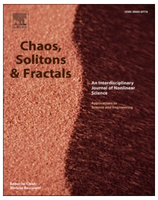
Chaos, Solitons and Fractals 200 (2025) 117100

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

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



Decomposition of respiratory motor patterns during birdsong production in terms of excitable transients

Agustin Carpio Andrada [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, Mostoles, Spain

ARTICLE INFO

Keywords:
Nonlinear dynamics
Excitability
Birdsong
Respiration

ABSTRACT

The production of birdsong involves complex respiratory motor gestures shaped by intricate neural and muscular coordination. In this work, we propose that these gestures can be decomposed into simpler dynamical elements generated by a minimal excitable system. Using air sac pressure recordings from singing canaries (*Serinus canaria*), we model each syllable as a linear superposition of transients arising from perturbations of a two-dimensional Wilson–Cowan-type excitable system. By fitting these transients to observed pressure patterns using a differential evolution algorithm, we obtain accurate reconstructions of the motor gestures underlying song. We then apply unsupervised dimensionality reduction and clustering to the extracted transients, identifying a compact set of gesture types—motor primitives—shared across individuals with distinct vocal learning histories. Our results suggest that complex motor sequences in birdsong may be constructed from a limited, reusable repertoire of dynamical modules, offering insights into the structure of learned behavior and the neural encoding of motor control. This framework introduces a principled, dynamical decomposition of vocal behavior with potential applications in the study of motor learning and neuroethology.

1. Introduction

The production of song in oscine birds requires exquisite coordination among the nervous system and a wide array of muscles. Some of these muscles regulate expiration and inspiration, generating the delicate pressure gestures necessary to drive the labia of the syrinx into oscillatory regimes. Others control the syringeal muscles themselves, reconfiguring the vocal organ and inducing specific tensions within the labia, which in turn give rise to particular sound frequencies. These motor gestures are remarkably complex. In several species, such as the canary and the zebra finch, both syringeal muscle activity and air sac pressure (the latter interpretable as a proxy for respiratory muscle activity) have been described in detail [1–7].

Interestingly, experiments conducted in sleeping birds have suggested that some of these complex gestures can be analyzed in terms of simpler substructures [8]. When a sleeping bird is exposed to a playback of its own song, one can elicit not only neural responses in

---

[*] Corresponding author at: Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física, Ciudad Universitaria, 1428 Buenos Aires, Argentina.

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

https://doi.org/10.1016/j.chaos.2025.117100
Received 29 May 2025; Received in revised form 15 August 2025; Accepted 19 August 2025

song-related brain areas but also activity in the syringeal muscles. In response to a faithful copy of the bird's own song, the recorded patterns of syringeal muscle activity closely resemble those used during actual song production. However, when the acoustic stimulus is continuously degraded, the evoked muscle patterns do not degrade gradually. Instead, it has been reported that motor gestures degrade through the loss of entire subsegments, a phenomenon described as being controlled by a "switch." A motor gesture is either elicited or not, but it does not fade progressively [8].

In nonlinear dynamics, there is a natural framework for describing such all-or-none responses: excitability. In an excitable system, a perturbation may or may not elicit a response. But if it does, the response follows a stereotyped trajectory. One may fail to perturb the system enough to elicit the response, but once triggered, the full response unfolds in a fixed form, eventually returning to the fixed point the system was before being perturbed [9]. Motivated by this observation, we set out to explore whether the complex respiratory gestures observed in the canary's song could be decomposed into a sequence of transients generated by a perturbed, simple excitable system. Furthermore, we asked whether such a decomposition could rely on the same basic elements across different individuals exposed to different vocal learning experiences.

This work is organized as follows. We begin by describing the observables used in our analysis. Then, we present the minimal dynamical elements with which we aim to decompose these observables, and we proceed to perform the decomposition. Having shown that our basis permits such a decomposition, we use an unsupervised technique to analyze the elements of the functional basis, and we demonstrate that a single discrete basis can be used to decompose the motor gestures of different individuals, regardless of their specific vocal learning histories. We conclude by discussing the implications of this finding in the context of a theory of vocal learning.

