REVIEW

The dynamics behind diversity in suboscine songs

Ana Amador[1,2,*] and Gabriel B. Mindlin[1,2]

ABSTRACT

Vocal behavior plays a crucial evolutionary role. In the case of birds, song is critically important in courtship, male–male competition and other key behaviors linked to reproduction. However, under natural conditions, a variety of avian species live in close proximity and share an 'acoustic landscape'. Therefore, they need to be able to differentiate their calls or songs from those of other species and also from those of other individuals of the same species. To do this efficiently, birds display a remarkable diversity of sounds. For example, in the case of vocal learners, such as oscine passerines (i.e. songbirds), complex sequences and subtle acoustic effects are produced through the generation of complex neuromuscular instructions driving the vocal organ, which is remarkably conserved across approximately 4000 oscine species. By contrast, the majority of the sister clade of oscines, the suboscine passerines, are thought not to be vocal learners. Despite this, different suboscine species can generate a rich variety of songs and quite subtle acoustic effects. In the last few years, different suboscine species have been shown to possess morphological adaptations that allow them to produce a diversity of acoustic characteristics. Here, we briefly review the mechanisms of sound production in birds, before considering three suboscine species in more detail. The examples discussed in this Review, integrating biological experiments and biomechanical modeling using non-linear dynamical systems, illustrate how a morphological adaptation can produce complex acoustic properties without the need for complex neuromuscular control.

Introduction

Birds can produce an impressive diversity of sounds, including tones, clicks, trills and screeches. In this Review, we will discuss some examples of how non-linear mechanisms contribute to the complexity of birdsong. We briefly touch on song production in oscine birds (i.e. songbirds; see Glossary), before focusing our attention on their sister clade, the suboscine birds (see Glossary). Both clades belong to the order Passeriformes. Suboscine birds include more than 1100 species (Tobias et al., 2012), whereas there are approximately 4000 oscine species. In contrast to oscines, which are vocal learners, suboscine passerines typically develop normal songs without learning. Several lines of evidence support the lack of vocal learning in suboscines: (1) suboscines display a general absence of vocal dialects (i.e. vocal variations across the geographical distribution of a species); (2) acoustic deprivation experiments show that their songs are largely innate (Kroodsma, 1984, 1985; Kroodsma and Konishi, 1991; Touchton et al., 2014); and (3) neuroanatomical research demonstrates that some species of suboscines lack the forebrain song nuclei used in vocal learning by oscine passerines (e.g. Gahr et al., 1993; Liu et al., 2013). However, there are some exceptions reported (e.g. Saranathan et al., 2007; Kroodsma et al., 2013). Thus, the suboscines provide the opportunity to study mechanisms of vocal differentiation that may occur without the vocal flexibility achieved through learning.

Birdsong produced by oscine birds is an ideal model for studying how a brain reconfigures itself during learning in order to achieve the complex motor tasks involved in vocal communication. The avian vocal organ, the syrinx (see Glossary), is quite similar across oscine species (Stein, 1968; Ames, 1971). The broad range of vocalizations displayed by these species is therefore believed to be associated with vocal learning and the complex and subtle neural control of the vocal organ achieved in the process. There has been extensive effort to unveil the nature of the neural coding of song in the oscine brain, where cortical and brainstem structures interact extensively in the generation of the instructions sent to the periphery in order to produce the song (e.g. Amador et al., 2017; Mooney, 2022; Sakata et al., 2020; Zeigler and Marler, 2008).

During avian vocalizations, the neural instructions engage the respiratory pathway in order to achieve the level of air sac pressure needed to generate sounds. Expiratory muscles compress the air sacs so air flows through the trachea and the vocal sources (see Fig. 1A). Oscine birds can produce morphologically rich pressure patterns, which are crucial in determining the tempo, timing, as well as – to some degree – the modulation of acoustic features (Hartley and Suthers, 1989; Goller and Suthers, 1996a; Mindlin et al., 2003; Amador and Margoliash, 2013). The neural instructions generated by the oscine brain during singing also reach the muscular apparatus acting on the syrinx. Fig. 1B shows a typical oscine syrinx with its complex set of muscles, allowing fine control of the sound source. Much research has focused on the action of the muscles determining the configuration of the syrinx and, therefore, the acoustic features of the sound (Goller and Suthers, 1996a,b; Suthers and Zollinger, 2004; Düring et al., 2013, 2017). This research suggests that much of the complexity in the song of oscine birds is rooted in the richness of the neuromuscular control (Amador et al., 2017).

For the suboscines, the situation is quite different. Unlike their sister group, the suboscines possess a less complex set of syringeal muscles but display remarkable morphological diversity of the syrinx (Ames, 1971; King, 1989; Goller et al., 2021). Recent studies on different suboscine species suggest that diversification of acoustic features in their vocal repertoires might have occurred through different morphological adaptations or adaptations of the way in which the elements of the vocal organ are used (Garcia et al., 2017; Goller et al., 2021). In the absence of vocal learning, such adaptations allow suboscines to generate a wide range of acoustic features. This is in contrast to oscine birds, where such diversity is mainly achieved through diverse and rich neuromuscular control as the vocal organ is highly conserved across oscine species.

[1]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física, Ciudad Universitaria, 1428, Buenos Aires, Argentina. [2]CONICET - Universidad de Buenos Aires, Instituto de Física Interdisciplinaria y Aplicada (INFINA), Ciudad Universitaria, 1428, Buenos Aires, Argentina.

*Author for correspondence (anita@df.uba.ar)

A.A., 0000-0002-1040-2359; G.B.M., 0000-0002-7808-5708

Glossary

Acoustic resonance
A phenomenon in which an acoustic system amplifies sound waves whose frequency matches one of its own natural frequencies of vibration (known as the resonance frequency).

Bandpass filter
A filter that allows a specific range or 'band' of frequencies to pass through while attenuating frequencies outside this range. The passband of the filter is defined by its center frequency and the bandwidth, which specifies the range of frequencies that are allowed to pass through the filter.

Bernoulli's principle
In fluid dynamics, an increase in the speed of a fluid occurs simultaneously with a decrease in pressure or a decrease in the fluid's potential energy. For vocal production in humans and birds, as the air passes through the very narrow opening between the membranes in the vocal tract, it must accelerate to get through. This high-speed air creates a suction effect, bringing the membranes together.

Fundamental frequency
The lowest frequency of a periodic waveform.

Helmholtz resonator
An acoustic resonator consisting of a cavity, an opening and a neck, which is a small channel connecting the cavity to the opening. The speed of the sound in the medium and the resonator geometry (volume of the cavity, length and diameter of the neck) will define the value of the resonant frequency.

Non-linear oscillator
An oscillator in which the restitution force is a non-linear function.

Oroesopharyngeal cavity
Air cavity present in birds consisting of the oral cavity, the esophagus and the pharynx.

Oscine birds
Those belonging to the suborder Passeri (from the order Passeriformes), which includes most songbirds. Birds belonging to this suborder are found all over the world, in a wide range of habitats and are characterized by a highly developed vocal apparatus and by vocal learning.

Passive sound filter
A device that highly attenuates frequencies from an audio signal without the use of any energy input (or external power source in the case of electronic devices).

Suboscine birds
Those belonging to the suborder of passerine birds that are closely related to the oscine birds, but have simpler vocal organs and mostly do not learn their vocalizations. Suboscine birds are found mainly in Central and South America.

Syrinx
The vocal organ of birds. It is located at the base of the trachea. The specific shape, structure and musculature vary across bird groups, but in all of them the biophysical mechanism of sound production is oscillating membranes that modulate the airflow.

Timbre
The quality of a sound that distinguishes it from other sounds of the same fundamental frequency and loudness. It is often described as the 'tone color' of a sound. The physical characteristics of sound that determine the perception of timbre are multidimensional and include frequency spectrum and envelope.

The comparative study of these phylogenetically close groups allows us to study the evolutionary impact of the presence or absence of vocal learning.

Here, we first discuss the mechanisms of vocal production in oscines, before considering some case studies in suboscine birds. When discussing suboscines, we focus on one specific acoustic feature of the vocalizations: slow soundwave modulations coexisting with the high-frequency oscillations that are typical of avian vocalizations (kilohertz range). As we will show, different

suboscine species achieve this acoustic feature through a variety of anatomical strategies or by taking advantage of the non-linearity of the vocal source. The overall aim of this Review is to highlight how suboscine species use biomechanical adaptations to generate very rich and diverse songs.

