



# Technological advances for getting insight into the learning capacities of birds in the vocal domain

Sébastien Derégnaucourt, Alice Araguas, Bahia Guellaï

## ► To cite this version:

Sébastien Derégnaucourt, Alice Araguas, Bahia Guellaï. Technological advances for getting insight into the learning capacities of birds in the vocal domain. *Interaction Studies*, 2023, 24 (2), pp.289-310. 10.1075/is.22043.der . hal-04410046

**HAL Id: hal-04410046**

**<https://hal.science/hal-04410046>**

Submitted on 22 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Technological advances for gaining insight into the learning capacities of birds in the vocal and auditory domains**

Abstract: Birds produce different types of sounds in different contexts such as begging for food in youngsters, alerting to a danger, defending a territory or attracting a sexual partner. About half of the bird species are able to transform their vocalizations through imitation, improvisation or invention of sounds. Here we review the different experimental procedures that have been used to study the learning capacities of birds in the vocal domain and in the auditory domain, with a particular emphasis on recent technological developments. Nowadays, it is possible to record individual vocalizations of birds living in social groups or to record continuously the vocal ontogeny of birdsong. In conditioning experiments, new paradigms have successfully replaced food rewards with a socio-sexual reward. It is possible to engage in vocal interactions with a bird using dedicated computer systems. In both the laboratory and more recently in the field, different techniques have been used to train young oscine songbirds to learn from acoustic models. The use of virtual social environments and robots as social agents are also promising avenues. All together, these new techniques will permit researchers to explore more deeply the umwelt of bird species.

## 1. Introduction

Birds have fascinated and inspired humans for centuries. Of particular interest have been oscine songbirds (belonging to the suborder Passeri of the perching birds, Passeriformes) that have historically been raised by humans, primarily for aesthetics reasons linked to their plumage and their natural song. But humans have also tried to educate them to produce human words and musical melodies. For example, wind instruments such as “flageolets” (woodwind instruments and members of the fipple flute family) and “serinettes” (‘serin’ is French for canary - a music box similar to a little barrel organ) were used to play tunes to birds to stimulate or train them to sing. These practices were very popular in Europe in the aristocratic milieu in the 18<sup>th</sup> and the 19<sup>th</sup> centuries. They took place in the context of conquering, taming and domesticating nature, and to some extent to "civilize" the animals but also to fuel the debate regarding the role of education on human nature (Clouzot & Kreutzer, 2019). Breeders of the past, especially from the 15<sup>th</sup> to the 18<sup>th</sup> centuries, discovered that innate predispositions to learn can be manipulated through an acute control of social factors during particular moments of the bird’s life (Canary *Serinus canaria*: Hervieux de Chanteloup, 1709; Nightingale *Luscinia megarhynchos*: de Nobleville, 1751, see also the contribution of Baron Von Perneau: Stresemann, 1947; see Birkhead & Van Balen, 2008 for a review). From these reports, we learned that: (1) birds need to be isolated from conspecifics during the vocal practice of these musical melodies but not necessarily during exposure and (2) there is no need to overexpose the birds to the models to imitate: few lessons a day are good enough to trigger good imitations. These results were later confirmed using experiments in the controlled conditions of the laboratory (Thorpe, 1958; Tchernichovski et al., 1999).

Particularly during the 18<sup>th</sup> century, several species were captured in the wild and bred in captivity. Through several different and successive selection processes, wild caged birds of some species then became domesticated ‘cage birds’. Breeders developed techniques to raise

44 and to teach these birds some sounds far removed from their usual vocalizations, taking  
45 advantage of vocal production learning. Such learning is also an important component of  
46 language, particularly with respect to the ability to imitate environmental sounds especially  
47 conspecific ones, an ability exhibited by few taxonomic groups of birds (oscine songbirds,  
48 hummingbirds and parrots) and of mammals (whales, bats, seals, elephants and humans) but  
49 not in monkeys and apes (Petkov & Jarvis, 2012). Nevertheless, the traditional dichotomy  
50 between vocal learners and non-vocal learners seems obsolete since several cases of social  
51 influences on some aspects of vocal development have been reported in some species of birds  
52 or primates traditionally classified as non-vocal learners (see Watson et al., 2015; Desmedt et  
53 al., 2020; Kroodsma et al., 2013 for example). A continuum with respect to graded vocal  
54 control seems a more appropriate description nowadays (Petkov & Jarvis, 2012; Tyack, 2020;  
55 ten Cate, 2021).

56 Since the invention of the sound spectrograph, birdsong became a useful model system to  
57 investigate how experience and genetically determined circuitry interact to shape a learned  
58 behavior (Mets & Brainard, 2019) and thousands of birds from tens of species have been  
59 raised in controlled conditions of the laboratory to determine how their social environment,  
60 and especially their auditory experience, shape their vocalizations. Birdsong acquisition  
61 exhibits many similarities with speech and language development (Bolhuis & Everaert, 2013;  
62 Marler, 1970; Doupe & Kuhl, 1999; Hyland Bruno et al., 2021; Lipkind et al., 2020).  
63 Birdsong learning is based on an auditory-vocal integration: song models heard during a  
64 sensory phase are compared to the bird's own vocalizations during the sensorimotor phase  
65 when intense practice occurs (Konishi, 1965). As in human infants, vocal developmental  
66 learning goes through different stages in juvenile oscine songbirds: following the first weeks  
67 of avian babbling, plastic versions of the future song notes and syllables (a terminology  
68 inherited from studies on language and music and attested in Medieval and Modern

encyclopedias: Fritz, 2016) appear before a process of stabilization called “crystallization” (Thorpe, 1958). From these studies, it was discovered that some species can learn to imitate a unique song during a short sensitive period of early life (“age-limited learners”) whereas others can learn to modify their vocal repertoire during their whole life (“open-ended learners”). It was also discovered that some species do not need to be exposed for weeks to models and can modify their vocalizations through imitation almost immediately (Balsby & Bradbury, 2009). For a long time, it was believed that oscine songbirds could not learn beyond the sensitive period. But successive studies have shown that song learning could still occur in young birds raised in social isolation or exposed to inappropriate models such as tape-recorded songs during the so-called sensitive period before being exposed to an appropriate model. For example, Baptista and Petrinovitch (1984) demonstrated that young white-crowned sparrows (*Zonotrichia leucophrys*) could learn songs after the end of the sensitive period if they were exposed to live tutors but not if they were exposed to taped songs. Most of these studies, done in the laboratory and/or with data collected in the field, provided insight into the processes of song learning, and more generally of vocal learning, in a more relevant context (Carouso-Peck et al., 2021).