2. Methodology

We recorded songs from five adult male domestic canaries (*Serinus canaria*). To monitor subsyringeal air sac pressure, a flexible silastic cannula (outer diameter 1.65 mm) was inserted through the abdominal wall into the anterior thoracic air sac under ketamine/xylazine anesthesia. The free end of the cannula was connected to a miniature piezoresistive pressure transducer (Fujikura FPM-02PG) mounted on the bird's back. Transducer voltage signals were amplified and modulated to be compatible with audio recording hardware (MAYA 1010 sound card). Birds resumed normal singing behavior within 24–48 h post-surgery. Songs and air sac pressure were recorded continuously. Demodulation yielded relative voltage traces representing air sac pressure. All procedures adhered to NIH guidelines and were approved by the Institutional Animal Care and Use Committee (IACUC) (full procedure described in [10]) (Fig. 1).

Both acoustic and physiological signals were recorded at a sampling frequency of 44.1 kHz. Air sac pressure signals were normalized and smoothed using a Savitzky–Golay filter [11] (window length = 450 samples, polynomial order = 1), corresponding to a smoothing window of approximately 10.2 ms given the 44.1 kHz sampling rate. The zero of each signal was aligned with the median value of the normal respiratory pressure recorded outside of song.

We segmented song signals into syllables using local minima in the air sac pressure trace—corresponding to intersyllabic inspiratory phases—as temporal boundaries. This procedure resulted in a dataset comprising 303 syllables across eight recordings from five individual birds.

We hypothesized that syllables could be reconstructed as linear combinations of dynamic components—or *gestures*—each arising from a minimal excitable system. We implemented a two-dimensional Wilson–Cowan-type [12,13] model (Eq. (1)) constrained to operate in an excitable regime. Each gesture corresponds to the $x(t)$ component of the system's response to a transient perturbation (a "kick") applied at time T_{ij}, representing a fast inhibition removal. In this notation, the index j refers to the jth syllable, and the index i refers to the order of appearance of the gesture within the syllable.

Model equations for single gestures

$$\begin{aligned} \dot{x} &= \mu_j\Big(-x + S\Big(\rho_{x_{ij}} + ax - by\Big)\Big) \\ \dot{y} &= \mu_j\big(-y + S(\rho_y + cx - dy)\big) + \kappa_{ij}\delta(t - T_{ij}), \end{aligned} \tag{1}$$

where the function $S(\cdot)$ is the standard logistic function (i.e. $S(x) = \frac{1}{1+e^{-x}}$), and the parameters ($\rho_x$, μ, κ, T) control the system's phase space, timescale, impulse intensity and impulse timing, respectively. Representative dynamics under varying ρ_x and κ values are shown in Fig. 2, illustrating how different gestures emerge from the system's structure.

For each syllable, a number of gestures -determined by the fitting procedure- was fitted to the air sac pressure signal using a Differential Evolution (DE) [14] strategy, with $\rho_y = -6.7$, $a = b = c = 10$, $d = -10$. The objective function minimized the mean squared error (MSE) between the original syllable pressure trace and the sum of simulated gestures, augmented with a penalty for nonzero gesture values beyond the syllable's offset time t_f.

Importantly, DE explored the joint parameter space of all gestures assigned to a given syllable simultaneously. That is, if a syllable j was modeled with n gestures, DE operated over a single $5n$-dimensional space — where each gesture i contributed five constrained parameters: $\rho_{x_{ij}}$, μ_j, κ_{ij}, T_{ij}, and an amplitude scaling factor. The parameter search was constrained to excitable regimes (i.e., avoiding bifurcations), and the Wilson–Cowan system was integrated using a fourth-order Runge-Kutta method. The number of gestures per syllable was determined incrementally until the reconstruction error fell below a threshold of 0.05.

