol.* 14:1100969. doi: 10.3389/fpsyg.2023.1100969
- Zhou, X., Fu, X., Lin, C., Zhou, X., Liu, J., Wang, L., et al. (2017). Remodeling of dendritic spines in the avian vocal motor cortex following deafening depends on the basal ganglia circuit. *Cereb. Cortex* 27, 2820–2830. doi: 10.1093/cercor/bhw130