Significant knowledge has been gained on the sensorimotor aspects of vocal production and learning these last 25 years with the development of computers and software. Here we present these different improvements for different purposes linked to learning in the vocal domain such as vocal recording and analysis, song tutoring, conditioning experiments and controlled vocal interactions. Even if most research deals with birdsong learning in oscine songbirds, our review is not restricted to vocal production learning capacities and could embrace other types of learning in the vocal domain such as contextual learning (Janik & Slater, 1997) that could be applied to a wide range of bird species including species traditionally classified as non-

vocal-learners. We will conclude by suggesting potential lines of research on the use of computers in the field of avian communication that seem feasible in the future.

## **2. Continuous vocal recording and analysis: from a single individual to group members**

Before the age of computer technology, sound recordings were saved on tapes which drastically limited quantitative accumulation and analysis of data. It is also worth noting that the Kay Sona-Graph (which was also prohibitively expensive for many researchers) could not display a sonogram representing more than 1.5sec of sound, which for a long time restricted birdsong research to short vocalizations and thus often to songs of only a few species. A major step was reached when vocal recording and analysis were completely digitalized (Tchernichovski et al., 2004; Gardner et al., 2005). Such continuous recording and analysis permitted, for example, the discovery of the role of sleep on developmental song learning (Derégnaucourt et al., 2005; Shank & Margoliash, 2009), the role of melatonin on vocal production and learning in both vocal and non-vocal learners (Derégnaucourt et al., 2012) and the ability to track cultural evolution of birdsong in the laboratory (Fehér et al., 2009). Such methods also allowed the exploration of vocal ontogeny in both vocal and non-vocal learner species (Gardner et al., 2005; Derégnaucourt et al., 2009) and the discovery of some vocal changes that would have been unnoticed without continuous recording. Such systems were developed to record a bird kept singly in a sound-proof chamber, and to analyze the sound recorded online, which also permitted researchers to detect vocalizations and to discard cage noises. The invention of spectrographic analysis also allowed researchers to focus on the accuracy of imitation during vocal learning. Initially, imitation was determined by researchers eyeballing spectrograms with some issues on inter-observer reliabilities depending on their

117 experience (Jones et al., 2001). Some progress has been made by automatizing digital  
118 similarity measurements between different sounds, e.g., between a tutor's song and its  
119 imitation by a pupil (Tchernichovski et al., 2000; Lachlan et al., 2010). Such an approach  
120 depends on specific song features selected by experimenters. In a more recent attempt, Mets  
121 and Brainard (2018) used a statistical model to broadly capture the structure of each song, and  
122 then estimated the divergence between the model and its imitation. The authors emphasized  
123 that this measure could be used to quantify changes in other behaviors such as human speech  
124 (Mets & Brainard, 2018). Continuous recordings at the group level is now permitted thanks to  
125 the development lightweight wireless microphone backpacks (Anisimov et al., 2014; Gill et  
126 al., 2016). These methods allow birds to behave freely in a social group and enable studies of  
127 vocal interactions. For example, Gill et al. (2015) observed that call occurrences and  
128 repertoire changed over the breeding cycle in a group of zebra finches (*Taeniopygia guttata*).  
129 In another study, simultaneous recordings of individual vocalizations and multiunit vocal  
130 premotor activity in breeding pairs of white-browed sparrow-weaver (*P. mahali*) ranging  
131 freely in their natural habitat permitted researchers to document how the precise inter-  
132 individual coordination of motor activity during duet singing is neuronally controlled  
133 (Hoffman et al., 2019). The authors observed that rhythmic social interactions that require  
134 temporal coordination are not just associated with a coherent oscillation of activity in large  
135 areas of both interacting partners' brains but with an inter-individually synchronized  
136 activation of small groups of neurons within the same brain nucleus in both partners. Again,  
137 such study would have not been possible without continuous recording. In the field,  
138 microphone arrays are also increasingly used to remotely monitor acoustically active animals  
139 including birds (Blumstein et al., 2011) but to our knowledge, they had never been used to  
140 study directly the learning capacities of birds in the vocal domain even if changes in the

soundscape tracked by studies in the field of eco-acoustics (Stowell & Sueur, 2020) might capture these aspects.