We collected all fitted gestures and embedded them in a low-dimensional space applying Uniform Manifold Approximation and Projection (UMAP) [15] over normalized and aligned—at maxima—zero-padded signals to ensure uniform duration across samples. A 1-dimensional projection was enough to capture the diversity of our adjusted signals. Gestures were then clustered using agglomerative hierarchical clustering, yielding a discrete set of motor primitives.$K = 5$ was the optimal number of clusters based on internal

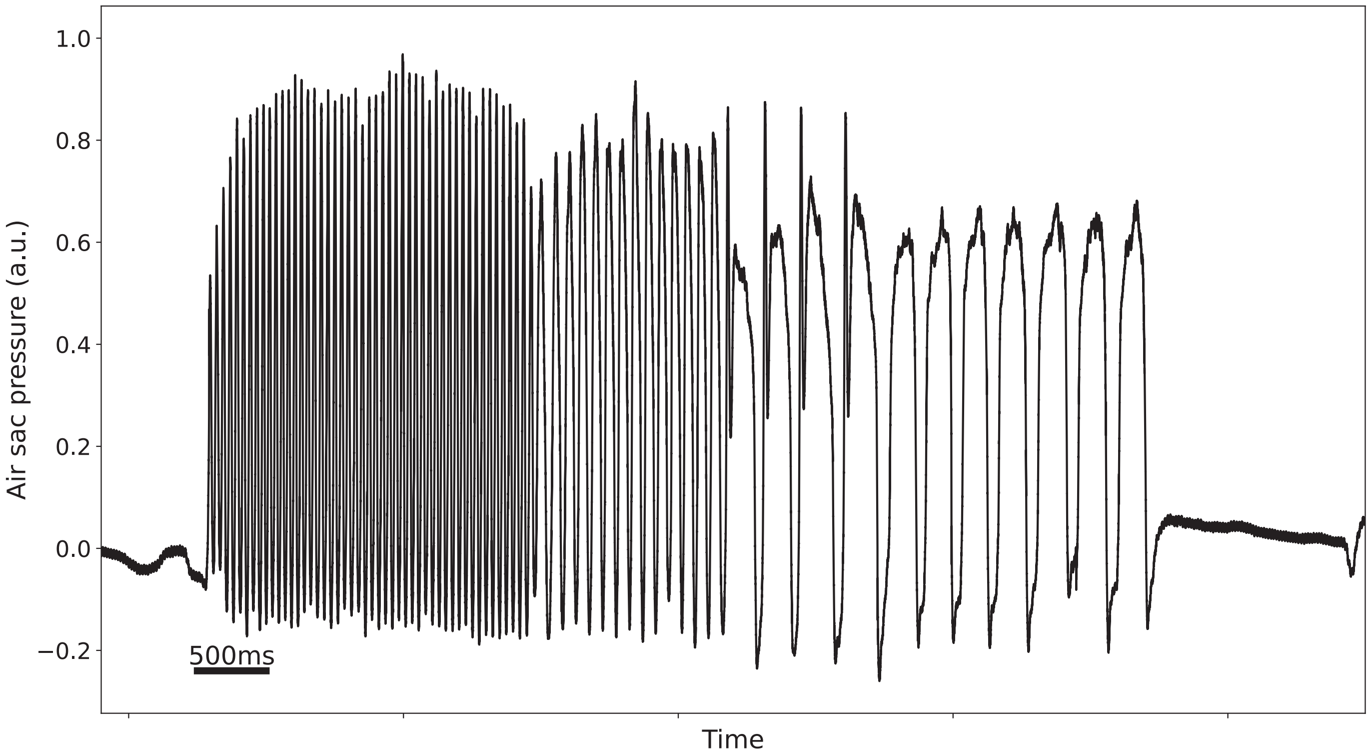


Fig. 1. Example of a full air sac pressure recording during a canary song. The trace can be interpreted as a concatenation of individual syllables of different types, separated by short inspiratory intervals. Syllables extracted from this kind of signal are the input for our excitatory gesture-based decomposition.

