Physica D 473 (2025) 134555

Contents lists available at ScienceDirect

Physica D: Nonlinear Phenomena

journal homepage: www.elsevier.com/locate/physd

Conservation of avian vocal heritage through synthetic song reintroduction

Roberto Bistel [a,b], Ana Amador [a,b], Gabriel B. Mindlin [a,b,c,*]

[a] Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física. Ciudad Universitaria, 1428, Buenos Aires, Argentina
[b] CONICET - Universidad de Buenos Aires, Instituto de Física Interdisciplinaria y Aplicada (INFINA). Ciudad Universitaria, 1428, Buenos Aires, Argentina
[c] Departamento de Matemática Aplicada, Universidad Rey Juan Carlos, Madrid, Spain

ARTICLE INFO

Keywords:
Dynamical systems models
Birdsong learning

ABSTRACT

Until the spread of digital recording technology, our knowledge of the history of avian vocal culture was based on onomatopoeic descriptions or notations inspired by musical notation. In the 1960s, hand-drawn annotations of the frequency modulations in the songs of Rufous-collared sparrows (*Zonotrichia capensis*) were made in a natural reserve in Argentina. Some of these song themes have been preserved to the present day, while others have not appeared in recent recordings. In this work, we used a dynamical system based on an avian vocal production model to generate synthetic songs. We designed a song that matches the description of a currently absent theme and used it as a vocal tutor for wild juveniles. The success of our approach suggests a promising tool for preserving the vocal repertoire of wild birds.

1. Introduction

Biodiversity is crucial for maintaining ecosystem balance and ensuring species survival. However, human activities such as deforestation, pollution, and climate change continuously threaten this diversity, leading to its rapid decline [1]. The loss of biodiversity is often linked to the loss of genetic information. The extinction of a species results in the complete disappearance of its genome, while a reduction in population decreases the genetic variability necessary for adaptation. However, biodiversity also includes behavioral diversity, both within and across species [2], with learned behaviors being particularly vulnerable. Reintroduction programs for restoring individuals to an ecosystem do not always guarantee the recovery of these learned behaviors.

The song of oscine birds is a behavior that heavily relies on learning, involving approximately 4,000 of the 10,000 known bird species [3]. While some aspects of song are genetically determined, others are learned through interaction between juveniles and adults. If the learning process is interrupted, reintroducing individuals does not necessarily ensure the recovery of these behaviors. In this work, we propose a strategy to reintroduce absent songs into a population of wild birds.

A recent study reported the persistence of certain song themes in a *Zonotrichia capensis* population in a natural reserve in Buenos Aires Province, Argentina [4]. In this location, hand-drawn annotations of frequency modulations in the songs of individuals of this oscine species were made in the 1960s [5].

Rufous-collared sparrows (*Zonotrichia capensis*) - also called chingolos - is a widespread and non-threatened South American bird species, capable of adapting to diverse environments. As a species highly adapted to anthropogenic environments, it is resilient to moderate disturbances, reducing the likelihood of adverse effects on its population dynamics or behavior.

The song of *Zonotrichia capensis* consists of two parts: an initial set of two to four notes with frequency modulation (theme) and a second part consisting of repeated notes with decreasing frequency (trill). While the trill remains consistent among individuals in the same region, the themes vary. Each individual typically sings one theme, although some can sing two or even three. Each song lasts approximately two seconds. Some of the themes recorded by [5] persist, while others have disappeared over time [4]. Only three of the themes described by [5] are still present in the current population [4,6].

In this work, we delimited an area of the reserve [4–6] and selected a theme, previously reported by [5], that was no longer sung, with the aim of reintroducing it into the population's repertoire. This theme had not been recorded between 2020 and 2022, so the initial challenge was to synthesize a song compatible with its acoustic properties. To generate this song, we used a mathematical model for synthetic song generation. The model describes the physical dynamics of bird song, allowing us to reconstruct the temporal parameters necessary to create a synthetic song similar to the schematics in [5]. After achieving the desired timbre, the

* Corresponding author.
E-mail address: gabo@df.uba.ar (G.B. Mindlin).

https://doi.org/10.1016/j.physd.2025.134555
Received 26 August 2024; Received in revised form 19 December 2024; Accepted 27 January 2025
Available online 29 January 2025
0167-2789/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

individuality of the modeled song was established by replicating the time evolution of the frequencies in the introductory syllables.

Based on the physics of avian phonation, the model generates a sound signal that emulates the temporal and acoustic properties of the original song [7–10]. In laboratory studies, the synthetic song evoked responses in song-selective neurons similar to those triggered by real song [11,12]. Additionally, positive results were observed through electromyographic recordings of the syrinx muscles, which receive direct input from the neural network responsible for generating and processing birdsong [13].

The validation of the model has been a gradual process. Initially, it was validated under laboratory conditions, as previously mentioned, and later confirmed through field experiments using synthetic songs that mimic those of *Zonotrichia capensis* [14]. During the breeding and nesting seasons, the birds' behavioral responses to both real and synthetic songs were quantified, revealing territorial defense responses with no significant differences between the two. Both real and synthetic conspecific songs triggered an increase in song rate, while heterospecific songs did not elicit responses from the individuals [14]. These results suggest that synthetic songs are an adequate substitute for real songs in ecological contexts and can serve as vocal tutors for wild juveniles.