Biophysical mechanisms for vocal production in birds

Until the late 1990s, there was debate regarding the phonation mechanisms of songbirds (see Box 1). However, nowadays there is consensus on the idea that singing oscine birds use the same physical mechanism as humans when producing voiced sounds (Mindlin and Laje, 2006; Riede and Goller, 2010; Titze and Martin, 1994): the oscillation of soft tissues modulates the air flow, generating sound that is then filtered by the upper vocal tract. The syrinx of oscine birds contains two sets of labia, i.e. two pairs, each consisting of the lateral labium and medial labium (LL and ML in Fig. 1C,D). When the bird is breathing silently, the labia are so far apart that it is physically impossible to induce their oscillation with the expiratory airflow (Fig. 1C). In order to sing, the bird activates syringeal muscles that rotate the third cartilage in such a way as to be carried inward. In this way, the lateral labia are positioned in the bronchial lumen in a 'pre-phonatory' position (Goller and Larsen, 1997; Suthers et al., 1999; Larsen and Goller, 2002; Fig. 1D). The closeness of the labia allows auto-sustained oscillations to begin when the air flow reaches a sufficiently high pressure, thus generating sound (for more details, see Mindlin and Laje, 2006). In this way, sound is generated as the airflow is modulated by the labial oscillations. The pressure fluctuations are then filtered by the trachea, the oroesopharyngeal cavity (OEC; see Glossary) and the beak. Riede, Suthers and collaborators showed that the OEC is dynamically adjusted in order to emphasize the fundamental frequency (see Glossary) in some songbirds, thus generating more 'tonal' sounds (e.g. Riede et al., 2006; Riede and Suthers, 2009). Songbirds can very rapidly modulate the fundamental frequency of their sounds; thus, the underlying volume changes in the OEC need to occur very fast. This is observed for the tonal songs of northern cardinals (*Cardinalis cardinalis*) and white-throated sparrow (*Zonotrichia albicollis*), for example. In these birds, fine motor control of the syringeal muscles is coordinated with fine motor control of the upper vocal tract. An alternative type of song is generated by zebra finches (*Taeniopygia guttata*); these songs contain syllables that have a wide spectral content. For these birds, frequencies between 2.5 and 5 kHz are generally emphasized, presumably as a result of the filtering properties of the OEC (Riede et al., 2013).

The vocalizations of oscine species contain a wide range of timbres (see Glossary) that are compatible with the mechanisms described above. But as we explore outside the oscine world, we find a variety of strategies to control the vocal organ, and anatomical adaptations that further expand the timbrical landscape. Below, we describe three examples of sounds produced by suboscines that illustrate different control mechanisms used during song production, all of which involve oscillating membranes. These examples show similarities and differences when compared with song production by oscine birds. Because oscines and suboscines are so closely related, these comparisons may shed light on the evolutionary strategies used to achieve variability in birdsong production in the absence of learning.

Three suboscine stories

Below, we review three examples of passerine suboscine vocalizations: (1) the white-tipped plantcutter (*Phytotoma rutila*),
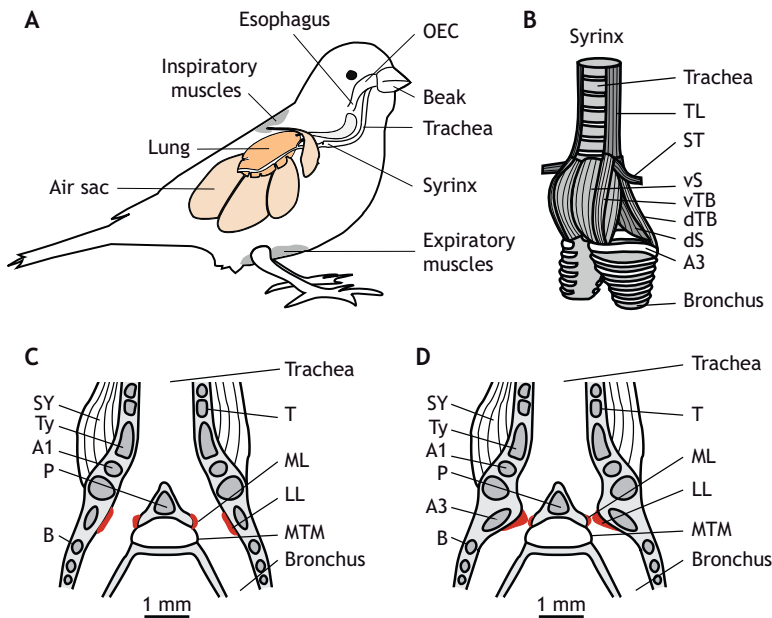
Fig. 1. Respiratory system and vocal organ of songbirds.
(A) Songbird respiratory system showing the respiratory muscles (inspiratory and expiratory), lungs, air sacs and vocal filters: the trachea, oroesopharyngeal cavity (OEC) and beak. The syrinx is located between the trachea and the bronchi. Modified with permission from Fainstein et al. (2021). (B) External ventrolateral view of a songbird syrinx indicating the syringeal muscles. (C,D) Schematic ventral views of a songbird syrinx in quiet respiratory (C) and phonatory (D) configurations. During vocalization, the medial labia (ML) and lateral labia (LL) are set into vibration mode when they are adducted into the expiratory air stream. Cartilage components of the syrinx include three tracheo-bronchial semi-rings (A1–A3) and the tympanum (Ty). B, bronchial cartilage; dS, m. syringealis dorsalis; dTB, m. tracheobronchialis dorsalis; MTM, medial tympaniform membrane; P, pessulus; ST, m. sternotrachealis; SY, syringeal muscle; T, tracheal cartilage; TL, m. tracheolateralis; vS, m. syringealis ventralis; vTB, m. tracheobronchialis ventralis. (B–D) Modified with permission from Suthers et al. (1999).

(2) the great kiskadee (*Pitangus sulphuratus*) and (3) the tracheophones, a group that includes the rufous hornero (*Furnarius rufus*). As discussed above, when considering suboscine vocalizations, we will focus on one specific acoustic feature: the appearance of sound modulations with frequencies between 100 and 200 Hz coexisting with high-frequency (i.e. of the order of kilohertz) components. This feature is easy to identify in a song, and is present in several suboscine species, making it ideal for comparing across species.

The roughest sounds

The white-tipped plantcutter (Fig. 2A) is a passerine bird belonging to the Cotingidae family. It is distributed throughout south-central South America, particularly from western Bolivia to northern and central Argentina, and also in some parts of Paraguay and Uruguay and extreme southwestern Brazil (Rodríguez-Cajarville et al., 2019). One of the most notable characteristics of this species is its vocalization, which consists of a rough, long note, with a sound comparable to that of a rusty hinge (see spectrogram in Fig. 2B; an example recording is available from www.xeno-canto.org/50074). Fig. 2C shows the initial fragment of the vocalization, which consists of a succession of short sound fragments of increasing speed. The production rate of these fragments is of the order of 100 Hz. A detailed inspection (Fig. 2D) shows that each sound fragment is made of fluctuations of a few kilohertz, the amplitudes of which decrease significantly before the start of the subsequent fragment. This observation is key to suggesting the mechanism behind this vocalization.

In order to explore the possible mechanisms used to generate this sound, dynamical systems models are very helpful. A dynamical systems model can be used to capture the most relevant features of the behavior in a given biomechanical process; it can therefore be used to analyze the viability of any proposed mechanism. Within this framework, a dynamical system consistent with a given biomechanical process can be used to synthesize artificial sounds that can be compared with the sound under study. For example, Fig. 2E shows a sequence of pulses generated by a biomechanical model representing the sound source of the white-tipped plantcutter vocalization. The vocal production of low-rate pulsating sounds

requires simple motor gestures, such as coordinating respiratory muscles that will interact with the tissue membranes of the vocal source (i.e. the syringeal labia in this case). The physical mechanism used to generate pulse-like sound consists of holding the syringeal labia together while increasing the pressure in the air sac system through the activation of the expiratory muscles. In this way, sub-syringeal pressure builds up until the force exerted on the labia can overcome the force exerted by the syringeal muscles that hold the labia together. When the air flow is established, the pressure between the labia decreases (according to Bernoulli's principle; see Glossary), the labia collide against each other, and the process begins again (for further details; see Gardner et al., 2001; Amador and Mindlin, 2008; Mindlin, 2017; Perl et al., 2011). This mechanism has been directly observed in birds during sound generation (Jensen et al., 2007; Goller and Riede, 2013). Note that there is a big difference between a pulse sequence generated by the sound source (Fig. 2E) and the sound displayed in the vocalization (Fig. 2D). This is because of the effect of the OEC (see Fig. 1A), which filters the sound signal generated by the syringeal labia. This effect is illustrated in Fig. 2F, which shows the result of filtering the series of pulses in Fig. 2E with a biomechanical model using a Helmholtz resonator (see Glossary) to represent the OEC. The details of the dynamical systems model used to synthesize the pulses and the acoustic filters of the vocal cavities are described in Uribarri et al. (2020).