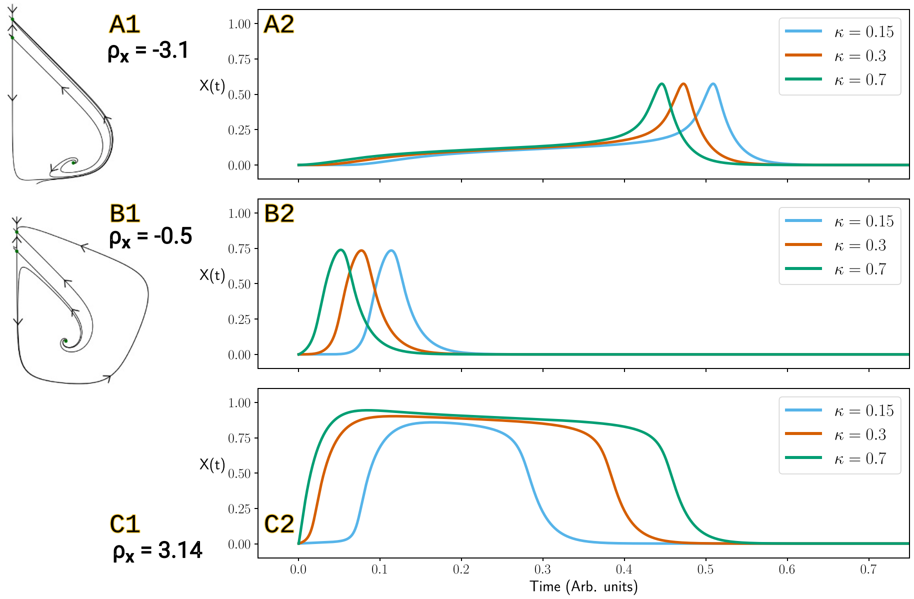


Fig. 2. Phase portraits (A1–C1) and corresponding x(t) waveforms (A2–C2) of the Wilson–Cowan model. Rows A–C vary ρ_x, producing distinct regimes; green dots denote fixed points and arrows indicate the flow. In A2–C2, colored traces show responses for different impulse amplitudes κ. Together, these panels illustrate the excitable transients that constitute our motor primitives. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

validity and stability metrics, despite the observation that under this classification some of them exhibit similar waveform shapes. These clusters, shown in Fig. 4, represent the functional basis used to reconstruct the full set of syllables. Since UMAP embeddings are seed-sensitive, we carried out 120 pairwise comparisons between clusterings from different UMAP runs with varying random seeds. The resulting Adjusted Rand Index [16] values distribution had a median of 0.8, indicating that the partitions are generally consistent.

To test for generality, we examined gesture distributions across individuals comparing them with the whole set. Despite each canary's unique vocal learning experience, the distribution of gestures in UMAP space remained qualitatively close, with all individuals employing elements from the same clustered basis. This is illustrated in Fig. 5.

3. Results

We successfully decomposed all syllables into additive combinations of dynamical elements ("gestures"), each generated by an excitable Wilson-Cowan system. Fig. 3 (top) shows a representative reconstruction of a canary song of 50 syllables, where the cumulative signal from multiple gestures accurately captures the original air sac pressure trace. For 94 % of the syllables, one or two gestures were sufficient to reconstruct a signal with an error below the 0.05 threshold, confirming the capacity of a compact excitable basis to account for the observed respiratory dynamics during canary singing. In rare cases, three gestures were required; preliminary results in other species with more complex syllables suggest that a larger number of gestures may be necessary, as expected for multimodal signals.

The generated gestures displayed variability in shape and duration, despite being generated by the same model. To assess whether this diversity could be reduced to a few clusters, we embedded the gestures in a low-dimensional space using UMAP. Alignment and normalization were essential to factor out trivial variability in phase and amplitude. The resulting 1D UMAP projection revealed a structured distribution of gestures. Using agglomerative hierarchical clustering, we identified a discrete set of gesture clusters that formed a functional basis. Each cluster represented a group of stereotyped waveform shapes (Fig. 4), interpretable as a family of motor primitives for respiratory gesture construction in birdsong.

Remarkably, the repertoire of gestures proved highly consistent across individuals. Despite differences in vocal learning histories, only minor variations in cluster usage frequency were observed when analyzing gesture distributions for each bird. All individuals employed gestures from the same global basis set, suggesting the existence of a shared low-dimensional structure underlying canary respiratory gestures.