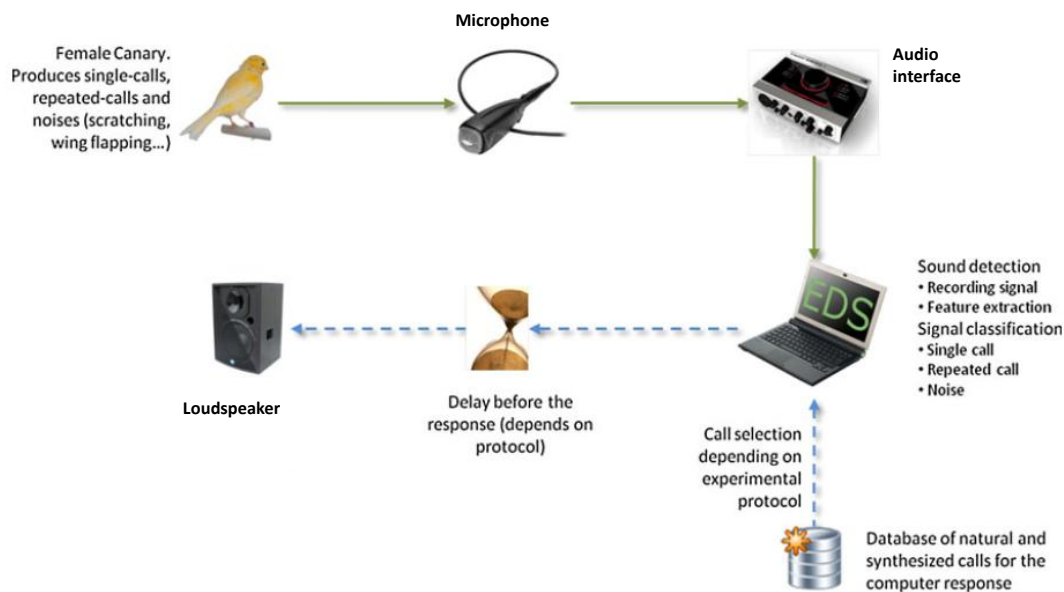
### 3. Controlled vocal interactions

As mentioned above, continuous recordings permitted researchers to document vocal interactions between two individuals and at the group level. This necessitates automatic detection of sounds and adapted response with the possibility to accurately control the delay between sound detection and response. Turn-taking is an important component of language and has been demonstrated in early infancy in humans and also in other species (Levinson, 2016). Such a process has also been the subject of much research in birds, especially in the context of territory defense in oscine songbirds and has been simulated using interactive playback (Slater et al., 2003). For example, song type matching is found in some species such as the song sparrow (*Melospiza melodia*; Akçay et al., 2013; Beecher et al., 2000). It has been observed that birds approached a type-matching stimulus much more aggressively than a non-matching stimulus (Vehrencamp, 2001). Using a call detector implemented into an automated playback paradigm, Okanoya and Kimura (1993) observed that Bengalese finches (*Lonchura striata*) responded only to the Bengalese finch calls, but not the zebra finch calls, and female calls were more efficient in inducing responses than male calls. Using a dedicated software that permits continuous recording and real-time analysis, and a plastic model of a zebra finch equipped with a speaker, Benichov and colleagues (2016) investigated the neurobiological basis of turn-taking in zebra finches. Using an isochronous call pattern (e.g. one call per second), they observed that finches adjusted their call timing rapidly. Then, they changed the timing to enhance jamming, namely to increase the probability to overlap the birds' calls with calls sent by the system. All zebra finches adjusted their call pattern to avoid jamming, either by calling earlier, later or both. When the experimenters used a pattern of alternating single



and paired calls, finches adjusted the timing of their calls differently for the single compared to the paired calls. Interestingly, females performed better than males at these tasks. The authors observed that fully functional forebrain regions, part of the so-called “song control system” – a set of several brain nuclei involved in song production, perception and learning - are necessary for this precise timing in both males and females (Benichov et al., 2016).

Such a close-loop computer system was also used successfully in domestic canaries to investigate the functions and context of production of two types of calls: “single calls” and “repeated calls” (Figure 1, Lerch et al., 2011).



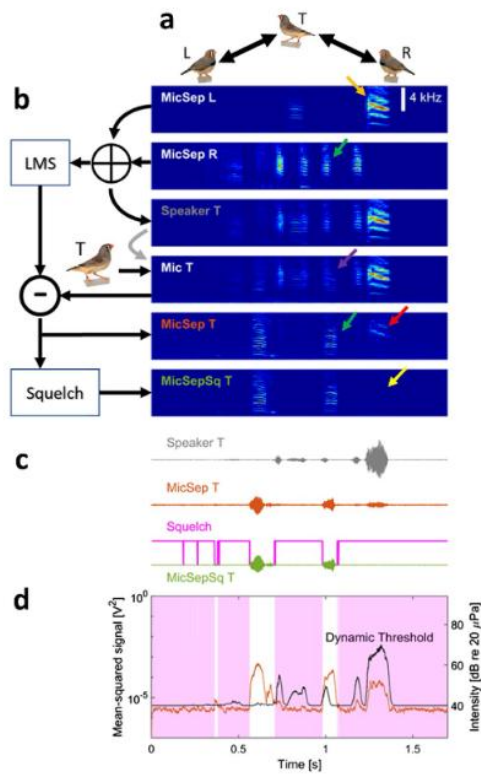
*Figure 1. Scheme of the closed-loop interaction between a canary and the computer system. Solid green arrows indicate the flow of acoustic signals produced by the bird. Dashed blue arrows indicate the signal produced by the system as a response to the bird’s vocalization. Modified from Lerch et al., Animal Cognition, 2011.*

The experimental procedures involved different stimuli (either natural or artificially synthesized) and different scenarios (e.g. positive interaction: the computer responded with a

single call when the bird produced a single call and with a repeated call when the bird produced a repeated call, and negative interaction: the computer responded with a single call when the bird produced a repeated call, and with a repeated call when the bird produced a single call). Only females were tested: they produced significantly more single than repeated calls (either in response to natural and synthesized calls), and their rate of single calls was associated with the context (more often higher in the case of positive interaction), whereas repeated calls were context independent. This experiment confirmed a long-term assumption that “single calls” act as contact calls in canaries, whereas “repeated calls” are mostly used in a context of stress (Mulligan & Olsen, 1969).

More recently, Rychen and colleagues (2021) developed a system for real-time control of vocal interactions among separately housed animals. The system is implemented on a Field-Programmable Gate Array (FPGA) and it can control communication networks by restricting vocal exchanges among chosen individuals up to four animals (Figure 2). For example, an individual could interact with other individuals that could not hear each other. Such system also gives the possibility to mimic eavesdropping situations when one individual can listen to vocal exchanges between other individuals without being noticed by the others.

198



200

201 *Figure 2. System controlling vocal communication networks. (a) The female T on top is*  
 202 *bidirectionally connected (arrows) with two males L and R that are not supposed to hear each*  
 203 *other (hierarchical network). (b–d) Example in which T and L call simultaneously (green*  
 204 *arrows). (b) Even though Mic T records a superposition of both calls (magenta arrow), L's*  
 205 *call gets cleanly filtered out (MicSep T) by the LMS (Least Mean Squares) filter. In contrast,*  
 206 *a louder call by L (orange arrow) leaves a significant echo on MicSep T (red arrow). The*  
 207 *squelch (an algorithm that prevents transitive sound propagation in linked chambers for*  
 208 *example here bird L cannot hear bird R) removes the residual of that call on MicSepSq T*  
 209 *(yellow arrow), which constitutes a clean signal, just as if L had been silent. Spectrograms*  
 210 *have been normalized to individual color scales. (c) Shown are the Speaker T (gray), MicSep*  
 211 *T (red), and MicSepSq T (green) sound waveforms. (d) The light magenta regions indicate*  
 212 *times when the squelch is active and signal transmission is blocked. Modified from Rychen et*  
 213 *al. Scientific Reports 2021.*

## 4. Enhancing song tutoring

In the study of song ontogeny, there have been methodological advances and improvements since the early days of simple playback of taped songs. This includes control of song playback by subject, playback augmented by visual stimuli, playback of vocal interactions and more recently the development of robots as song tutors.