There are precedents for developing a training protocol for wild juveniles based on recordings. In a study conducted with *Savannah sparrows*, playback transmission enabled learning in wild juveniles [15]. However, in [15], the recordings featured material sung by conspecific individuals. It has also been demonstrated that captive juveniles of canaries (*Serinus canaria*) can learn synthetic songs and later adjust them during adulthood to conform to the structural rules of adult song [16]. This principle supports our research, in which we explore the capacity of *Zonotrichia capensis* juveniles to learn and adapt synthetic songs as part of their vocal repertoire.

Therefore, we generated a synthetic song compatible with one of the themes described by [5], which is absent from the current population, and played it during the critical sensory period from October to February. During this period, juveniles learn their song by imitating an acoustic model. The song was played through electronic devices capable

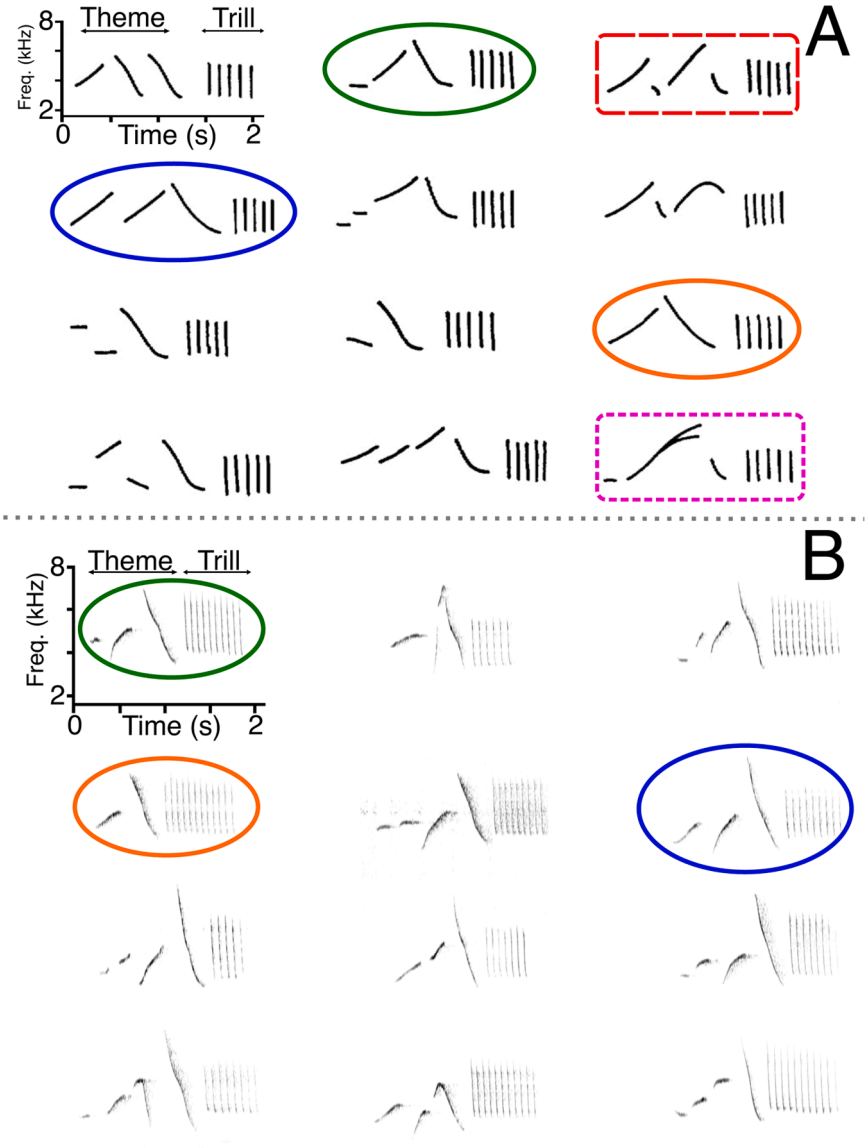

Fig. 1. Vocal repertoire of *Zonotrichia capensis* from the 1960s to the present. The song consists of a theme followed by a trill. (**A**) Hand-drawn annotations depicting the frequency modulations of *Zonotrichia capensis* songs [5]. The theme selected for insemination is marked in a red square with dotted lines. Additionally, a theme exhibiting the biphonic capability of the syrinx is highlighted in a fuchsia square with dotted lines. Persistent themes identified in the 1960s and still present in the current population are marked with circles in the same colors in both (A) and (B). (**B**) Sonograms of songs recorded between 2020 and 2022. Additional sonograms from the same area are provided in Fig. S1.

of simultaneous audio recording and playback [14], which acted as vocal tutors for the juveniles. An analysis of the recordings revealed that at least two juveniles acquired a song with acoustic properties similar to the synthetic song used as a tutor.

The paper is organized as follows: Section II describes the materials and methods used in the study, including a detailed explanation of the synthetic song generation model and the process of selecting and synthesizing the song used as a tutor. The protocol followed for the playback of the synthetic song is also presented. Section III presents the results obtained, and Section IV discusses the conclusions derived from the study.

2. Materials and methods

2.1. Synthetic song generation model. Synthetic song used as a tutor

The mathematical model for synthetic song generation is based on the dynamics of sound production in the syrinx labia. The syrinx, the vocal organ of birds, is located at the junction of the primary bronchi and the trachea. The oscine syrinx contains two sound sources, each consisting of a pair of soft tissues, similar to human vocal cords, known as labia [17,18]. This unique bilateral structure allows independent control of each side, enabling the simultaneous production of two distinct sound frequencies, a phenomenon known as *biphonation*.