4. Conclusions

The finding reported in this work supports the hypothesis that transients ruled by simple excitable dynamical systems can serve as a generative basis for reconstructing respiratory gestures in singing canaries. We found that the motor vocabulary repertoire is shared by individuals with different vocal learning histories, being composed by gestures grouped in a few clusters. Additionally, the fact that

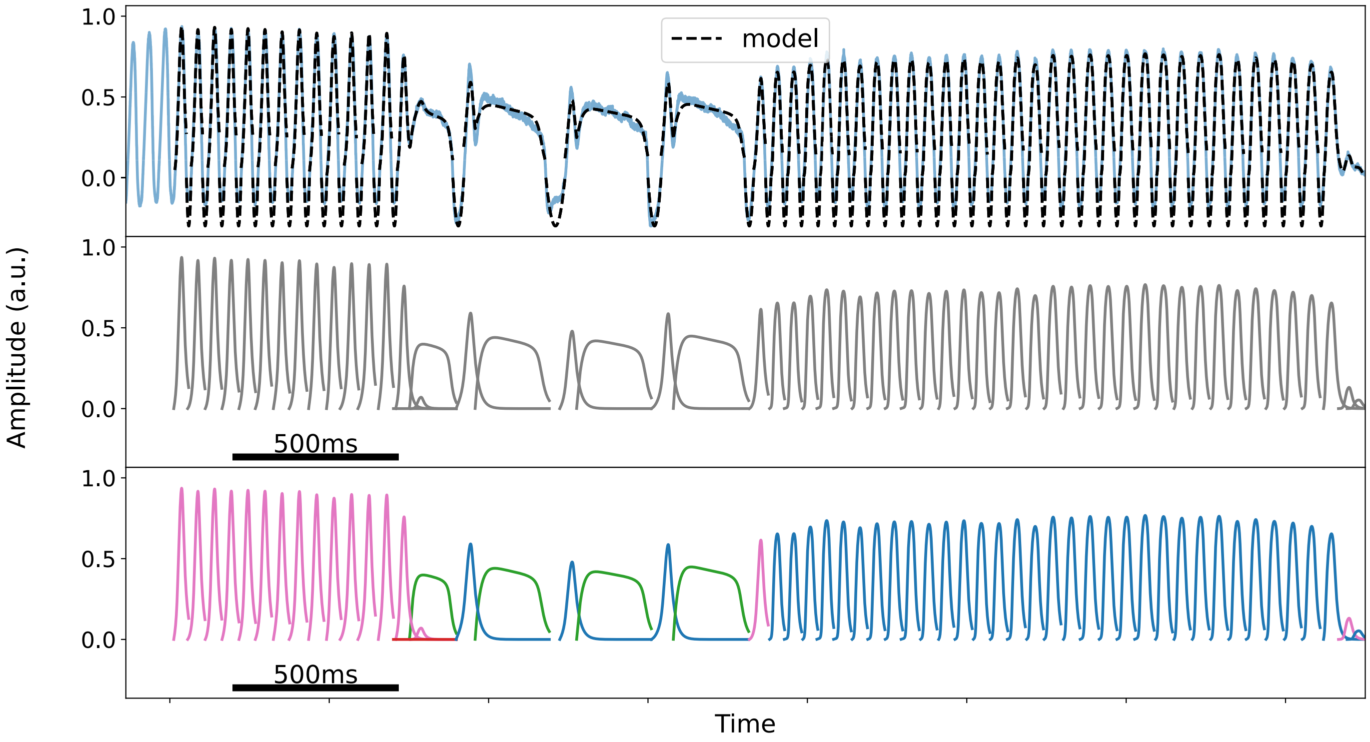


Fig. 3. Gesture-based reconstruction of a song fragment. Top: original signal (blue) overlaid with reconstructed signal (dashed black) composed of additive gestures. Inspiratory segments, corresponding to negative amplitude values, were separately fitted with a simple quadratic model, which captures the stereotyped, parabolic pressure trajectory of inspiratory phases that occur between song syllables in songbirds [2]. Middle: individual unclassified gesture waveforms (gray). Bottom: same gestures color-coded according to their cluster assignment, representing discrete motor primitive types. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