Note that it would be possible to produce a vocalization similar to that of the white-tipped plantcutter by an alternative mechanism. In fact, there are numerous species that can induce oscillations in the labia of the order of kilohertz, as well as an oscillation of less than 200 Hz. This pattern of vocalization can result from the action of ultra-fast muscles (Elemans et al., 2008). However, this mechanism is unable to account for the specific decay in the amplitude of each of the sound segments that compose the vocalization of the white-tipped plantcutter. In fact, indirect evidence supports the hypothesis that the fast oscillatory component is due to the effect of a passive sound filter (see Glossary). This evidence has been obtained from the analysis of songs from individuals recorded in different habitats. If the OEC is the main determinant of the passive filtering, variations in animal size would be reflected in changes in the

Box 1. Unravelling the phonation mechanisms of songbirds – the historical context

Sound can be generated through very different dynamical mechanisms and different animal species take advantage of this fact. The phonation mechanisms in birds were debated for many years. Essentially, there were two hypotheses: (1) the sound source behaved as an aerodynamic whistle, or (2) the sound source consisted of oscillating membranes modulating the airflow to generate a sound wave. In both cases, the sound was then filtered by the upper vocal tract. The hypothesis that birds generated sounds by the same physical mechanism used by a whistle was proposed to explain the origin of tonal sounds (i.e. sounds containing a single frequency that is commonly the fundamental frequency; in nature, this is generally achieved by tuning passive acoustic filters to concentrate most of the sound energy in the fundamental frequency, leaving the harmonics with almost no energy). This hypothesis gained strength because there were theoretical difficulties in explaining how a system based on an oscillating membrane could generate sounds with an almost total absence of harmonics (Casey and Gaunt, 1985; Gaunt and Gaunt, 1985; Fletcher, 1988, 1989). It was Stephen Nowicki who, in 1987, provided experimental evidence supporting the idea that a source–filter mechanism could achieve tonal sound production. He studied nine species of songbirds singing in a heliox atmosphere. Heliox is a gas mixture composed of 80% helium and 20% oxygen. Helium is lighter than oxygen, and produces less resistance in the airway, therefore changing the air velocity. This allowed the 'missing' harmonic overtones to be revealed. These harmonics appear when the bandpass filter (see Glossary) provided by the vocal tract, normally tuned to the fundamental frequency of the sound source, shifts upwards as a result of the change of sound velocity in the heliox atmosphere (Nowicki, 1987; Brittain-Powell et al., 1997). It was a compelling yet indirect experiment. There was neither a direct observation of the sound sources during vocal production nor a clear consensus on what tissues were actually responsible for the modulation of the air flow.

In 1997, Goller and Larsen conducted an experiment where the medial tympaniform membrane was damaged (MTM in Fig. 1C,D). This generated only small modifications to the song of zebra finches (*Taeniopygia guttata*) and cardinals (*Cardinalis cardinalis*), which are two oscine species capable of generating very different songs (Goller and Larsen, 1997). Later, using videography, Goller and Larsen (1999) were able to directly confirm that it was the lateral labium membrane together with the medial labium membrane that modulates the air flow to generate sound (ML and LL in Fig. 1D). This new conception of sound generation in birds was strengthened by long-term experiments (Goller and Larsen, 2002), working with various species, together with new mathematical models (Gardner et al., 2001; Mindlin and Laje, 2006; Mindlin et al., 2003) that explained how tonal syllables could be generated through a membrane vibration mechanism (Suthers and Margoliash, 2002).

vocalization frequency, because commensurate changes in OEC size would modify its resonant frequency (see 'acoustic resonance' in Glossary). In Uribarri et al. (2020), the frequencies in the kilohertz range were analyzed from different individuals from the xeno-canto database (https://www.xeno-canto.org/) and the Macaulay Library at the Cornell Lab of Ornithology (https://www.macaulaylibrary.org/). For each vocalization, the mean peak frequency was noted, along with the location where the vocalization was recorded. A simple analysis showed a linear correlation between mean peak frequency (range: 3.5–4.1 kHz) and habitat altitude (range: 0–3500 m above sea level). This is significant because of Bergmann's rule (Meiri and Dayan, 2003), which states that populations of larger size are found in colder environments.

Working with the hypothesis that the high frequencies of the vocalizations are the result of the OEC filter, it is possible to use

scaling arguments to estimate how the resonant frequencies of the filter should change with the size of the individual. Modeling the OEC as a Helmholtz resonator, its resonant frequency depends on its volume, the area and the neck of the entrance as:

$$\omega = v_{sound}\sqrt{\frac{A}{LV}}, \qquad (1)$$

where v_{sound} is the sound velocity, and A, L and V represent geometrical values of the cavity: V is the OEC volume, A and L are the area and length of the resonator's neck. In this way, if the linear size of the animal scales with $\lambda \approx (1+\epsilon)$ with $\epsilon \ll 1$, the corresponding geometric quantities will scale accordingly: the area A with λ^2, the length L with λ and the volume V with λ^3. Replacing this in Eqn 1, the frequency will change as:

$$\omega \rightarrow \omega\sqrt{\frac{\lambda^2}{\lambda\,\lambda^3}} = \frac{\omega}{\lambda} \approx (1-\epsilon)\omega. \qquad (2)$$

And therefore:

$$\frac{\Delta\omega}{\omega} \approx -\epsilon. \qquad (3)$$

In other words, an ϵ percent change in the size of the animal will translate into an ϵ percent change in the frequency emphasized by the upper vocal tract.

To estimate body size, in Uribarri et al. (2020), lengths of tarsometatarsus bones from museum specimens were analyzed. The tarsometatarsus (often referred to as 'tarsus') is a bone that is found in the lower leg of birds and some non-avian dinosaurs (Proctor et al., 1993), and its size correlates with animal size. In Uribarri et al. (2020), it was found that the tarsi lengths correlated with the altitudes at which the specimens were collected. For animals captured between 0 and 3500 m above sea level, the change in size was approximately 13%. This result is consistent with Bergmann's rule (Rodríguez-Cajarville et al., 2019). Then, the frequencies of vocalizations were analyzed, and a linear correlation was found between the resonant frequency and altitude. The change in the whole range of the mean peak frequency was 13.5%. Beyond building confidence in the biomechanical model and in the fact that the high frequencies in these vocalizations are the result of filtering signals from a sound source with a pulsatile structure, this result suggests that this species is capable of reliably transmitting information about body size through acoustic properties. This may not be the case for many oscine species, for which the correlation between acoustic features and size tends to be weak and variable (e.g. Cardoso, 2012; Linhart and Fuchs, 2015; Liu et al., 2017). Moreover, songbirds use their ability to adjust filter properties dynamically, e.g. by changing the OEC volume during singing (e.g. Riede et al., 2006), which obscures the relationship between vocalization frequency and size.