For example, the theme observed in the lower right corner of Fig. 1A (marked in a fuchsia square in dotted lines), where the upsweep shows two bands, can be explained by this biphonic capability. Each bronchial side of the syrinx can oscillate at different frequencies, generating two sound bands that can combine into a single vocalization. This behavior is well-documented in studies on the biomechanical functioning of the avian syrinx [19,20].

The labia oscillate when air from the air sacs flows through them. The oscillations generate sound waves that are filtered as they pass through the trachea and the oropharyngeal cavity, before being emitted as song through the beak [9,20–22].

In the model, air sac pressure and syrinx muscle tension are treated as time-dependent parameters, adjusted to ensure that the sound exhibits the temporal and acoustic properties of the target song. Many of the timbral characteristics of the sound arise from the nonlinear properties of the model [10,22,23].

The labia are assumed to be in a stationary position when the bird is silent. Once the parameter representing air sac pressure is increased, a threshold for oscillatory motion is reached. If the problem parameters remain in the phonation region of the parameter space, the airflow is modulated, and sound is produced. As the pressure decreases, the sound eventually stops (that is, the syllable ends). A qualitative change in dynamics when the parameters are varied is known as a bifurcation. Near the values of the parameters where the bifurcation occurs, the model can be transformed into simple equations that describe the dynamics of the system.

The mean position of each labium is represented by the variable x, and the dynamic system describing these oscillations follows Newton's equations [24,25]:

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = -\gamma^2\kappa x - \gamma x^2 y + \beta\gamma y$$

Here, γ represents the system's time scale; κ is the restitution constant, proportional to the tension in the labia; and β is the negative damping, proportional to the pressure in the air sacs. Variations in the position of x modify the effective area between the labia, causing fluctuations in the airflow and, consequently, generating pressure waves. Assuming that the trachea acts as a tube of length L, the pressure at its entrance can be expressed as:

$$P(t) \propto y + P_{back}\left(t - \frac{L}{c}\right)$$

$$P_{back}(t) = -rP\left(t - \frac{L}{c}\right)$$

where c is the speed of sound, and r is the reflection coefficient at the upper end of the trachea. The sound produced in the syrinx propagates through the trachea and is further filtered by the oro-esophageal cavity (OEC), which is dynamically adjusted by the bird to emphasize specific frequencies [21,26]. This filtering process is modeled as a Helmholtz resonator coupled to the beak, whose parameters influence the acoustic properties of the emitted sound [21,23].

The vocal tract dynamics can be described using an electrical circuit analogy, where the acoustic pressure is represented as an electric potential and the airflow as an electric current [21]. The resonating cavity is connected to the trachea through the glottis, which is modeled as an inductance L_g (associated with the length or area of the glottis) in series with a capacitance C_h (representing the volume of the OEC). To account for energy dissipation in the resonator, a resistance R_h is included in series with these elements. The Helmholtz resonator is coupled to the external environment via the beak, modeled as an inductance L_b (associated with the length of the beak) and a resistance R_b, which are connected in parallel to the resonator's capacitor [21,23].

Fig. 2 shows the equivalent electrical model used to describe the vocal tract dynamics [23]. The governing equations of the system are as follows:

$$\frac{di_1}{dt} = i_2$$

$$\frac{di_2}{dt} = -\frac{1}{C_h L_g} i_1 - \left(\frac{r_{dis}}{L_b} + \frac{r_{dis}}{L_g}\right) i_2 + \left(\frac{1}{C_h L_g} - \frac{r_b\, r_{dis}}{L_b L_g}\right) i_3 + \frac{1}{L_g}\frac{dP_{back}}{dt} + \frac{r_{dis}}{L_b L_g} P_{back}$$

$$\frac{di_3}{dt} = -\frac{L_g}{L_b} i_1 - \frac{R_b}{L_b} i_3 + \frac{1}{L_b} P_{back}$$

In these equations, i_3 represents the output sound, while r_{dis} and r_b correspond to the resistances of the resonator and the beak, respectively. The resonant frequency depends strongly on C_h and L_g. Specifically, a decrease in glottal opening (L_g) lowers the resonant frequency, as does an increase in the OEC volume (C_h) when L_g is held constant. The resistance r_{dis} controls the resonance amplitude of the OEC, influencing its relative weight compared to the frequencies amplified by the trachea. Changes in r_b reduce the total amplitude of the sound, while variations in L_b induce slight spectral changes at higher frequencies [21,23]. The

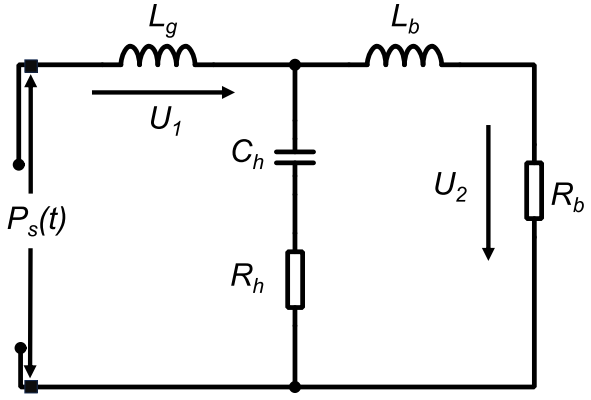

Fig. 2. Equivalent electrical circuit model of the bird's vocal tract, representing the trachea, oropharyngeal cavity, glottis, and beak. The circuit diagram captures the dynamics of the Helmholtz resonator, which simulates the filtering properties of the oropharyngeal cavity.

values for these parameters were taken in this work as $\left(\frac{1}{C_h}, \frac{1}{L_b}, \frac{1}{L_g}, r_b, r_{dis}\right) = \left(7\ 10^9, 10^{-4}, 5^{-2}, 10^7, 1.2^7\right)$.