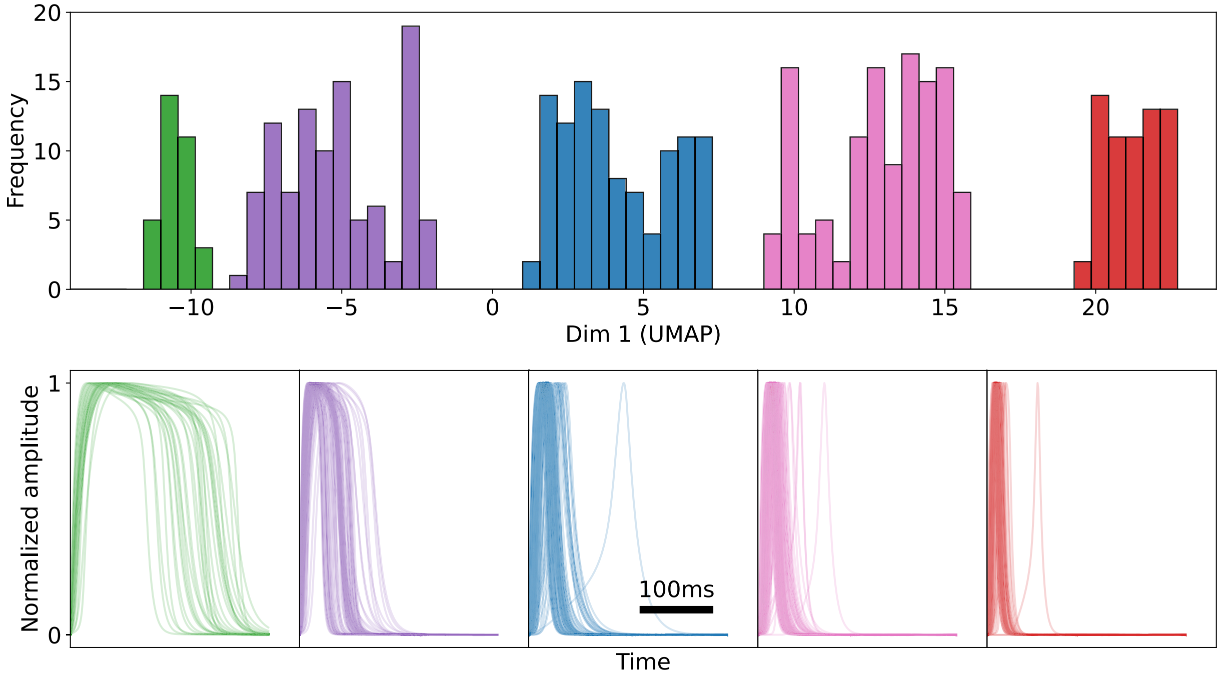


Fig. 4. Distribution of gestures in low-dimensional space. Top: histogram of gestures along UMAP dimension 1. Bottom: normalized waveforms of gestures in each cluster. Colors denote different clusters obtained via hierarchical clustering. Each cluster represents a stereotyped excitable response used to reconstruct motor gestures across syllables.

gestures from a small number of clusters sufficed to reconstruct the full syllabic repertoire suggests a strong compression of motor control complexity. This work establishes a methodological framework to address broader problems in the realm of dynamical decomposition of complex motor behaviors into reusable functional elements.

Previous studies have reproduced qualitative features of canary motor gestures by periodically forcing additive models of interacting excitatory and inhibitory populations [17,18]. These approaches, however, did not attain the quantitative accuracy achieved here, where the forcing is simplified to discrete “kicks.” In the present framework, this formulation not only yields closer approximations to the recorded signals but also reveals the existence of basic motor elements that can be elicited by temporally focused inhibitory suppression kicks.

The use of a common set of elements in the construction of these complex motor patterns suggests that these elements might be genetically programmed, while it is their combination that might be subject to learning. In other words, learned vocal behaviors may be built from combinations of stable, reusable motor modules, potentially simplifying both learning and execution. To strengthen the hypothesis of universality in this set of elements, larger studies should be conducted using birdsongs from different regions and exposed to different learning experiences. The present work can be viewed as introducing a methodology capable of testing—and potentially refuting—the idea that differences in birdsong communication arise from symbolic differences rather than from variations in the temporal arrangement or ordering of these elements.