It is noteworthy that the high frequencies of the white-tipped plantcutter call do not show significant modulations during vocalization. In other words, the filter is kept constant while a smooth modulation occurs in the rate of generation of sound segments; that is, of the pulses that will be filtered by the cavity. In other suboscine species, this modulation is generated by varying the pressure of the air sacs (Amador et al., 2008), which is consistent with the observation that in the white-tipped plantcutter, the first and last pulses are those generated at a lower firing rate. In this way, Uribarri et al. (2020) show that in the white-tipped plantcutter, the low frequency is associated with the series of pulses generated by
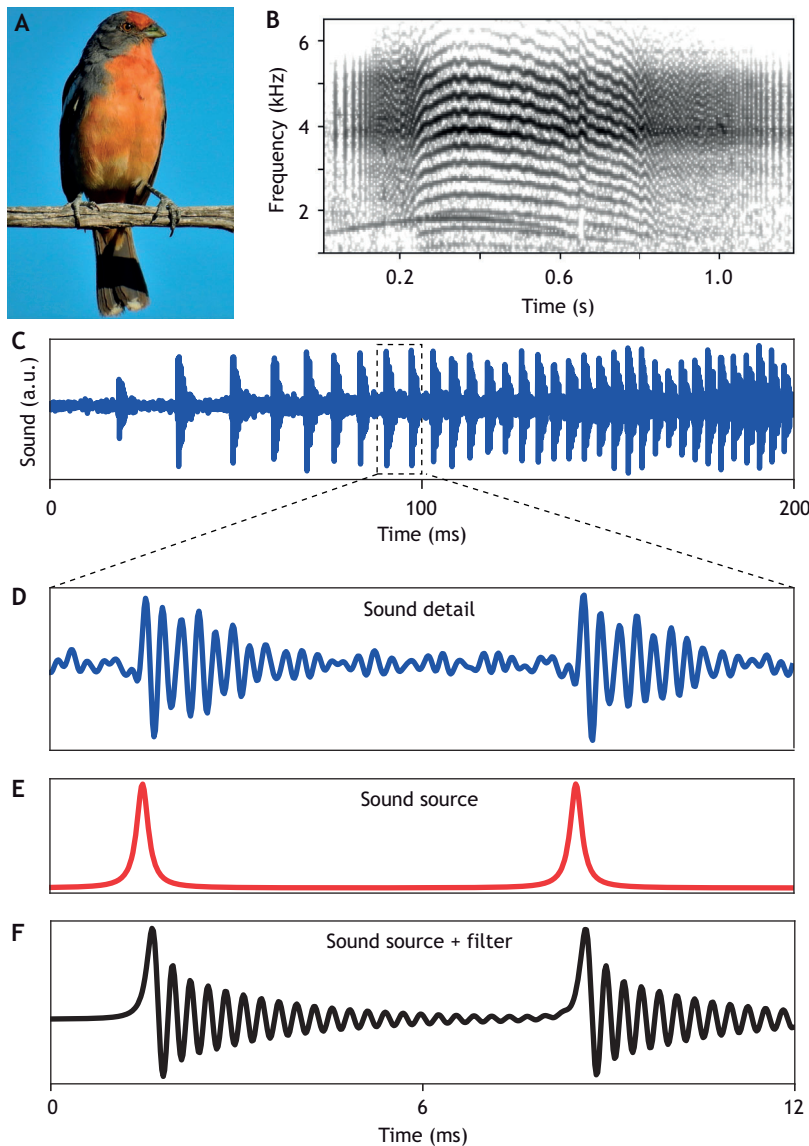
Fig. 2. White-tipped plantcutter, its song and a dynamical hypothesis. (A) The white-tipped plantcutter (*Phytotoma rutila*). (B) Spectrogram of a *P. rutila* song. (C) The waveform of the initial 200 ms segment of the song. a.u., arbitrary units. (D) A detail of two sound segments. A rapidly decaying fast oscillation is compatible with a pulse being filtered by a dissipative resonator. (E) A pulse, as synthesized by a dynamical systems model. (F) Time traces representing pulses, filtered by a damped oscillator of the appropriate resonant frequency. Photo credit: Pablo Alejandro Pla (Macaulay Library at the Cornell Lab of Ornithology, ML80414721). Figure adapted with permission from Uribarri et al. (2020).

the syringeal labia, produced at a very low rate that correlates with the air sac pressure; by contrast, the high-frequency component is due to the passive filtering of these pulses by the OEC.

Perfectly out of tune

The great kiskadee (Fig. 3A) is another example of a suboscine species that produces songs in which low-frequency components coexist with frequencies of the order of kilohertz (example recordings available from www.xeno-canto.org/272848). Here, we will analyze its song and the dynamical origin of these frequency components.

The great kiskadee is a species of passerine bird belonging to the Tyrannidae family. It is native to tropical America (Neotropics), and widely distributed from southern USA to central Argentina. In 2008, a rather surprising result was published: the modulation of the fundamental frequency (see Glossary) of the great kiskadee song does not seem to require the activation of syringeal muscles (Amador et al., 2008). Indeed, after the syringeal muscles were inactivated, by sectioning both branches of the tracheosyringeal nerve, the birds did not show drastic changes in their songs. Specifically, the fundamental frequency modulations of the

vocalizations remained unchanged. The result was surprising, because in oscine birds, the syringeal muscles play an important role in controlling the phonology of the vocalizations. In a foundational piece of work, Goller and Suthers (1996b) showed that electromyographic (EMG) activity in the muscle syringealis ventralis (vS in Fig. 1B), the largest syringeal muscle, increased exponentially with the fundamental frequency of the generated sound and also correlated with the frequency modulation. Moreover, if syringeal muscles are inactivated in oscine birds, the songs suffer major acoustic changes and birds may have difficulty breathing (Goller and Cooper, 2004; Bhama et al., 2011). Although the functions of the different muscles seem to be more complex than indicated in the seminal studies (e.g. Düring et al., 2013), the activity of the ventral syringeal muscles correlates with the fundamental frequency in oscine birds (Goller and Suthers, 1996a,b; Suthers et al., 1999). By contrast, the work of Amador and collaborators (2008) with great kiskadees found something very different: a linear correlation between the modulation of the fundamental frequency and the modulation of the air sac pressure measured during singing. This suggested a novel mechanism for pitch modulation during song production in birds. Some years later,
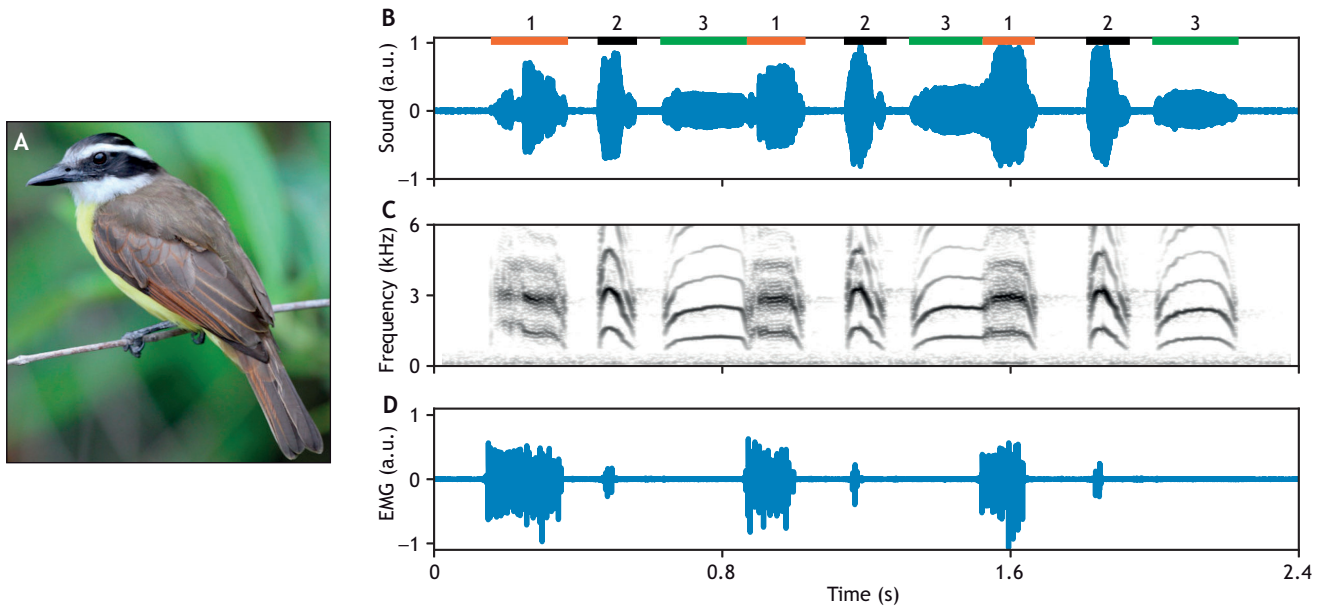
Fig. 3. Simultaneous measurement of sound and electromyographic (EMG) activity of the syringeal muscle during three renditions of the great kiskadee song. (A) The great kiskadee (*Pitangus sulphuratus*). (B) The sound waveform of the song. Numbers indicate syllable identification within a song. (C) Spectrogram of the sound in B. Notice the spectral richness of syllable 1 in each repetition of the song. (D) Simultaneous EMG activity of the syringeal muscle (the obliquus ventralis muscle, ovm) recorded during singing. The most prominent activity occurs during syllable 1, whereas none is present during syllable 3. Photo credit: Alex Wiebe (Macaulay Library at the Cornell Lab of Ornithology, ML71601401). a.u., arbitrary units. Figure adapted with permission from Döppler et al. (2020).

the same mechanism was found in zebra finches (Amador and Margoliash, 2013), which, in comparison with great kiskadees, have a much more complex syrinx musculature (see Fig. 1B) involved in the phonology of the vocalizations. Moreover, the mechanism for frequency modulation through pressure is also found in humans (Titze, 1989). Thus, pitch modulation through modulation of the respiratory pressure potentially represents a general mechanism for frequency modulation when sound is generated by oscillating membranes. The linear correlation between air sac pressure and the fundamental frequency of the vocalizations explains how pitch could be modulated with no activation of the syringeal muscles. This result is particularly intriguing as great kiskadees possess large ventral syringeal muscles. What, then, is their role?