The Helmholtz resonator model is essential for capturing the dynamic filtering properties of the vocal tract, which play a critical role in shaping the spectral structure of the emitted song. By incorporating these biomechanical and acoustic constraints, the synthetic song model accurately reproduces the nonlinear features observed in *Zonotrichia capensis* vocalizations, such as abrupt transitions in frequency modulation and amplitude envelopes.

For many species, the various acoustic modulations in song are translated into a set of basic physiological instructions called "gestures" [7]. In the case of the *Zonotrichia capensis*, these acoustic modulations can be defined using three frequency modulation patterns: sinusoidal, linear, and exponential downsweep. The specific parameters for each modulation pattern are detailed in Table 1 [6,27].

– Song Selection for Synthesis:

To select the song for synthesis, we compared the schematic notes from [5] shown in Fig. 1A with the calculated spectrograms of all themes recorded between 2020 and 2022 in the study area [4,6] (see Fig. 1B and Fig. S1). We aimed to identify a theme with distinctive acoustic characteristics that clearly set it apart from the current sonograms, ultimately selecting the third theme in the first row of Fig. 1A (marked in a red square in dotted lines). This theme features an alternation between syllables with increasing and decreasing frequency modulation.

Since no recordings of an identical theme existed, functions were generated to emulate the tension of the syrinx muscles using exponential and linear modulations. To estimate the frequency limits of the four syllables of the theme, a grid with a 500 Hz step size was employed. The initial and final frequency values for each syllable were: $\left(f_{i,1}, f_{f,1}\right) = (2kHz, 5kHz)$, $\left(f_{i,2}, f_{f,2}\right) = (2.5kHz, 1.5kHz)$, $\left(f_{i,3}, f_{f,3}\right) = (2.5kHz, 6kHz)$, $\left(f_{i,4}, f_{f,4}\right) = (3.5kHz, 2kHz)$.

Using these values, a frequency list was generated for each syllable. The values of κ that yield solutions with fundamental frequencies ω follow the function:

$$\kappa(\omega) = 6.56 \times 10^{-8}\omega^2 + 4.\ 23 \times 10^{-5}\omega + 2.67 \times 10^{-2}$$

This (κ, ω) relationship results from multiple numerical simulations of the model. Thus, the model parameters are adjusted to synthesize a song with the desired fundamental frequency. The song's timbral content arises from the nonlinear dynamics underlying the movement of the syrinx lips and the successive filtering occurring in the trachea and oropharyngeal cavity [10,11,22,23]. In Fig. 3, we show the sonogram of the synthetic sound used for insemination.

2.2. Synthetic song playback protocol

The synthetic song insemination protocol employed three autonomous electronic systems for audio playback and recording [14]. Audio was recorded in "wav" format, mono, 8-bit, with a sampling frequency of 22 kHz and a file duration of 2 minutes. The synthetic song was converted into a "wav" file for playback, in mono, 16-bit, with a sampling frequency of 48 kHz and a duration of 2.1 seconds.

Table 1
"Basic patterns for the gestures used to synthesize the song of the Zonotrichia capensis."

Modulation pattern	Frequency	Parameters
Sinusoidal	$w(t) = w_f + (w_i - w_f)\left(\frac{t - t_i}{t_f - t_i}\right)$	w_f, w_i, t_f, t_i
Linear	$w(t) = w_{av} + A\sin\left(\alpha_i + (\alpha_f - \alpha_i)\left(\frac{t - t_i}{t_f - t_i}\right)\right)$	$w_{av}, A, \alpha_f, \alpha_i, t_f, t_i$
Exponential	$w(t) = w_f + (w_i - w_f)e^{-\frac{3(t-t_i)}{(t_f - t_i)}}$	w_f, w_i, t_f, t_i

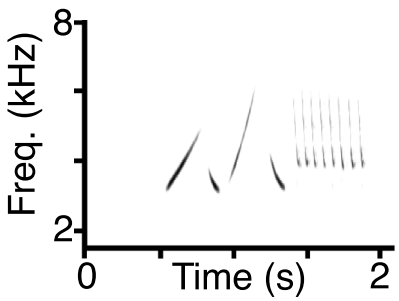

Fig. 3. Sonograms of the synthetic sound used for insemination. The theme corresponds to a song described in [5] (see Fig. 1A), and features pitch and temporal characteristics that have been absent from the songs recorded between 2020 and 2022.

The synthetic song insemination covered an area of approximately 1353 m^2 in Parque Pereyra Iraola, Buenos Aires Province, Argentina [*GPS* coordinates: (-34.863232, -58.121825); (-34.863630, -58.121872); (-34.863422, -58.122526)]. This area, previously studied in [14], has been systematically used to record the songs of *Zonotrichia capensis* since 2020. The site was chosen for its dense shrub and weed vegetation and proximity to agricultural fields, which are favorable for the presence of juveniles. The trees with the greatest number in the areas are coronillos (*Scutia buxifolia*), espinillos (*Vachellia caven*), talas (*Celtis tala, Celtis ehrenbergiana*), pines (*Pinus*) and eucalyptus (*Eucalyptus*).