All the transients used as minimal dynamical units in our decomposition are the large exploration in phase space that takes place after an excitable system is perturbed beyond a threshold, from its quiescent state. The different gestures correspond to the transients of the same excitable structure, but for different parameter values. Interestingly, the cluster of gestures used in our decomposition for all the songs, of all the individuals analyzed, can be generated by varying just one parameter in the model. All the gestures are transients of a dynamical system displaying three fixed points: a repulsor, a saddle and a node, with the unstable manifold of the saddle being part of the stable manifold of the attractor. Interestingly, the system, for parameter values beyond those used in our decomposition, can present additional saddle node bifurcations. Their preparation provides a critical slowing down that allows reproducing the longest pulses in our decomposition. It is for this reason that we used as a model of excitability the Wilson Cowan neural oscillator. It has been shown that its bifurcation diagram is rich enough to display up to five fixed points, allowing to reproduce pulses with the desired shapes close to the bifurcation curves [19]. We do not make claims on a neural interpretation of this dynamical system in terms of how these gestures are generated in a bird’s brain. Our construction is dynamic in nature, and used to emphasize the interplay between basic common gestures and learned combinations in this spectacularly rich behavior.

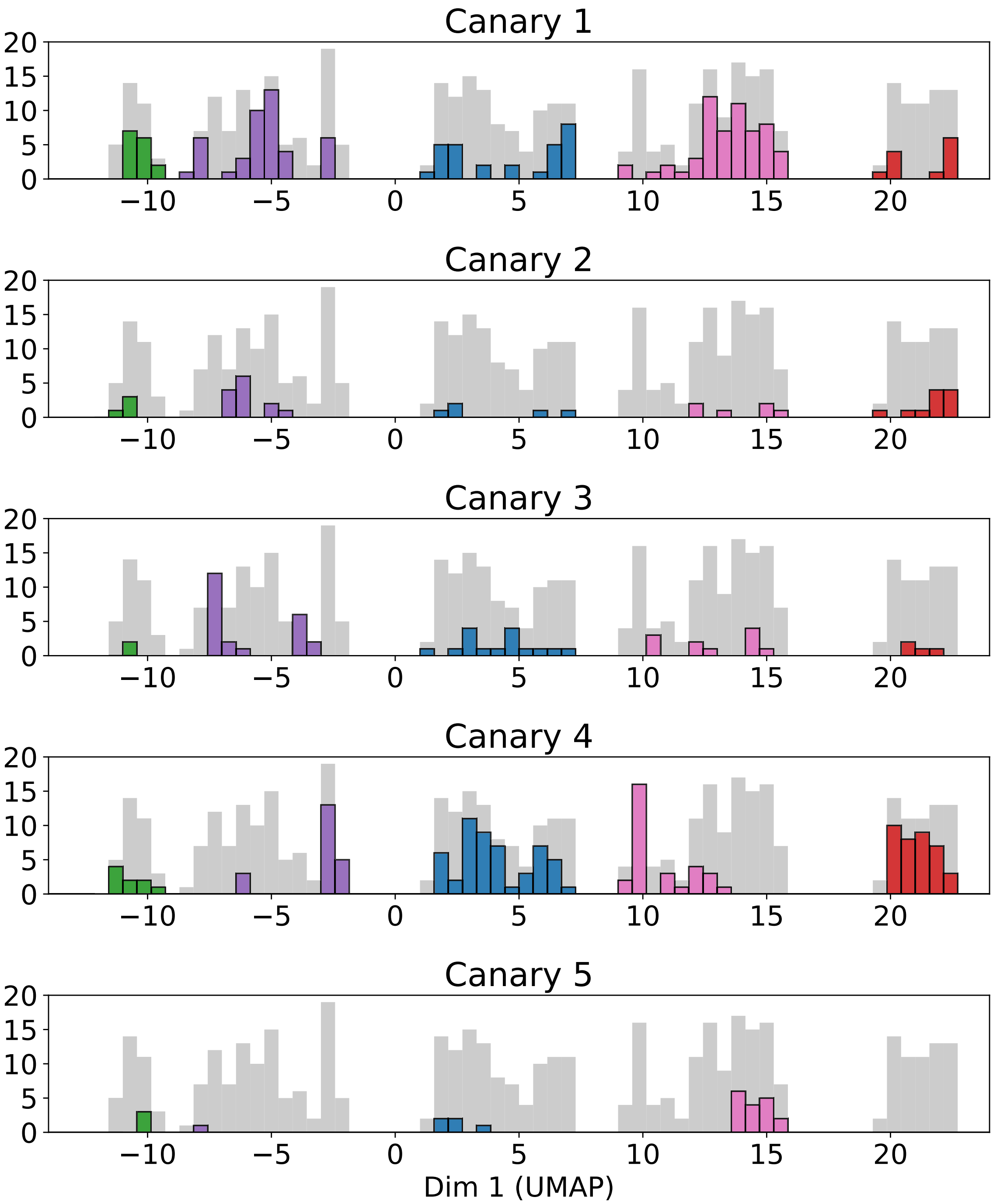


Fig. 5. Individual-level gesture usage compared to the general distribution. Each panel shows the histogram of UMAP projections for one canary (colored bars) overlaid on the distribution across all birds (gray background). It is seen that gestures are not clustered by individuals, since the distribution is heterogeneous. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Agustin Carpio Andrada: Writing – original draft, Software, Investigation, Formal analysis, Data curation. **Gabriel B. Mindlin:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The work was partially funded by UBACyT, Universidad de Buenos Aires, Argentina.

References

[1] Hartley RS. Expiratory muscle activity during song production in the canary. Respir Physiol 1990;81(2):177–87.
[2] Suthers RA. Peripheral control and lateralization of birdsong. J Neurobiol 1997;33:632–52.
[3] Suthers RA, Goller F. Motor correlates of vocal diversity in songbirds. In: Nolan V, Ketterson ED, Thompson CF, editors. Current ornithology. vol. 14. Boston, MA: Springer; 1997. p. 235–88. https://doi.org/10.1007/978-1-4757-9915-6_6.