### 4.1. Song playback

Different artificial methods have been developed in the laboratory to train young birds to learn songs. Since the pioneering works of Peter Marler using tape playback to train young sparrows to learn a song (Marler, 1970; Marler & Peters, 1977), tens of species have been exposed to their specific songs or other models during the sensitive period with mixed results: if some species do learn pretty well using such a passive exposure, some do not (e.g., Zebra Finch; Derégnaucourt et al., 2013). In this species, when song presentation was made contingent upon key pressing, young males copied to a greater extent than controls receiving the same stimulation in a passive way (Adret, 1993a) and the effectiveness of the contingency for initiating song imitation was greater for key pressing rather than hopping on a reinforcing perch (Adret, 1993b, but see James & Sakata, 2017). As previously mentioned, few daily exposure of song models are sufficient to trigger good learning (Tchernichovski et al., 1999; Deshpande et al., 2014). Using this paradigm and other experimental and computational approaches to manipulate and analyze the birds' vocal development, several studies have revealed different rules during birdsong developmental learning that exhibit analogies with language acquisition in humans (Lipkind et al., 2013, 2017, 2020; James & Sakata, 2017; Mets & Brainard, 2018).

237 For example, zebra finches could be exposed to non-conspecific songs as models such as  
238 heterospecific sounds (Gehrold et al., 2013) and canaries to synthetic sounds (Gardner et al.,  
239 2005). It was observed that this ability to learn non-conspecific songs is restricted to the  
240 production of avian babbling and does not persist into adulthood (Gardner et al., 2005). In  
241 another experiment, young male zebra finches were exposed to synthesized songs consisting  
242 of five species-typical syllables arranged in every possible five-syllable sequence ( $n = 120$   
243 possible sequences) (James & Sakata, 2017). The authors observed that all birds converged  
244 towards similar song sequences showing that there exist learning biases underlying song  
245 organization in this species. Dina Lipkind and colleagues have used extensively a paradigm  
246 that consists of switching song models during the sensitive period for song learning to study  
247 how young birds cope with vocal changes both in the temporal and the spectral domains  
248 during ontogeny (Lipkind et al., 2013, 2017, 2020). For example, young finches were exposed  
249 to a song model composed of three syllables (A-B-C) and when they started to produce the  
250 right sequence, they were exposed to another song model with the same syllables in a  
251 different order (A-C-B). Some birds succeeded to modified their song but after a huge  
252 practice involving their learning of new syllable transitions such as A to C and C to B  
253 (Lipkind et al., 2013). Such vocal changes imposed experimentally in birds exhibit analogies  
254 with language acquisition in humans (Lipkind et al., 2013). The study of interactions between  
255 genetic predispositions and experience is a huge topic in ethology and in birdsong research in  
256 particular. In a series of experiments with Bengalese finches, Mets and Brainard used both  
257 cross-fostering and computerized instruction with synthetic songs to demonstrate that  
258 matching the tutor song to individual predispositions can improve learning across genetic  
259 backgrounds (2019). All together, these examples show how being able to manipulate sound  
260 playback to birds during song ontogeny could shed lights on general mechanisms of learning.

In the field, experimental evidence of birdsong learning was demonstrated recently in wild savannah sparrows (*Passerculus sandwichensis*, Mennil et al., 2018). The authors used an array of loudspeakers to broadcast conspecific songs recorded in North America to birds living on an island. They also observed that these new songs spread in the subsequent generations.

#### 4.2. Influence of visual stimuli

Several experiments have shown that birdsong acquisition is not an auditory phenomenon per se and that multimodal cues are also involved in the learning process, including visual stimuli. Their influence on birdsong learning was demonstrated for example in nightingales: song acquisition was enhanced when playback was associated with strobe light produced synchronously with the temporal and amplitude patterns of the song models (Hultsch et al., 1999). Several experiments used visual stimuli combined with audio playback including taxidermic mounts or plastic dummies (see Derégnaucourt, 2010 for a review). Adret (1997) trained finches to peck a key to trigger a video movie and observed poor learning. This result could be explained by the flickering frequency of Cathode Ray Tube (CRT) used for this experiment, that exhibited a time resolution too low for birds. TFT (Thin Film Transistors) have been used successfully to induce behavioral reactions in birds (e.g. courtship songs in zebra finches and Bengalese finches: Ikebuchi & Okanoya, 1999) and for the purpose of song tutoring, even if, to our knowledge, there have been only a couple of studies dealing with these aspects that have been published and presented below. In an experiment in zebra finches, Deshpande et al. (2014) used different paradigms during a sole day of training (20 key pecks allowed): following a key-peck as described above, the audio-visual components of

song models were played either simultaneously (Sim) or in a staggered fashion where birds received either audio before video (AV) or video before audio (VA). When trained at 45 dph (day post-hatch), AV and Sim groups showed significantly higher similarity to the song model than did an untrained group of birds kept in social isolation during the sensitive period. However, there were no significant differences between the different groups of trained birds. By contrast, when trained at 35 dph, differences in the efficacy of these audio-visual training stimuli became apparent. Surprisingly, a single training session at this age was effective only in the AV condition. This result is reminiscent of the Colavita visual dominance effect (Colavita, 1974): when one has to respond simultaneously to an auditory signal and to a visual signal, one tends to respond only to the visual signal which illustrates visual dominance in attentional processes. But as mentioned by the authors, further work is required to explore these aspects.

In a recent experiment, dyads consisting of a young male and an unrelated young female were exposed passively to audio-video recordings of a singing adult male. If birds approached the setup more often and spent more time close to it than juveniles exposed to audio playback only or audio playback combined with pixelated and time-reversed videos, they did not exhibit better imitation of the tutor song (Varkevisser et al., 2021).