While the area contains anthropogenic noise, human disturbance and animal traffic remain minimal, providing a realistic yet controlled ecological context for the study. The high density of *Zonotrichia capensis*, alongside the presence of other species, further supported the site's suitability. The area also harbors other bird species, including the House Sparrow (*Passer domesticus*), Rufous-bellied Thrush (*Turdus rufiventris*), Great Kiskadee (*Pitangus sulphuratus*), Rufous hornero (*Furnarius rufus*), Monk Parakeet (*Myiopsitta monachus*) and Spot-winged pigeon (*Patagioenas maculosa*), creating a diverse avian community.

The experiments conducted in our work did not involve direct manipulation of individuals. The experiments adhered to established ethical guidelines for minimizing disturbance in avian field research [28–30]. Playback sessions were conducted during the peak singing hours - early morning- and were limited to a maximum duration of 8 hours. The intervals between songs were set at a slightly higher rate than the natural singing intervals of *Zonotrichia capensis*, ensuring an appropriate balance between stimulating vocal responses and avoiding overstimulation.

The three playback devices were placed in pine trees, with the volume adjusted to allow each system to be heard within a 65-meter radius, ensuring that all three systems were audible from any point within the area. The exact locations of the playback devices are: (-34.863422, -58.122526); (-34.863630, -58.121872); (-34.863232, -58.121825). More details of the location of the site are provided in the project's GitHub repository, where satellite images of the area are accessible [31]. Playback volumes were measured and adjusted using a decibel meter to match the natural vocalizations of *Zonotrichia capensis*, avoiding acoustic interference.

The protocol was divided into two stages:

– **First stage** (November 3, 2022 - February 3, 2023):

During this phase, the synthetic song playback began between 6:00 am and 8:00 am, and ended between 12:00 pm and 2:00 pm. The

interval between songs was randomly set to 6 – 12 seconds. These values resulted in a higher song frequency compared to the normal intervals for *Zonotrichia capensis*, which range from 10 to 15 seconds [5,32,33]. The systems recorded continuously throughout the playback schedule.

– **Second stage** (July - September 2023):

In this phase, two recording-playback devices were used, following the same schedule as in the first stage, but with playback intervals adjusted to the normal range of 10 to 15 seconds.

The two stages of the protocol correspond to the vocal learning and consolidation periods of *Zonotrichia capensis* juveniles. In the southern hemisphere, the first stage coincides with the critical sensory period, from October to February, when juveniles learn an acoustic model from their environment. From February to July, the individuals enter a period of silence. The second stage, from July to September, marks the end of winter and the start of spring, when juveniles born the previous year resume practicing and refine their song based on the learned model [32].

3. Results and discussion

Data collection via the recording-playback devices occurred at intervals of 7 to 10 days, enabling progressive analysis to identify whether juveniles had begun practicing the synthetic song.

On January 20, 2023, one of the recording devices at Site 1 captured, for the first time, the song of a juvenile vocalizing a theme with characteristics similar to the synthetic song. Fig. 4A shows a four-second spectrogram of this recording. The first two seconds correspond to the juvenile's theme (without the trill, typical of juvenile practice [32,34]), while the remaining two seconds capture the synthetic theme played by the device.

A few days later, 50 meters from Site 1, a juvenile was observed practicing a theme similar to the synthetic song in Site 2. Continuous recordings until early February (see Fig. 4B) revealed that the juvenile began incorporating a trill of a few notes, a typical feature of song practice.

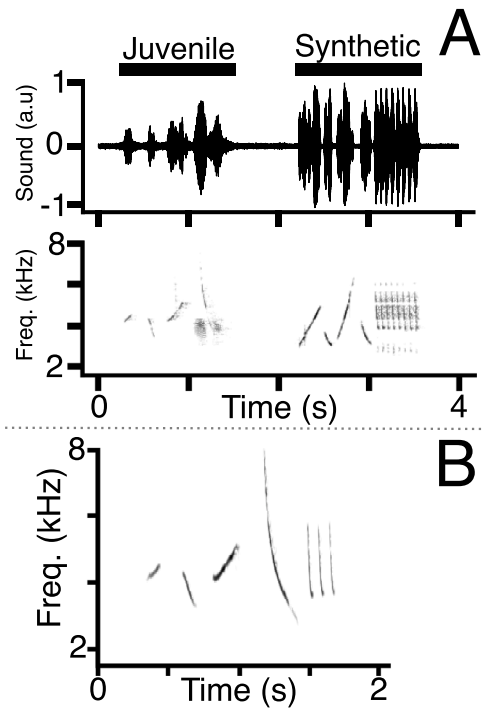

Fig. 4. Sonograms of juvenile recordings during the first stage, spanning from late January to early February 2023. (**A**) The first recording of a juvenile at Site 1, where it sings with the synthetic tutor, replicating the frequency modulation of the synthetic song. (**B**) A sonogram from Site 2 recorded during the first stage. The juvenile's trill consists of fewer, more spaced notes, a characteristic feature of an undeveloped juvenile song.

The juvenile's acquired song alternated between syllables with increasing and decreasing frequencies, consistent with the synthetic theme. None of the songs recorded between 2020 and 2022 exhibited this feature. The final downsweep syllable in the recorded song spanned a broader frequency range compared to the synthetic version. This feature is shared by all adults recorded in the field. It is possible that this characteristic is species-specific, although it cannot be ruled out that the juvenile combined qualitative elements of the synthetic theme with quantitative features from the songs of other adults in the territory.