[4] Wild JM. Neural pathways for the control of birdsong production. J Neurobiol 1997;33(5):653–70. https://doi.org/10.1002/(sici)1097-4695(19971105)33:5<653::aid-neu11>3.0.co;2-a.
[5] Gardner T, Cecchi G, Magnasco M, Laje R, Mindlin GB. Simple motor gestures for birdsongs. Phys Rev Lett 2001;87(20):208101. https://doi.org/10.1103/PhysRevLett.87.208101.
[6] Suthers RA, Margoliash D. Motor control of birdsong. Curr Opin Neurobiol 2002;12(6):684–90.
[7] Döppler JF, Bush A, Goller F, Mindlin GB. From electromyographic activity to frequency modulation in zebra finch song. J Comp Physiol A 2018;204(2):209–17. https://doi.org/10.1007/s00359-017-1231-3.
[8] Bush A, Döppler JF, Goller F, Mindlin GB. Syringeal EMGs and synthetic stimuli reveal a switch-like activation of the songbird's vocal motor program. Proc Natl Acad Sci U S A 2018;115(37):9340–5. https://doi.org/10.1073/pnas.1801251115.
[9] Prescott SA. Excitability: types I, II, and III. In: Jaeger D, Jung R, editors. Encyclopedia of computational neuroscience. New York: Springer; 2013. p. 1–7.
[10] Alliende JA, Méndez JM, Goller F, Mindlin GB. Hormonal acceleration of song development illuminates motor control mechanism in canaries. Dev Neurobiol 2010;70(14):943–60. https://doi.org/10.1002/dneu.20835.
[11] Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 1964;36(8):1627–39. https://doi.org/10.1021/ac60214a047.
[12] Wilson HR, Cowan JD. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 1973;13:55–80.
[13] Hoppensteadt FC, Izhikevich EM. Weakly connected neural networks. In: Applied mathematical sciences. vol. 126. New York: Springer; 2012.
[14] Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 1997;11:341–59.
[15] McInnes L, Healy J. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv 1802.03426. 2018.
[16] Rand WM. Objective criteria for the evaluation of clustering methods. J Am Stat Assoc 1971;66(336):846–50. https://doi.org/10.1080/01621459.1971.10482356.
[17] Trevisan MA, Mindlin GB, Goller F. Nonlinear model predicts diverse respiratory patterns of birdsong. Phys Rev E 2006;74:021908. https://doi.org/10.1103/PhysRevE.74.021908.
[18] Alonso RG, Trevisan MA, Amador A, Goller F, Mindlin GB. A circular model for song motor control in Serinus canaria. Front Comput Neurosci 2015;9:41. https://doi.org/10.3389/fncom.2015.00041.
[19] Carpio Andrada A. Dinámica compleja en sistemas con retroalimentación retrasada y sus aplicaciones en la integración sensoriomotora. Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales; 2022 [Tesis de Grado].