#### 4.3. Playback of vocal interactions

Several studies suggest that listening to vocal interactions between adults or between an adult and a young could enhance vocal learning. Such paradigm, inspired by the seminal work of Dietmar Todt (Model/rival approach; Todt, 1975) has been used with success by Irene Pepperberg to train Alex and other grey parrots (*Psittacus erithacus*) to learn human words and to associate them with different types of objects (social modeling theory, Pepperberg,

1985). In this context, parrots could observe interactions between two humans, one person being the ‘trainer’ and one being the ‘model’ as well as the ‘rival.’(Pepperberg, 1985). Young European starlings (*Sturnus vulgaris*) raised in social isolation learned better when they could hear live interactions than while exposed to playback of taped songs (Poirier et al., 2004). In another experiment with song sparrows (*Melospiza melodia*), hand-raised birds were exposed sequentially to two pairs of adult males in Summer (Burt et al., 2007). During the first three months of the next year, a computer program was used to simulate a live, interactive tutor. Individuals were exposed to computer presentations of songs from two of the original live tutors: one of the two “virtual” tutors sang interactively with the young bird while the other sang asynchronously. The song classifier program used a spectrograph cross-correlation algorithm to compare the detected subject’s song against examples from the tutor’s repertoire. The classifier returned a score indicating the similarity of the detected subject song to each of the tutor examples. If the score of the most similar tutor example type was above a fixed threshold, then the virtual tutor concluded that the subject had sung that type, otherwise it considered the subject had sung an unknown type. The virtual tutor then acted upon this information. The classifier parameters and similarity threshold were calibrated a priori using a test set of different song sparrow songs and the classifier output was periodically assessed during tutoring to verify that it was working reasonably (Burt et al., 2007). Subjects learned relatively more from their interactive tutor and his early partner than from the non-interactive one (Burt et al., 2007). More recently, it has been shown that wild savannah sparrows can learn novel songs broadcast by speakers but that songs broadcast in a simulated countersinging context were not better learned than songs broadcast by speakers simulating solo singing (Mennil et al., 2019). To perform this experiment, the authors created playback stimuli by assembling tutor songs into stereo playback tracks. Each track featured two



333 simulated tutors engaged in simultaneous bouts of interactive singing, or in alternating bouts  
334 of non-interactive singing.

#### 336 4.4.Towards an integrative approach to study the influence of multimodality: use of robots

337 As mentioned above, birdsong learning is a multimodal task involving behavioral contingency  
338 and several studies have emphasized that a live tutor could be better than taped playback or a  
339 computer to train a bird to imitate songs (Baptista & Petrinovitch, 1986; Derégnaucourt et al.,  
340 2013). More recently, Mets and Brainard (2018b) demonstrated that increasing the richness of  
341 the learning experience from poor (tutoring by computer) to a richer level of interactions  
342 (tutoring by a live bird) reduces the genetic influence in favor of experiential influence in  
343 Bengalese finches. Experiments involving a human tutor are the most emblematic to illustrate  
344 this phenomenon, e.g. those of Jürgen Nicolai training bullfinches to imitate whistled  
345 melodies (Nicolai et al., 2014). To study the multimodal aspects of song tutoring, one should  
346 be able to handle each behaviour of the tutor independently, and assess the impact of each  
347 stimulus not only on the establishment and maintenance of the relationship between the young  
348 bird and the adult bird, but also on the success of song imitation. There is a growing interest  
349 for the use of robots in ethology (see Romano et al., 2019 for a review) and ethorobotics -  
350 even if it is primarily focused on human-robot relationships - is a fast-growing discipline. For  
351 example, in birds, robots were used successfully to study courtship behaviour in bowerbirds  
352 (Patricelli et al., 2002, 2006; Perry et al., 2019), filial imprinting in juveniles quails (Jolly et  
353 al., 2016), collective and social behaviour in chickens (Gribovskiy et al., 2018) and public  
354 information in starlings (Butler & Fernández-Juricic, 2014). We recently developed a robotic  
355 zebra finch that we used as a song tutor for young finches. The robot is that of a life-sized  
356 zebra finch and is painted to represent an adult male; it is fixed on an axis and can move  
357 different parts of its plastic body. It can be programmed to behave contingently: for example,

when the young bird sits on a perch close to it, it moves towards him and sings. The program is also implemented with a vocal loop (see above): when the bird produces a contact call, the robot replied with a contact call and it is possible to manipulate the latency of the response. Using such a contingent program, we observed that young finches exposed to a robot one hour a day for five days a week during five weeks provided a good imitation of the song model. We did not find a significant difference in song imitation between birds tutored with a robot and birds tutored by a live male (Araguas et al., 2022), this latter condition being known until now as the best way so far to get a close-to-perfect imitation of a song model in zebra finches (Tchernichovski et al., 1999; Derégnaucourt et al., 2013).

## **5. Conditioning experiments**

The usual paradigm used in conditioning experiments consists in providing a reward for good responses (e.g. food) and a penalty for bad responses (lights off or electric shock). In go-no go procedures, animals are usually food-restricted to make sure that they would be motivated to peck on keys or press on levers to participate in experiments. Recently, new techniques have been developed by taking into account the socio-sexual tendencies of zebra finches. In these scenarios, finches do not work for food but to see and/or interact vocally with conspecifics. For example, Tokarev and Tchernichovski (2014) developed a new design to measure auditory discrimination (Figure 3).