To answer this question, Döppler et al. (2020) recorded vocalizations simultaneously with EMG activity in the syringeal muscles of great kiskadees (more precisely in the obliquus ventralis muscle; ovm). Fig. 3B–D shows three repetitions of the great kiskadee song. Fig. 3B shows a soundwave of the song recorded with a microphone, Fig. 3C shows its spectrogram, and Fig. 3D shows the corresponding electrical activity in the ovm. The most noticeable muscle activity (Fig. 3D) appears simultaneously with a subtle acoustic property: a 'roughness' in the fundamental frequency (see syllable 1 in Fig. 3C). This acoustic property results from a low-frequency modulation of the amplitude of the sound signal, as shown in Fig. 4. The simultaneous study of the EMG activity shows a pattern of activity occurring at the same rate as the modulation of the sound envelope. However, the action of the ovm cannot be solely responsible for this acoustic feature: after resection of the tracheosyringeal nerve (which inactivates the ovm), the amplitude of the modulations decreases, but this feature does not disappear (see details in Döppler et al., 2020).

In order to test the hypothesis that roughness is actively controlled by the ovm during song production in kiskadees, a dynamical systems model was developed (Döppler et al., 2020). Tyrannid

suboscines have two sound sources, each one at the end of the bronchi, where they meet to form the trachea (Ames, 1971). Each sound source contains membranes (labia) that can perform auto-sustained oscillations if enough energy is fed into the system by the air flow of an expiratory pulse. Consistent with previous studies, it was assumed that each sound source could be modelled as a non-linear oscillator (see Glossary; for detailed explanations and equations, see Amador et al., 2008; Mindlin and Laje, 2006). The properties of the modelled oscillators allow slow modulation of the air sac pressure to be transduced into low-frequency modulation of the song (see Döppler et al., 2020, for further details). The model also assumed (1) a coupling between the two sound sources, as the medial labia are supported by a common bone structure, and (2) slightly different elastic properties for each of the two labia. Thus, the physical architecture of the kiskadee vocal source shows important similarities with the oscine syrinx. Yet, its control is quite unique. Döppler et al. (2020) tested the hypothesis that it was possible to reproduce all the acoustic and physiological properties of the kiskadee song by assuming two out-of-tune sound sources driven at a specific frequency (see Doppler et al., 2020, for modeling details).

The biomechanical model presented in Döppler et al. (2020) allowed several observations to be accounted for. If the parameters under which each of the sources operated were slightly different, the model could explain why it is possible that the modulations appear in the first syllable, as it is this syllable that is generated with the highest values of air sac pressure. In this dynamical system, the frequencies of the oscillators depend on the pressure: the higher the pressure, the greater the difference between the frequencies of each oscillator. When two oscillators are coupled, the smaller the frequency difference between them, the easier it is to keep them synchronized. Thus, based on this model, the first syllable, and a small segment of the second syllable, are likely to show an amplitude modulation, because of the difficulty in synchronizing
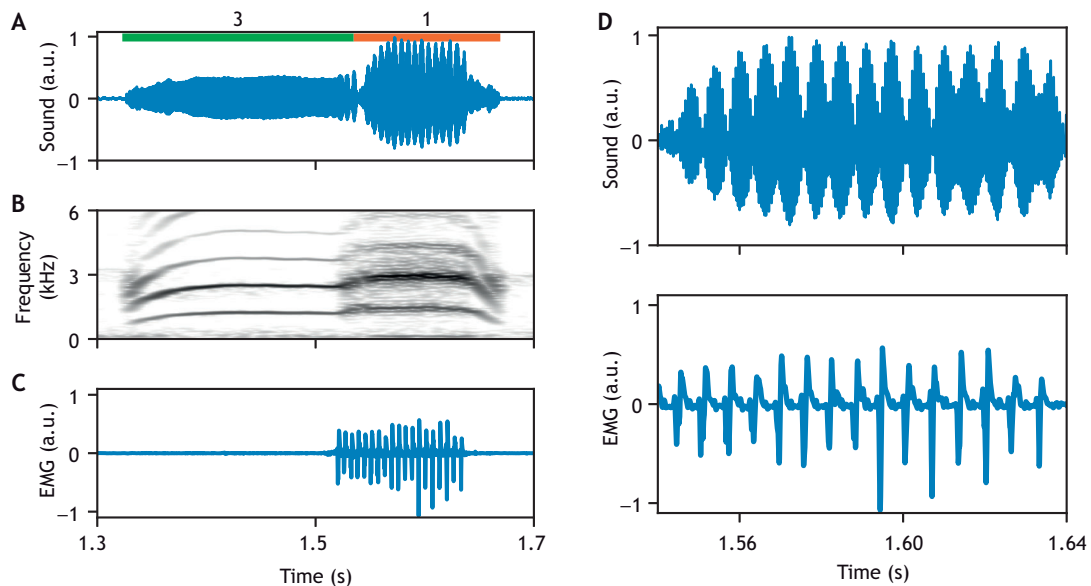
Fig. 4. Comparison of syllables 1 and 3 of the great kiskadee song. (A) The soundwave amplitude is modulated during production of syllable 1. (B) This modulation is observed as a more complex spectrum of the sound. (C) EMG activity of the ovm consists of a series of peaks at a frequency similar to that of the amplitude modulation. This activity is only present during syllable 1. (D) Detail of the sound and EMG activity during syllable 1 to show the correspondence between the EMG pulse and soundwave modulation. Figure adapted with permission from Döppler et al. (2020).

two oscillators with very different natural frequencies. These modulations, a product of the lack of synchrony between the sound sources, are emphasized by the action of the ovm. In this way, the results from experiments and mathematical simulations suggest that the ovm produces a controlled detuning during the execution of the first syllable of the song.

Three sound sources

Within the suboscines, a group of birds called the tracheophones possess a vocal organ characterized by having a pair of membranes, one ventral and one dorsal, in the trachea, just above where the bronchi meet to merge into the trachea (Ames, 1971; Garcia et al., 2017). These tracheal membranes, the membrana trachealis (MT; shown in Fig. 5), were thought to constitute the sound source of their vocalizations, mostly for anatomical reasons (Rüppell, 1933).