During the second stage (July - September 2023), the protocol resumed at Sites 1 and 2. For the first three weeks of July 2023, no significant song activity from *Zonotrichia capensis* was detected, likely due to the cold winter temperatures. However, by late July, an increase in song activity was observed. At Site 1, no further recordings of themes compatible with the synthetic song were found.

At Site 2, beginning on August 1, 2023, recordings captured multiple songs from an individual exhibiting a theme compatible with the synthetic song, particularly its downsweep in the second syllable. Upon detailed analysis of the recordings, we detected subtle vocalizations from a second individual. This led us to return to the field, where we successfully recorded the interaction between the two birds. Differences in syllable morphology and alternating song frequencies confirmed the presence of two distinct individuals. The sonograms of both individuals recorded during the second stage at Site 2 are presented in Fig. S2. Supplementary materials include a recording that captures the interaction between the two individuals.

To confirm and quantify these differences, we used a Siamese neural network, previously applied to analyze theme persistence in *Zonotrichia capensis* [4]. This type of neural network compares two inputs to determine their similarity. It consists of two convolutional neural networks with shared weights and parameters, processing inputs independently. Their outputs are then combined using a similarity metric, such as Euclidean distance, to produce a scalar value indicating similarity. The network embeds inputs into a lower-dimensional feature space, where similar inputs cluster closely, and dissimilar ones are farther apart.

In our case, the network was trained with synthetic spectrograms generated by our dynamical model, validating its ability to compare field recordings with historical data. The network generates *3D* embeddings, which we reduced to *2D* using *PCA* to visualize theme clusters effectively.

This approach efficiently processes large datasets, such as the extensive recordings collected during the singing season. By detecting subtle differences in frequency modulation and temporal patterns, it provides an objective, quantitative alternative to traditional subjective methods.

Fig. 5 presents a *2D* visualization of the theme similarity analysis. The network identified three clusters: one showing an overlap of themes from the recordings at Site 1 and Site 2 during January - February (first stage); and two distinct clusters corresponding to individuals recorded at *Site 2* during July - September (second stage). This suggests that the juvenile initially recorded at Site 1 moved to Site 2, incorporating trill syllables during practice.

While the Siamese network clearly differentiates the two individuals recorded during the second stage, it remains uncertain whether the juvenile's song from the first stage developed into one of the second-stage recordings. The results confirm that at least two individuals successfully learned and replicated the synthetic theme introduced.

It is important to note that this introduced theme has not appeared in systematic recordings of the area since at least 2020. As discussed in [15], when introducing a novelty into a population, careful consideration of the project's ethical implications is essential. In our case, the *Zonotrichia capensis* population is not endangered, and the introduced

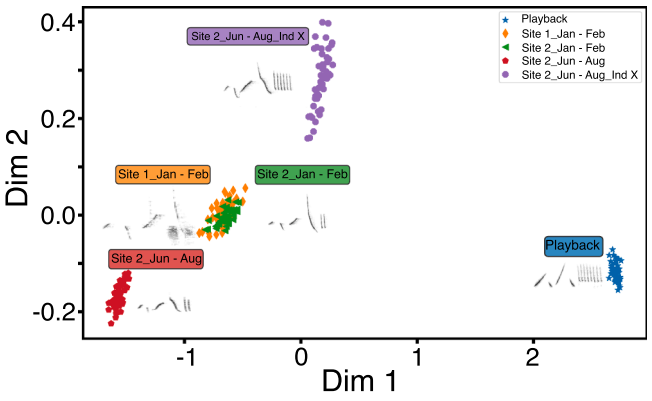

Fig. 5. *2D* visualization of the theme similarity analysis for all individuals recorded with a theme compatible with the synthetic song.

vocal feature -while novel- is subtle and consistent with songs recorded decades ago in the same area [5].

Similar to natural learning between tutors and juveniles, the synthetic song was not copied exactly. The juveniles' songs featured a short upsweep followed by a long downsweep before the trill, differing from the synthetic version's long upsweep and short downsweep. This reflects the flexibility of vocal learning in *Zonotrichia capensis*. Juveniles incorporate elements of the synthetic song, but adapt them to align with local vocal patterns and species-specific constraints.

The broader frequency span of the final downsweep observed aligns with the adult repertoire, as shown in previous recordings (see Fig. 1B and Fig. S1). Thus, it is possible that this feature was acquired from other real tutors or represents an expression influenced by the species' genetic programming. The juveniles may be integrating the qualitative structure of the synthetic song with quantitative traits learned from conspecifics in their environment. Such flexibility is well-documented in oscine birds, where final songs result from a combination of innate predispositions, tutor input, and social feedback during development [35, 36]. Juveniles often modify tutor songs, incorporating partial elements while refining them to match population-specific song structures.

The results emphasize the importance of considering the interaction between synthetic stimuli and the natural acoustic environment. The synthetic song serves as a vocal template, providing juveniles with a pattern to imitate. However, the final song reflects both the influence of the synthetic song and the juvenile's exposure to the existing vocal culture of its population. This is consistent with observations in other species, like Savannah sparrows (*Passerculus sandwichensis*), where experimentally introduced songs are adapted within existing repertoires [15].

This reflects the natural plasticity of song learning and the constraints imposed by the biomechanical and neural systems of the species [9]. The observed modifications likely arise from biomechanical preferences or optimization strategies within the physiological limits of the syrinx and vocal tract [20]. In any case, the presence of syllables with alternating frequency modulations in two individuals recorded near our recording-playback devices suggests that this song may have been effectively inseminated through exposure to our stimulus.