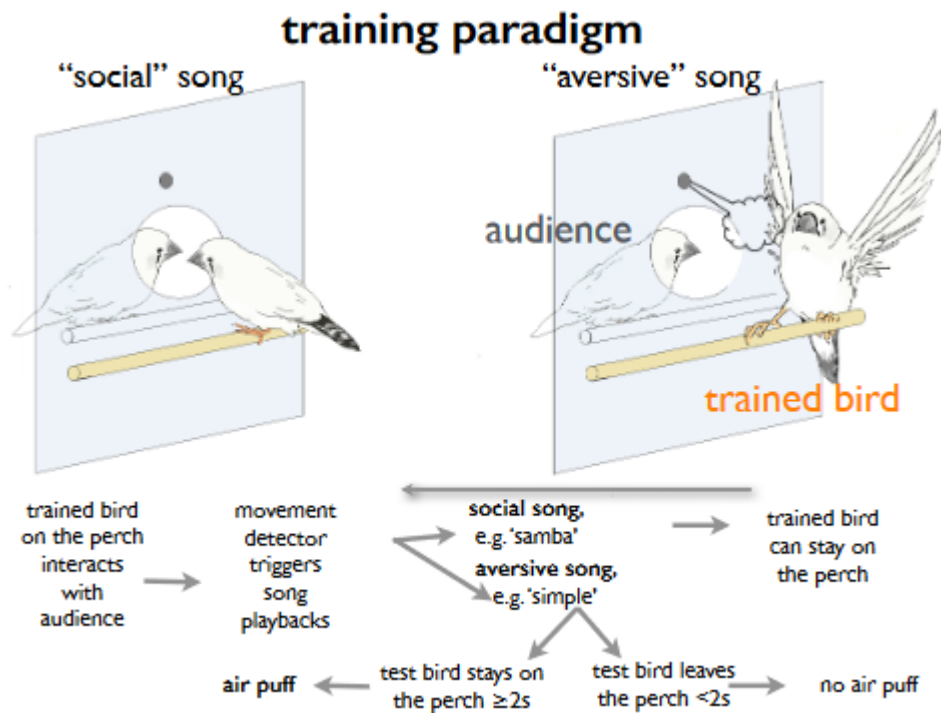


Figure 3. Social reinforcement apparatus: playbacks of the ‘social’ song allow birds to socialize via a window; playbacks of the ‘aversive’ song are followed by an air puff, which the bird learns to avoid. Artwork courtesy of Lotem Tchernichovski. Source: Tokarev and Tchernichovski, *BioRxiv* : 004176, 2014.

A male is allowed to interact with a female while perching next to a small window. Playbacks of one syllable were followed by an air puff (aversive syllable), causing the bird to escape to the safe perch, while playbacks of the other syllable allowed the interaction to continue. Using this paradigm, it was shown that learning auditory discriminations from observation is efficient but less robust than learning from experience (Narula et al., 2018). The authors trained an “experimenter” zebra finch to discriminate between auditory stimuli in the presence of an “observer” finch. Experimenters were slow to successfully discriminate the stimuli, but immediately generalize their ability to a new set of similar stimuli. By contrast, observers subjected to the same task were able to discriminate the initial stimulus set, but require more

time for successful generalization (Narula et al., 2018). Taking again into account the natural tendency for isolated finches to seek socio-sexual contacts, Macedo-Lima and Remage-Healey (2020) replaced the food reward by a social reward for an experiment to measure auditory discrimination in zebra finches. A male and a female were put in adjacent cages separated by a sheet of opaque polarized glass. If a GO tone was played, the glass would turn transparent for 6 s, resulting in a HIT (i.e., period of visual engagement with the adjacent social partner through the transparent barrier). The authors did not document whether training was faster with a social reward rather than a food reward but it is likely that a dedicated experiment to test this prediction should be run in order to answer this question.

## **6. Future directions**

Birdsong has been established as a biological model of choice to study vocal learning and, as a consequence, the goal of several experimental procedures is to enhance song imitation or even to obtain a perfect copy of the model (Derégnaucourt et al., 2013; Mets and Brainard, 2019). The robotic zebra finch, which permits both vocal and behavioural contingencies with a young bird's behaviour, offers new opportunities to reach this goal and to explore meticulously the multimodal cues involved in birdsong learning (Araguas et al., 2022). Such multimodal aspects of communication in birds could also be explored using other computer systems, such as the one recently developed to control vocal exchanges in a group (Rychen et al., 2021); as mentioned by the authors, it would be easy to combine the visual communication channel using cameras and computer screens. Despite the importance of flight in the life of birds, birdsong has mainly been studied in perching birds, and vocal production learning in flight presents unique challenges and opportunities for birds and for researchers (Berg et al., 2019; Arnold et al., 2022). Affordable, high-speed, high-resolution cameras and

mic-video array technology now available to tackle a number of questions in spatial and biomechanical issues related to highly mobile communication strategies.

Since the pioneering works of Adret (1993a, b) showing that young male zebra finches learn quickly to press a button to trigger song playback, several experiments have highlighted the active role of young songbirds on birdsong learning. Automatic learning devices also have potential to advance the study of social learning. For example, devices such as the one used with monkeys by Fagot and Paleressompoulle (2009) could allow experiments in which birds living in a social group choose when to participate.

Experiments involving young birds raised in social isolation during several weeks or months raise ethical issues. In ethology, this experimental procedure is still considered as the baseline reference conditions even if it exhibits several pitfalls already pinpointed a long time ago (Lehrman, 1953). More recently, it has been shown that acute social isolation has immediate consequences on brain gene activity in zebra finches (George et al., 2020). The use of biomimetic robots appear as a promising avenue to fulfill the 3Rs' objectives (Replace, Reduce, Refine: Replacement alternatives refer to methods which avoid or replace the use of animals, reduction alternatives refer to any strategy that will result in fewer animals being used to obtain sufficient data to answer the research question, refinement alternatives refer to the modification of husbandry or experimental procedures to minimize pain and distress and to improve welfare; Bierbach et al., 2021); in birdsong research, robots could replace live companions and therefore potentially reduce the stress of social isolation without compromising highly standardized conditions. Interactions between wild birds and machines open also exciting opportunities of research such as the ability to track cultural evolution of birdsong following the introduction of new song models, but it also raises ethical questions (Mennill et al., 2019). The use of biomimetic robots could also be relevant in the field of biological conservation. For example, a recent study has shown that young regent honeyeaters

(*Anthochaera phrygia*) are more likely to learn songs from heterospecific models due to the lack of sufficient conspecific adult tutors (Crates et al., 2021). Robots mimicking conspecific adults could be used for song tutoring and could permit the maintenance of endangered dialects.

The use of Virtual Reality (VR) is also a promising avenue to study the learning capacities of birds in the vocal domain, especially in the field of birdsong learning (Ljubičić et al., 2016). For example, birds could be exposed to a Virtual Social Environment (VSE) that simulate the social context of song learning, through iterative interactions between the developing bird and a computerized visual and vocal interface comprising one or several birds (Ljubičić et al., 2016). Such a use is still at its infancy for behavioural studies in non-human animals. Studies on insects have pioneered the implementation of VR paradigms to study perceptual and cognitive capacities (Lafon et al., 2021).

In conclusion, our knowledge on learning capacities of birds in the vocal and auditory domains has been rapidly growing these last 20 years thanks to technological developments. Many setups developed recently are based on inexpensive devices and free and open-source softwares. Cloud platforms permit the storage of huge amount of data that can be easily shared and accessed by researchers. Such tools used both in the laboratory and the field will permit researchers to get more insight into the evolutionary processes that led to species differences regarding learning abilities and to explore more deeply the Umwelt of birds. These new techniques offer some opportunities to improve our knowledge of birds' behaviour in more natural settings, ultimately improving some ethical aspects of experimental research.