However, we now know that the mechanism underlying the vocalizations of tracheophones involves more complex dynamics. Garcia et al. (2017) showed that tracheophones actually have three sound sources: in addition to the MTs, these birds have two pairs of labia at the junction of the bronchi and trachea (see the two pairs of bronchial labia in Fig. 5; BL). These three pairs of oscillating membranes anatomically constitute three sound sources. Garcia et al. (2017) designed a set of experiments to test the hypothesis that all three sound sources are active during vocalizations. They used a fiberscope for direct visualization of the membranes while air was injected into the respiratory system to induce phonation (Fig. 5B). Under certain flow conditions, oscillations were induced in the MTs. In addition to these membranes, the BL were thought to also participate in phonation. To test this hypothesis, Garcia et al. (2017) used two strategies. Firstly, while injecting air into the respiratory
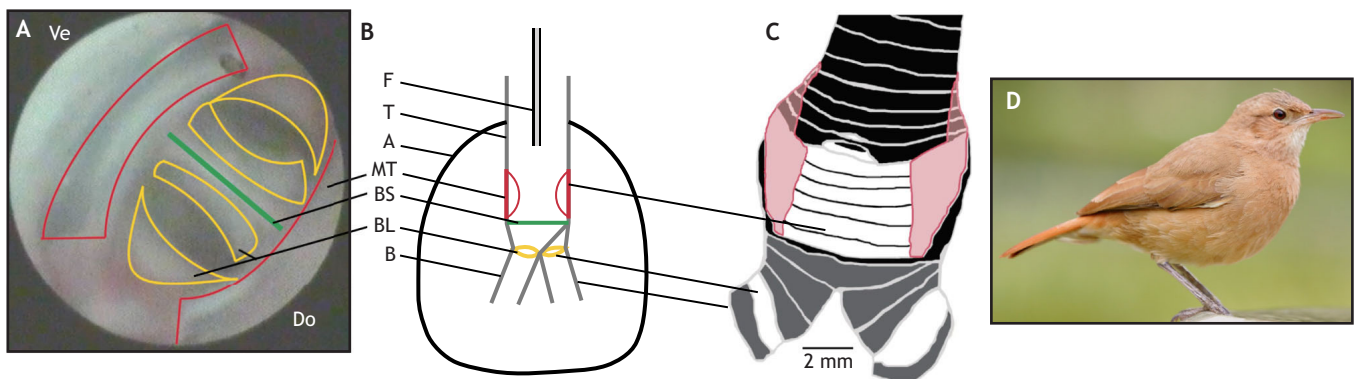
Fig. 5. Tracheophone syrinx: rufous hornero as an example. A fiber optic cable is surgically inserted into the trachea to internally visualize the syrinx and identify sound sources. Relevant structures are labeled and localized showing an image during quiet respiration (A), a schematic diagram of a lateral view of the tracheobronchial junction (B) and a drawing of a ventral view of the rufous hornero syrinx (C), where the syringeal muscles are colored in pink. (D) The rufous hornero (*Furnarius rufus*). Ve, ventral; Do, dorsal; F, fiber optic cable; T, trachea; A, air sac; MT, membrana trachealis; BS, bronchial septum; BL, bronchial labia; B, bronchus. Photo credit: Paulo Gusmão [Macaulay Library at the Cornell Lab of Ornithology (ML79992301)]. Figure adapted with permission from Garcia et al. (2017).
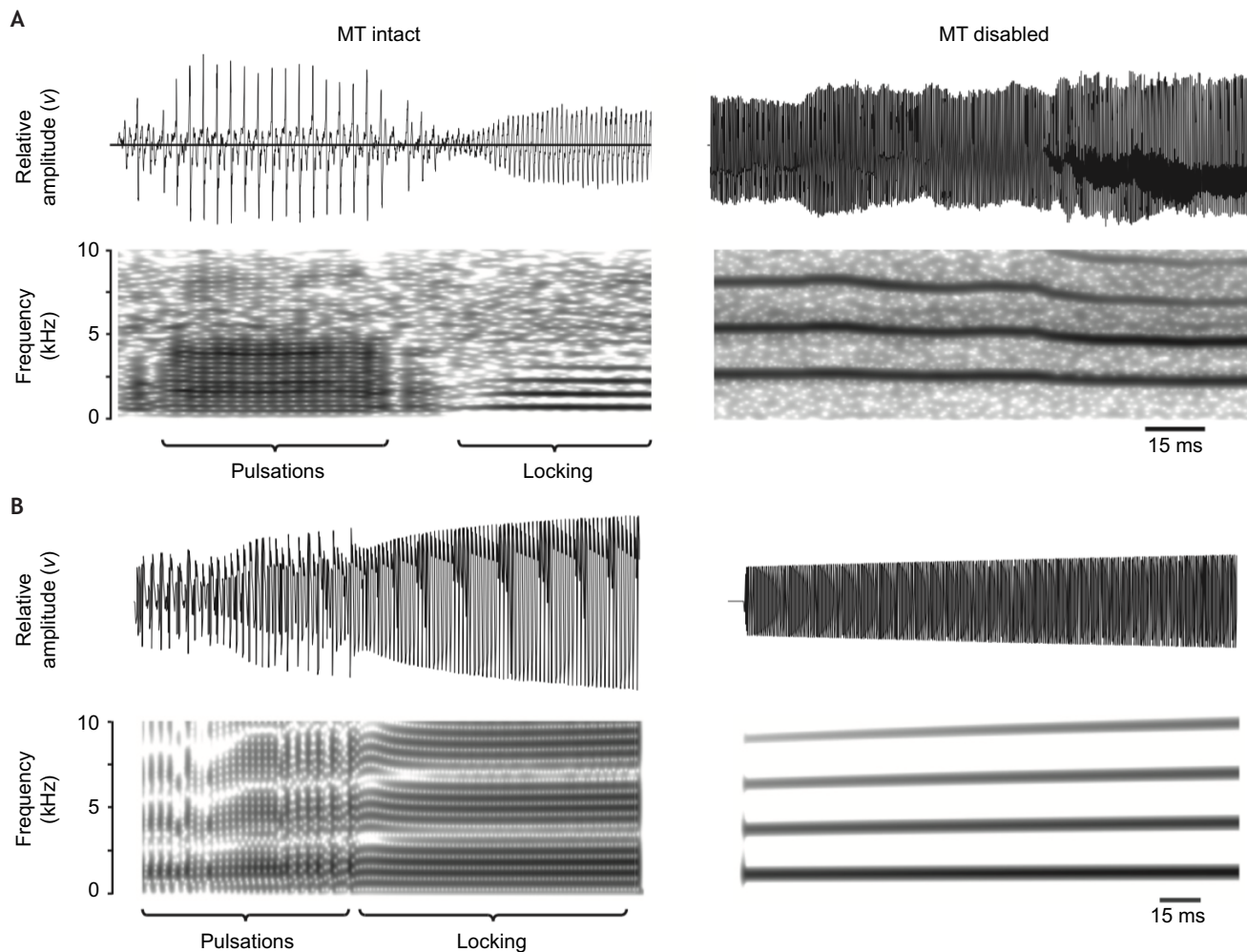
Fig. 6. Fiberscopic manipulation of the ventral MT reveals its role in tracheophone sound production. Examples of (A) induced phonation, and (B) mathematically modeled sound before (MT intact) and after (MT disabled) manipulation of the MT in an anesthetized rufous hornero. In each case, the sound wave is shown at the top and its spectrogram is shown at the bottom. Note the absence of the pulsations in the induced sound after membrane manipulation, as well as an increase in the fundamental frequency. Mathematical modeling of the MTs and a labial sound source interacting to produce sound (B) makes two clear predictions: (1) when the MTs oscillate at a frequency that differs from that of the bronchial sound sources, amplitude modulation will occur; and (2) when the natural frequency of the MTs is similar to that of the fundamental frequency of the bronchial sound sources, they will lock and oscillate at an intermediate frequency. Figure adapted with permission from Garcia et al. (2017).

system (at a flow level high enough to induce phonations), the fiberscope was positioned in such a way to obstruct the oscillations of the MTs, while at the same time allowing observation of the BL. The BL maintained their movement as the sound in these induced phonations continued. Secondly, the elastic properties of the MTs were altered by applying tissue adhesive. This manipulation did not prevent vocalizations; however, it did cause changes in the acoustic properties of the vocalizations, validating the idea that the three sound sources interact to generate the vocalizations of the species.

In all six species of suboscines studied in Garcia et al. (2017), the intact syrinx produced low-frequency sounds, characterized by pulse-like signals (see Fig. 6A, MT intact, 'pulsations'). Here, we show as an example some of the results for the rufous hornero (Fig. 5D), a medium-sized ovenbird belonging to Furnariidae tracheophones (typical duet between male and female rufous hornero available from www.xeno-canto.org/748509). When the MTs of these birds were non-functional, there were three important changes: (1) the 'pulsatile' characteristic of the oscillations was lost, (2) the sound amplitude decreased, and (3) the frequency of the

oscillations increased (see Fig. 6). It is noteworthy that the presence of pulsatile sounds is not only obtained by inducing vocalizations in anesthetized animals when air is injected into the respiratory system. The spontaneous distress vocalizations of the several species studied in Garcia et al. (2017) also contain important modulations in the low-frequency range (similar to the pulsation modulations shown in Fig. 6A, MT intact), and these disappear when the tracheal oscillations are prevented by surgical manipulations.

In addition to the experimental work discussed above, the plausibility of the proposed mechanisms of vocalization in tracheophones was explored using a biomechanical model of the vocal organ. All of the experimental observations were reproduced using a dynamical systems model (Fig. 6B). The model was designed to generate synthetic sound with two different types of oscillators representing the BL and the MTs. This model assumed that the two bronchial sources were synchronized, and therefore the two sets of BL were represented using one oscillator. This is possible because, if two identical sound sources are synchronized, they generate the same airflow modulation at the same time;

therefore, the result is indistinguishable from having one or two oscillators in the biomechanical model. This was the simplest biomechanical model that allowed the data to be reproduced. It is possible that for other vocalizations, the two pairs of BL are out of synchrony or are controlled independently, but this was not observed for the distress call in the studied species in Garcia et al. (2017). In this model for sound production, if the frequencies of the two oscillators (representing BL and MTs) were sufficiently different, the lower frequency of the MTs modulated the higher frequency of the BL sound source (Fig. 6B, MT intact, 'pulsations'). Indeed, this is seen in the pronounced amplitude modulation in the spontaneous calls of the rufous hornero (see Garcia et al., 2017), and in the induced vocalizations (Fig. 6A, MT intact). By contrast, if the frequencies were similar to each other, the two sources become 'locked', and the resulting vibrations were of lower frequency than if the labial source were to vibrate on its own (Fig. 6B, MT intact, 'locking'). This is consistent with the observed shift to higher frequencies after the MTs were disabled across the different species studied, as well as with the reduced pulse-like quality of the vibrations (Garcia et al., 2017). The pulsatile oscillations arise from the non-linear interaction of two oscillators representing the MTs and the BL. If the oscillator representing the MT is removed from the system, the pulsations and locking disappear (Fig. 6B, MT disabled).