Overall, our results demonstrate that synthetic songs, even when not perfectly replicated, effectively seed vocal patterns in juvenile birds. This supports the model's utility for studying vocal learning, cultural transmission, and the interaction between synthetic stimuli and natural song development. The ability of juveniles to integrate synthetic templates with local repertoire traits reinforces the biological realism of the model and its value for future studies on vocal behavior and song evolution.

While our study focuses on the direct impact of human-mediated cultural transmission, it aligns with the broader framework of cultural evolution in birdsong. The study of [37] have demonstrated the role of social learning in the cumulative cultural evolution of bird song. Our findings, while specific to a particular species and context, contribute to this broader understanding by showing how human intervention can accelerate cultural change.

A key difference between our study and [37] is the direct manipulation of the cultural environment. By introducing synthetic songs, we were able to test the limits of bird song learning and explore the potential for human-mediated cultural evolution. This approach offers a unique opportunity to study the mechanisms underlying cultural transmission and to inform conservation strategies for threatened bird species. Ultimately, both studies highlight the intricate interplay between genetic predisposition, social learning, and environmental factors in shaping bird song diversity. By understanding these factors, we can develop effective strategies to conserve and protect bird populations and their cultural heritage.

The synthetic song playback protocol did not disrupt the natural behavior of *Zonotrichia capensis* individuals. Analysis of the recordings shows that birds respond to synthetic song playback, and stay close to the device, including those that learned the synthetic theme (see Fig. S3). This response indicates that the playback effectively elicits natural territorial and communicative behaviors while ensuring that the protocol remains minimally invasive to their natural activities.

We continuously monitor the area with our automatics recorders devices, allowing us to quantify changes in vocal repertoire diversity, identify any potential impacts on conspecific and heterospecific individuals, and assess the long-term effects of synthetic song exposure. To date, our data do not show significant variations in the number of individuals or song patterns, neither in this species nor in others cohabiting the area.

4. Conclusions

A low-dimensional physical model of the avian vocal apparatus, capable of generating song-like syntheses, enables us to test the hypotheses underlying its construction. In this work, we evaluated the hypothesis that synthetic songs could serve as effective vocal tutors for *Zonotrichia capensis* juveniles. Although it is not obvious that learning can occur without visual interaction with an adult conspecific [38], previous studies with *Savannah sparrows* have demonstrated that recordings can be effective as tutors [15]. Our findings indicate that this was also the case for *Zonotrichia capensis*.

We demonstrated that *Zonotrichia capensis* juveniles can learn and reproduce synthetic songs generated by a mathematical model based on the dynamics of the avian syrinx. This model accurately recreated the acoustic and timbral properties of the species' songs, ensuring that the synthesized themes were realistic enough for the juveniles to learn.

Across the two stages of the song insemination protocol, we recorded at least two individuals in the population actively acquiring and practicing the synthesized theme. This finding is significant as it demonstrates the feasibility of introducing a synthetic song—designed in the laboratory with specific frequency modulations—into a wild population. Moreover, it suggests that juveniles not only have the capacity to learn artificial songs but also to adapt and integrate them into their natural vocal repertoires.

The results reinforce the feasibility of using synthetic songs as a tool for conserving and reintroducing vocal repertoires in wild bird populations. This approach opens up opportunities to preserve or reintroduce songs that have been lost or are at risk of disappearing, thereby contributing to the conservation of acoustic biodiversity. Future studies should investigate how these synthetic songs affect social interactions and the evolution of vocal repertoires in natural populations over time.

CRediT authorship contribution statement

Roberto Bistel: Writing – original draft, Visualization, Validation, Investigation, Data curation. **Ana Amador:** Methodology, Investigation,

Data curation. **Gabriel B. Mindlin:** Writing – original draft, Project administration, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

G Mindlin reports financial support and travel were provided by University of Buenos Aires. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially funded by Universidad de Buenos Aires (UBACyT), ANCyT, and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). G.B.M. is grateful for the hospitality of URJC during his sabbatical stays.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.physd.2025.134555.

Data availability

Data will be made available on request.