## Litterature cited

- Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour*, 46, 149-159. <https://doi.org/10.1006/anbe.1993.1170>
- Adret, P. (1993). Vocal learning induced with operant techniques: an overview. *Netherlands Journal of Zoology*, 43, 125-125.
- Adret, P. (1997) Discrimination of video images by zebra finches (*Taeniopygia guttata*): direct evidence from song performance. *Journal of Comparative Psychology*, 111, 115-125. <https://doi.org/10.1037/0735-7036.111.2.115>
- Akçay, Ç. L., Tom, M. E., Campbell, S. E., & Beecher, M. D. (2013). Song type matching is an honest early threat signal in a hierarchical animal communication system. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20122517. <https://doi.org/10.1098/rspb.2012.2517>
- Anisimov, V. N., Herbst, J. A., Abramchuk, A. N., Latanov, A. V., Hahnloser, R. H., & Vyssotski, A. L. (2014). Reconstruction of vocal interactions in a group of small songbirds. *Nature Methods*, 11, 1135-1137. doi: 10.1038/nmeth.3114
- Araguas, A., Guellai, B., Gauthier, P., Richer, F., Montone, G., Chopin, A., & Derégnaucourt, S. (2022). Design of a robotic zebra finch for experimental studies on developmental song learning. *Journal of Experimental Biology*, 225, jeb242949.
- Arnold, F., Staniszewski, M. S., Pelzl, L., Ramenda, C., Gahr, M., & Hoffmann, S. (2022). Vision and vocal communication guide three-dimensional spatial coordination of zebra finches during wind-tunnel flights. *Nature Ecology & Evolution*, 6, 1221-1230. <https://doi.org/10.1242/jeb.242949>

Balsby, T. J., & Bradbury, J. W. (2009). Vocal matching by orange-fronted conures (*Aratinga canicularis*). *Behavioural Processes*, 82, 133-139. <https://doi.org/10.1016/j.beproc.2009.05.005>

Baptista, L.F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow: social factors and sex differences. *Animal Behaviour*, 34, 1359-1371. [https://doi.org/10.1016/S0003-3472\(86\)80207-X](https://doi.org/10.1016/S0003-3472(86)80207-X)

Beecher, M. D., Campbell, S. E., Burt, J. M., Hill, C. E., & Nordby, J. C. (2000). Song-type matching between neighbouring song sparrows. *Animal Behaviour*, 59, 21-27. <https://doi.org/10.1006/anbe.1999.1276>

Beecher, M. D., & Akçay, Ç. (2021). Social factors in bird-song development: Learning to sing with friends and rivals. *Learning & Behavior*, 49, 137-149. <https://doi.org/10.3758/s13420-020-00441-6>

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, 26, 309-318. <https://doi.org/10.1016/j.cub.2015.12.037>

Berg, K. S., Delgado, S., & Mata-Betancourt, A. (2019). Phylogenetic and kinematic constraints on avian flight signals. *Proceedings of the Royal Society B*, 286, 20191083. <https://doi.org/10.1098/rspb.2019.1083>

Bierbach, D., Francisco, F., Lukas, J., Landgraf, T., Maxeiner, M., Romanczuk, P., ... & Krause, J. (2021). Biomimetic robots promote the 3Rs Principle in animal testing. In *ALIFE 2021: The 2021 Conference on Artificial Life*. MIT Press. [https://doi.org/10.1162/isal\\_a\\_00375](https://doi.org/10.1162/isal_a_00375)



Birkhead, T. R., & Van Balen, S. (2008). Bird-keeping and the development of ornithological science. *Archives of Natural History*, 35, 281-305. <https://doi.org/10.3366/E0260954108000399>

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., ... & Kirschel, A. N. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48, 758-767. doi: 10.1111/j.1365-2664.2011.01993.x

Bolhuis, J.J. & Everaert, M. (2013). Birdsong, speech, and language: exploring the evolution of mind and brain. MIT press. <https://doi.org/10.7551/mitpress/9322.001.0001>

Burt, J. M., O’Loghlen, A. L., Templeton, C. N., Campbell, S. E., & Beecher, M. D. (2007). Assessing the importance of social factors in bird song learning: a test using computer- simulated tutors. *Ethology*, 113, 917-925. doi: 10.1111/j.1439-0310.2007.01415.x

Butler, S. R., & Fernández-Juricic, E. (2014). European starlings recognize the location of robotic conspecific attention. *Biology Letters*, 10, 20140665. <https://doi.org/10.1098/rsbl.2014.0665>

Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology*, 29, 631-636. <https://doi.org/10.1016/j.cub.2018.12.026>

Clouzot, M., & Kreutzer, M. (2020). Croiser l’histoire des savoirs empiriques et les connaissances éthologiques pour concevoir le point de vue des oiseaux chanteurs. In : Croiser les sources pour lire les animaux (ed. E. Baratay). Presses Universitaires de la Sorbonne, pp.145-157.

Colavita, F.B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16, 409-412. <https://doi.org/10.3758/BF03203962>

Crates, R., Langmore, N., Ranjard, L., Stojanovic, D., Rayner, L., Ingwersen, D., & Heinsohn, R. (2021). Loss of vocal culture and fitness costs in a critically endangered songbird. *Proceedings of the Royal Society B*, 288, 20210225. <https://doi.org/10.1098/rspb.2021.0225>

Derégnaucourt, S. (2011). Birdsong learning in the laboratory, with especial reference to the song of the Zebra Finch (*Taeniopygia guttata*). *Interaction Studies*, 12, 324-350. <https://doi.org/10.1075/is.12.2.07der>

Derégnaucourt, S., Mitra, P. P., Fehér, O., Pytte, C., & Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. *Nature*, 433, 710-716. <https://doi.org/10.1038/nature03275>

Derégnaucourt, S., Poirier, C., Van der Kant, A., Van der Linden, A., & Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *Journal of Physiology-Paris*, 107, 210-218. <https://doi.org/10.1016/j.jphysparis.2012.08.003>

Derégnaucourt, S., Saar, S., & Gahr, M. (2009). Dynamics of crowing development in the domestic Japanese quail (*Coturnix coturnix japonica*). *Proceedings of the Royal Society B: Biological Sciences*, 276, 2153-2162. <https://doi.org/10.1098/rspb.2009.0016>