In this way, the interactions between the BL oscillations and those of the MTs constitute a morphological solution to the generation of intense sounds, with modulated amplitude, as well as pulse-like properties. In oscine species, many of these properties are achieved by neuromuscular control of their sound sources (Beckers, 2011; Goller and Riede, 2013; Suthers and Zollinger, 2008). The suboscines do not have the same level of muscular control, but work on tracheophones shows that they are capable of generating similar acoustic properties using specific morphological features.

Conclusions and perspective

In recent years, much progress has been made in understanding the physical mechanisms and dynamics involved in the generation of song by oscine birds. In parallel, a lot of effort has been applied to the study of the neuronal substrate that controls the vocal apparatus. Both the neuronal architecture underlying the control of the vocal apparatus and the anatomy of the syrinx are highly conserved across oscine species.

The picture is quite different when it comes to the sister clade of oscines, the suboscine birds, which are thought to be, in their great majority, non-vocal learners. These birds display a remarkable anatomical diversity, and the mechanisms used in the generation of their songs are extremely varied. To illustrate this phenomenon, here, we have concentrated on the study of a particular acoustic property: the existence of slow modulations (100–200 Hz) coexisting with fast oscillations (kHz range). This is an acoustic feature that is clearly identifiable and is present in several suboscine species. We have discussed the dynamical origin of this feature in three different cases: the screeching sound of the white-tipped plantcutter, the carefully controlled detuning between the two sound sources in the kiskadee's first syllable, and the sounds generated by the complex anatomical structure of the tracheophones. In all three examples, some set of properties in the sound signals suggested a starting point for building a non-linear dynamical model, which then generated quantitative predictions that could be used to test the assumed hypotheses regarding the mechanisms underlying the vocalizations. Taken together, these three stories – which integrate biological experiments and biomechanical modeling using

non-linear dynamical systems – illustrate how a morphological or dynamical adaptation can produce a complex acoustic property without the need for complex neuromuscular control.

An emerging hypothesis posits that it is through their different adaptations that suboscines are able to achieve a diverse range of acoustic properties in their repertoires. It is possible that it is precisely because of the lack of vocal learning that the suboscines must rely on morphological adaptations to produce acoustic characteristics that oscines can achieve through more complex neuromuscular control.

References
Amador, A. and Margoliash, D. (2013). A mechanism for frequency modulation in songbirds shared with humans. *J. Neurosci.* **33**, 11136-11144. doi:10.1523/JNEUROSCI.5906-12.2013

Amador, A. and Mindlin, G. B. (2008). Beyond harmonic sounds in a simple model for birdsong production. *Chaos* **18**, 043123. doi:10.1063/1.3041023

Amador, A., Goller, F. and Mindlin, G. B. (2008). Frequency modulation during song in a suboscine does not require vocal muscles. *J. Neurophysiol.* **99**, 2383-2389. doi:10.1152/jn.01002.2007

Amador, A., Boari, S. and Mindlin, G. B. (2017). From perception to action in songbird production: dynamics of a whole loop. *Curr. Opin. Syst. Biol.* **3**, 30-35. doi:10.1016/j.coisb.2017.03.004

Ames, P. L. (1971). The morphology of the syrinx in passerine birds. *Bulletin of the Peabody Museum of Natural History* 37. https://elischolar.library.yale.edu/peabody_museum_natural_history_bulletin/37

Beckers, G. J. L. (2011). Bird speech perception and vocal production: a comparison with humans. *Hum. Biol.* **83**, 191-212. doi:10.3378/027.083.0204

Bhama, P. K., Hillel, A. D., Merati, A. L. and Perkel, D. J. (2011). A novel model for examining recovery of phonation after vocal nerve damage. *J. Voice* **25**, 275-282. doi:10.1016/j.jvoice.2009.11.005

Brittain-Powell, E. F., Dooling, R. J., Larsen, O. N. and Heaton, J. T. (1997). Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*). *J. Acoust. Soc. Am.* **101**, 578-589. doi:10.1121/1.418121

Cardoso, G. C. (2012). Paradoxical calls: the opposite signaling role of sound frequency across bird species. *Behav. Ecol.* **23**, 237-241. doi:10.1093/beheco/arr200

Casey, G. and Gaunt, A. S. (1985). Theoretical models of the avian syrinx. *J. Theor. Biol.* **116**, 45-64. doi:10.1016/S0022-5193(85)80130-2

Döppler, J. F., Amador, A., Goller, F. and Mindlin, G. B. (2020). Dynamics behind rough sounds in the song of the *Pitangus sulphuratus*. *Phys. Rev. E* **102**, 062415. doi:10.1103/PhysRevE.102.062415

Düring, D. N., Ziegler, A., Thompson, C. K., Ziegler, A., Faber, C., Müller, J., Scharff, C. and Elemans, C. P. (2013). The songbird syrinx morphome: a three-dimensional, high-resolution, interactive morphological map of the zebra finch vocal organ. *BMC Biol.* **11**, 1-27. doi:10.1186/1741-7007-11-1

Düring, D. N., Knörlein, B. J. and Elemans, C. P. (2017). In situ vocal fold properties and pitch prediction by dynamic actuation of the songbird syrinx. *Sci. Rep.* **7**, 1-14. doi:10.1038/s41598-017-11258-1

Elemans, C. P., Mead, A. F., Rome, L. C. and Goller, F. (2008). Superfast vocal muscles control song production in songbirds. *PLoS ONE* **3**, e2581. doi:10.1371/journal.pone.0002581

Fainstein, F., Geli, S. M., Amador, A., Goller, F. and Mindlin, G. B. (2021). Birds breathe at an aerodynamic resonance. *Chaos* **31**, 123132.doi:10.1063/5.0069696

Fletcher, N. H. (1988). Bird song - a quantitative acoustic model. *J. Theor. Biol.* **135**, 455-481. doi:10.1016/S0022-5193(88)80270-4

Fletcher, N. H. (1989). Acoustics of bird song - some unresolved problems. *Commun. Theor. Biol.* **4**, 237-251.

Gahr, M., Güttinger, H. R. and Kroodsma, D. E. (1993). Estrogen receptors in the avian brain: survey reveals general distribution and forebrain areas unique to songbirds. *J. Comp. Neurol.* **327**, 112-122. doi:10.1002/cne.903270109

Garcia, S. M., Kopuchian, C., Mindlin, G. B., Fuxjager, M. J., Tubaro, P. L. and Goller, F. (2017). Evolution of vocal diversity through morphological adaptation without vocal learning or complex neural control. *Curr. Biol.* **27**, 2677-2683. doi:10.1016/j.cub.2017.07.059

Gardner, T., Cecchi, G., Magnasco, M., Laje, R. and Mindlin, G. B. (2001). Simple motor gestures for birdsongs. *Phys. Rev. Lett.* **87**, 208101. doi:10.1103/PhysRevLett.87.208101

Gaunt, A. S. and Gaunt, S. (1985). Syringeal structure and avian phonation. In *Current Ornithology*, Vol. 2 (ed. R. F. Johnston), pp. 213-245. New York: Plenum Press.