References

[1] O.E. Sala, et al., Global biodiversity scenarios for the year 2100, Science (1979) 287 (2000) 1770.
[2] K.N. Laland, W. Toyokawa, T. Oudman, Animal learning as a source of developmental bias, Evol. Dev. 22 (2020) 126.
[3] P.R. Marler, H. Slabbekoorn, Nature's Music: The Science of Birdsong, Elsevier, New York, NY, USA, 2004.
[4] R.A. Bistel, A. Martinez, G.B. Mindlin, An analysis of the persistence of Zonotrichia capensis themes using dynamical systems and machine learning tools, Chaos Solit. Fract. 165 (2022) 112803.
[5] F. Nottebohm, The Song of the Chingolo, Zonotrichia capensis, in Argentina: description and evaluation of a system of dialects, Condor. 71 (1969) 299.
[6] R.A. Bistel, A. Martinez, G.B. Mindlin, Neural networks that locate and identify birds through their songs, Eur. Phys. J. Spec. Top. 231 (2022) 185.
[7] T. Gardner, G. Cecchi, M. Magnasco, R. Laje, G.B. Mindlin, Simple motor gestures for birdsongs, Phys. Rev. Lett. 87 (2001) 208101.
[8] R. Laje, T.J. Gardner, G.B. Mindlin, Neuromuscular control of vocalizations in birdsong: a model, Phys. Rev. E 65 (2002) 051921.
[9] G.B. Mindlin, R. Laje, The Physics of Birdsong, Springer Science & Business Media, Berlin, 2005.
[10] A. Amador, G.B. Mindlin, Synthetic birdsongs as a tool to induce, and listen to, replay activity in sleeping birds, Front. Neurosci. 15 (2021).
[11] A. Amador, Y.S. Perl, G.B. Mindlin, D. Margoliash, Elemental gesture dynamics are encoded by song premotor cortical neurons, Nature 495 (2013) 59.
[12] S. Boari, Y.S. Perl, A. Amador, D. Margoliash, G.B. Mindlin, Automatic reconstruction of physiological gestures used in a model of birdsong production, J. Neurophysiol. 114 (2015) 2912.
[13] A. Bush, J.F. Döppler, F. Goller, G.B. Mindlin, Syringeal EMGs and synthetic stimuli reveal a switch-like activation of the songbird's vocal motor program, Proc. Natl. Acad. Sci. 115 (2018) 8436.
[14] R.A. Bistel, A. Amador, G.B. Mindlin, Response of wild songbirds to songs synthesized with a low-dimensional model, Phys. Rev. E 109 (2024) 054410.
[15] D.J. Mennill, S.M. Doucet, A.E.M. Newman, H. Williams, I.G. Moran, I.P. Thomas, B.K. Woodworth, D.R. Norris, Wild birds learn songs from experimental vocal tutors, Curr. Biol. 28 (2018) 3273.
[16] T.J. Gardner, F. Naef, F. Nottebohm, Freedom and rules: the acquisition and reprogramming of a bird's learned song, Science (1979) 308 (2005) 1046.
[17] J.L. van Leeuwen, R. Suthers, F. Goller, C. Pytte, The neuromuscular control of birdsong, Philos. Trans. R. Soc. Lond. Ser. B 354 (1999) 927.
[18] F. Goller, O. Larsen, New perspectives on mechanisms of sound generation in songbirds, J. Compar. Physiol. A 188 (2002) 841.
[19] R.A. Suthers, Contributions to birdsong from the left and right sides of the intact syrinx, Nature 347 (1990) 473.
[20] F. Goller, O.N. Larsen, A new mechanism of sound generation in songbirds, Proceed. Natl. Acad. Sci. 94 (1997) 14787.
[21] N.H. Fletcher, T. Riede, R.A. Suthers, Model for vocalization by a bird with distensible vocal cavity and open beak, J. Acoust. Soc. Am. 119 (2006) 1005.
[22] J.D. Sitt, A. Amador, F. Goller, G.B. Mindlin, Dynamical origin of spectrally rich vocalizations in birdsong, Phys. Rev. E 78 (2008) 011905.
[23] Y.S. Perl, E.M. Arneodo, A. Amador, F. Goller, G.B. Mindlin, Reconstruction of physiological instructions from zebra finch song, Phys. Rev. E 84 (2011) 051909.
[24] A. Amador, G.B. Mindlin, Beyond harmonic sounds in a simple model for birdsong production, Chaos 18 (2008) 043123.
[25] G.B. Mindlin, Nonlinear dynamics in the study of birdsong, Chaos 27 (2017) 092101.
[26] V.R. Ohms, P. Ch. Snelderwaard, C. ten Cate, G.J.L. Beckers, Vocal tract articulation in zebra finches, PLoS ONE 5 (2010) e11923.
[27] P.L. Tubaro, G.B. Mindlin, A dynamical system as the source of augmentation in a deep learning problem, Chaos Soliton. Fract. X 2 (2019) 100012.
[28] A. De Rosa, I. Castro, S. Marsland, The acoustic playback technique in avian fieldwork contexts: a systematic review and recommendations for best practice, Ibis 164 (2021).
[29] D.E. Kroodsma, Suggested experimental designs for song playbacks, Anim. Behav. 37 (1989) 600.
[30] P.K. McGregor, et al., Design of playback experiments: the Thornbridge Hall NATO ARW consensus, in: P.K. McGregor (Ed.), Playback and Studies of Animal Communication, Springer US, Boston, MA, 1992, pp. 1–9.
[31] LSD, *LSD GitHub*, https://github.com/LaboratorioSistemasDinamicos/UnifiedSystem_2023-204.
[32] P. L. Tubaro, Aspectos causales y funcionales de los patrones de variación del canto del chingolo (Zonotrichia capensis), (1990).
[33] P.L. Tubaro, F.M. Gabelli, I.M. Mozetich, E.T. Segura, Evaluación de la hipótesis de la detección alertada en el canto del Chingolo (Zonotrichia capensis): I. Experimentos con playbacks de cantos completos y fraccionados, El Hornero 14 (1996) 27.
[34] P. Handford, Aspects of the annual cycle in Highland populations of the Rufous-Collared Sparrow, Zonotrichia capensis, Auk 97 (1980) 400.
[35] P. Marler, S. Peters, Developmental overproduction and selective attrition: new processes in the epigenesis of birdsong, Dev. Psychobiol. 15 (1982) 369.
[36] O. Tchernichovski, P.P. Mitra, T. Lints, F. Nottebohm, Dynamics of the vocal imitation process: how a zebra finch learns its song, Science 291 (2001) 2564.
[37] H. Williams, R.F. Lachlan, Evidence for cumulative cultural evolution in bird song, Philos. Trans. R. Soc. B 377 (2022) 20200322.
[38] R. Simon, et al., RoboFinch: a versatile audio-visual synchronised robotic bird model for laboratory and field research on songbirds, Methods Ecol. Evol. 14 (2023) 1092.