Goller, F. and Cooper, B. G. (2004). Peripheral motor dynamics of song production in the zebra finch. *Ann. N. Y. Acad. Sci.* **1016**, 130-152. doi:10.1196/annals.1298.009

Goller, F. and Larsen, O. N. (1997). A new mechanism of sound generation in songbirds. *Proc. Natl. Acad. Sci. USA* **94**, 14787-14791. doi:10.1073/pnas.94.26.14787

Goller, F. and Larsen, O. N. (1999). Role of syringeal vibrations in bird vocalizations. *Proc. R. Soc. Lond. B* **266**, 1609-1615. doi:10.1098/rspb.1999.0822

Goller, F. and Larsen, O. N. (2002). New perspectives on mechanisms of sound generation in songbirds. *J. Comp. Physiol. A* **188**, 841-850. doi:10.1007/s00359-002-0350-6

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

Goller, F. and Suthers, R. A. (1996a). Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.* **75**, 867-876. doi:10.1152/jn.1996.75.2.867

Goller, F. and Suthers, R. A. (1996b). Role of syringeal muscles in controlling the phonology of bird song. *J. Neurophysiol.* **76**, 287-300. doi:10.1152/jn.1996.76.1.287

Goller, F., Love, J. and Mindlin, G. (2021). Different frequency control mechanisms and the exploitation of frequency space in passerines. *Ecol. Evol.* **11**, 6569-6578. doi:10.1002/ece3.7510

Hartley, R. S. and Suthers, R. A. (1989). Airflow and pressure during canary song: direct evidence for mini-breaths. *J. Comp. Physiol. A* **165**, 15-26. doi:10.1007/BF00613795

Jensen, K. K., Cooper, B. G., Larsen, O. N. and Goller, F. (2007). Songbirds use pulse tone register in two voices to generate low-frequency sound. *Proc. R. Soc. B* **274**, 2703-2710. doi:10.1098/rspb.2007.0781

King, A. S. (1989). Functional anatomy of the syrinx. In *Form and Function in Birds* (ed. A.S. King and J. McLelland), pp. 105-192. Academic Press.

Kroodsma, D. E. (1984). Songs of the Alder Flycatcher (Empidonax alnorum) and Willow Flycatcher (Empidonax traillii) are innate. *The Auk* **101**, 13-24. doi:10.1093/auk/101.1.13

Kroodsma, D. E. (1985). Development and use of two song forms by the Eastern Phoebe. *Wilson Bull.* **97**, 21-29.

Kroodsma, D. E. and Konishi, M. (1991). A suboscine bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Anim. Behav.* **42**, 477-487. doi:10.1016/S0003-3472(05)80047-8

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., Trainer, J. M. and Powell, G. V. (2013). Behavioral evidence for song learning in the suboscine bellbirds (*Procnias* spp.; Cotingidae). *Wilson J. Ornithol.* **125**, 1-14. doi:10.1676/12-033.1

Larsen, O. and Goller, F. (2002). Direct observation of syringeal muscle function in songbirds and a parrot. *J. Experim. Biol.* **205**, 25-35. doi:10.1242/jeb.205.1.25

Linhart, P. and Fuchs, R. (2015). Song pitch indicates body size and correlates with males' response to playback in a songbird. *Anim. Behav.* **103**, 91-98. doi:10.1016/j.anbehav.2015.01.038

Liu, W. C., Wada, K., Jarvis, E. D. and Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nat. Commun.* **4**, 1-12. doi:10.1038/ncomms3082

Liu, J. P., Ma, L. K., Zhang, Z. Q., Gu, D. H., Wang, J. J., Li, J. J., Goa, L. J. and Hou, J. H. (2017). Maximum frequency of songs reflects body size among male dusky warblers *Phylloscopus fuscatus* (Passeriformes: Phylloscopidae). *Eur. Zool. J.* **84**, 186-192. doi:10.1080/24750263.2017.1301578

Meiri, S. and Dayan, T. (2003). On the validity of Bergmann's rule. *J. Biogeogr.* **30**, 331-351. doi:10.1046/j.1365-2699.2003.00837.x

Mindlin, G. B. (2017). Nonlinear dynamics in the study of birdsong. *Chaos* **27**, 092101. doi:10.1063/1.4986932

Mindlin, G. B. and Laje, R. (2006). *The Physics of Birdsong*. Springer Science & Business Media.

Mindlin, G. B., Gardner, T. J., Goller, F. and Suthers, R. A. (2003). Experimental support for a model of birdsong production. *Phys. Rev. E* **68**, 041908. doi:10.1103/PhysRevE.68.041908

Mooney, R. (2022). Birdsong. *Curr. Biol.* **32**, R1090-R1094. doi:10.1016/j.cub.2022.07.006

Nowicki, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* **325**, 53-55. doi:10.1038/325053a0

Perl, Y. S., Arneodo, E. M., Amador, A., Goller, F. and Mindlin, G. B. (2011). Reconstruction of physiological instructions from Zebra finch song. *Phys. Rev. E* **84**, 051909. doi:10.1103/PhysRevE.84.051909

Proctor, N. S., Peterson, R. J. and Lynch, P. J. (1993). *Manual of Ornithology: Avian Structure and Function*. New Haven, CT: Yale University Press.

Riede, T. and Suthers, R. A. (2009). Vocal tract motor patterns and resonance during constant frequency song: the white-throated sparrow. *J. Comp. Physiol. A* **195**, 183-192. doi:10.1007/s00359-008-0397-0

Riede, T., Suthers, R. A., Fletcher, N. H. and Blevins, W. E. (2006). Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Natl Acad. Sci. USA* **103**, 5543-5548. doi:10.1073/pnas.0601262103

Riede, T. and Goller, F. (2010). Peripheral mechanisms for vocal production in birds–differences and similarities to human speech and singing. *Brain and Language* **115**, 69-80.

Riede, T., Schilling, N. and Goller, F. (2013). The acoustic effect of vocal tract adjustments in zebra finches. *J. Comp. Physiol. A* **199**, 57-69. doi:10.1007/s00359-012-0768-4

Rodríguez-Cajarville, M. J., Calderón, L., Tubaro, P. L. and Cabanne, G. S. (2019). Body size and genetic variation in the white-tipped plantcutter (*Phytotoma rutila*: Cotingidae) suggest ecological divergence across the Chaco–Andes dry forest belt. *J. Ornithology* **160**, 947-961. doi:10.1007/s10336-019-01694-3

Rüppell, W. (1933). Physiologie und Akustik der Vogelstimme. *J. Ornithol.* **81**, 433-542. doi:10.1007/BF01905461

Sakata, J. T., Woolley, S. C., Fay, R. R. and Popper, A. N. (eds.). (2020). *The Neuroethology of Birdsong*, Vol. 71. Springer Nature.

Saranathan, V., Hamilton, D., Powell, G. V., Kroodsma, D. E. and Prum, R. O. (2007). Genetic evidence supports song learning in the three–wattled bellbird *Procnias tricarunculata* (Cotingidae). *Mol. Ecol.* **16**, 3689-3702. doi:10.1111/j.1365-294X.2007.03415.x

Stein, R. C. (1968). Modulation in bird sounds. *The Auk* **85**, 229-243. doi:10.2307/4083583

Suthers, R. A. and Margoliash, D. (2002). Motor control of birdsong. *Curr. Opin. Neurobiol.* **12**, 684-690. doi:10.1016/S0959-4388(02)00386-0

Suthers, R. A. and Zollinger, S. A. (2004). Producing song: the vocal apparatus. *Ann. N. Y. Acad. Sci.* **1016**, 109-129. doi:10.1196/annals.1298.041

Suthers, R. A. and Zollinger, S. A. (2008). Mechanisms of song production in songbirds. In *Neuroscience of Birdsong* (ed. H.P. Zeigler and P. Marler), pp. 78-98. Cambridge University Press.

Suthers, R., Goller, F. and Pytte, C. (1999). The neuromuscular control of birdsong. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **354**, 927-939. doi:10.1098/rstb.1999.0444

Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *Acoust. Soc. Am* **85**, 901-906. doi:10.1121/1.397562

Titze, I. R. and Martin, D. W. (1994). *Principles of Voice Production*. Prentice Hall.

Tobias, J. A., Brawn, J. D., Brumfield, R., Derryberry, E. P., Kirschel, A. N. and Seddon, N. (2012). The importance of suboscine birds as study systems in ecology and evolution. *Ornitol. Neotrop.* **23**, 261-274.

Touchton, J. M., Seddon, N. and Tobias, J. A. (2014). Captive rearing experiments confirm song development without learning in a tracheophone suboscine bird. *PLoS ONE* **9**, e95746. doi:10.1371/journal.pone.0095746

Uribarri, G., Rodríguez-Cajarville, M. J., Tubaro, P. L., Goller, F. and Mindlin, G. B. (2020). Unusual avian vocal mechanism facilitates encoding of body size. *Phys. Rev. Lett.* **124**, 098101. doi:10.1103/PhysRevLett.124.098101

Zeigler, H. and Marler, P. E. (2008). *Neuroscience of Birdsong*. Cambridge University Press.