Derégnaucourt, S., Saar, S., & Gahr, M. (2012). Melatonin affects the temporal pattern of vocal signatures in birds. *Journal of Pineal Research*, 53, 245-258. <https://doi.org/10.1111/j.1600-079X.2012.00993.x>

Deshpande, M., Pirlepsov, F., & Lints, T. (2014). Rapid encoding of an internal model for imitative learning. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20132630. <https://doi.org/10.1098/rspb.2013.2630>

Desmedt, L., George, I., Mohamed Benkada, A., Hervé, M., Aubin, T., Derégnaucourt, S., & Lumineau, S. (2020). Maternal presence influences vocal development in the Japanese quail (*Coturnix c. japonica*). *Ethology*, 126, 553-562. <https://doi.org/10.1111/eth.13007>

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. <https://doi.org/10.1146/annurev.neuro.22.1.567>

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, 41, 396-404. <https://doi.org/10.3758/BRM.41.2.396>

Fehér, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature*, 459, 564-568. <https://doi.org/10.1038/nature07994>

Gardner, T. J., Naef, F., & Nottebohm, F. (2005). Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science*, 308, 1046-1049. <https://doi.org/10.1126/science.1108214>

George, J. M., Bell, Z. W., Condliffe, D., Dohrer, K., Abaurrea, T., Spencer, K., ... & Clayton, D. F. (2020). Acute social isolation alters neurogenomic state in songbird forebrain. *Proceedings of the National Academy of Sciences*, 117, 23311-23316. doi: 10.1073/pnas.1820841116.

Gill, L. F., Goymann, W., Ter Maat, A., & Gahr, M. (2015). Patterns of call communication between group-housed zebra finches change during the breeding cycle. *Elife*, 4, e07770. <https://doi.org/10.7554/eLife.07770>

Gill, L. F., D'Amelio, P. B., Adreani, N. M., Sagunsky, H., Gahr, M. C., & Ter Maat, A. (2016). A minimum- impact, flexible tool to study vocal communication of small animals with precise individual- level resolution. *Methods in Ecology and Evolution*, 7, 1349-1358. <https://doi.org/10.1111/2041-210X.12610>

Gribovskiy, A., Halloy, J., Deneubourg, J. L., & Mondada, F. (2018). Designing a socially integrated mobile robot for ethological research. *Robotics and Autonomous Systems*, 103, 42-55. <https://doi.org/10.1016/j.robot.2018.02.003>

Hervieux de Chanteloup, J. C. (1709). Nouveau traité des serins de canarie. Paris : Claude Prudhomme.

Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., ... & Gahr, M. (2019). Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nature communications*, 10, 1-11. <https://doi.org/10.1038/s41467-019-10593-3>

Hyland Bruno, J., Jarvis, E. D., Liberman, M., & Tchernichovski, O. (2021). Birdsong learning and culture: analogies with human spoken language. *Annual Review of Linguistics*, 7, 449-472. <http://dx.doi.org/10.1146/annurev-linguistics-090420-121034>

Hultsch, H., Schleuss, F., & Todt, D. (1999). Auditory–visual stimulus pairing enhances perceptual learning in a songbird. *Animal Behaviour*, 58, 143-149. <https://doi.org/10.1006/anbe.1999.1120>

Ikebuchi, M., & Okanoya, K. (1999). Male zebra finches and Bengalese finches emit directed songs to the video images of conspecific females projected onto a TFT display. *Zoological Science*, 16, 63-70. <https://doi.org/10.2108/zsj.16.63>

James, L. S., & Sakata, J. T. (2017). Learning biases underlie “universals” in avian vocal sequencing. *Current Biology*, 27, 3676-3682. <https://doi.org/10.1016/j.cub.2017.10.019>

Janik, V. M., & Slater, P. J. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour*, 26, 59-100. [https://doi.org/10.1016/S0065-3454\(08\)60377-0](https://doi.org/10.1016/S0065-3454(08)60377-0)

Jolly, L., Pittet, F., Caudal, J. P., Mouret, J. B., Houdelier, C., Lumineau, S., & de Margerie, E. (2016). Animal-to-robot social attachment: initial requisites in a gallinaceous bird. *Bioinspiration & biomimetics*, 11, 016007. <https://doi.org/10.1088/1748-3190/11/1/016007>

Jones, A. E., ten Cate, C., & Bijleveld, C. C. (2001). The interobserver reliability of scoring sonagrams by eye: a study on methods, illustrated on zebra finch songs. *Animal Behaviour*, 4, 791-801. <https://doi.org/10.1006/anbe.2001.1810>

Lachlan, R. F., Verhagen, L., Peters, S., & Cate, C. T. (2010). Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taeniopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *Journal of Comparative Psychology*, 124, 92. <https://doi.org/10.1037/a0016996>

Lafon, G., Howard, S. R., Paffhausen, B. H., Avarguès-Weber, A., & Giurfa, M. (2021). Motion cues from the background influence associative color learning of honey bees in a virtual-reality scenario. *Scientific Reports*, 11, 1-20. <https://doi.org/10.1038/s41598-021-00630-x>

Lehrman, D. S. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. *The Quarterly Review of Biology*, 28, 337-363. <http://www.jstor.org/stable/2813453>

Lerch, A., Roy, P., Pachet, F., & Nagle, L. (2011). Closed-loop bird–computer interactions: a new method to study the role of bird calls. *Animal Cognition*, 14, 203-211. <https://doi.org/10.1007/s10071-010-0353-6>

Levinson, S. C. (2016). Turn-taking in human communication—origins and implications for language processing. *Trends in Cognitive Sciences*, 20, 6-14. <https://doi.org/10.1016/j.tics.2015.10.010>

Lipkind, D., Marcus, G. F., Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M., ... & Tchernichovski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*, 498, 104-108. <https://doi.org/10.1038/nature12173>

Lipkind, D., Zai, A. T., Hanuschkin, A., Marcus, G. F., Tchernichovski, O., & Hahnloser, R. H. (2017). Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nature communications*, 8, 1-11. <https://doi.org/10.1038/s41467-017-01436-0>

Lipkind, D., Geambasu, A., & Levelt, C. C. (2020). The development of structured vocalizations in songbirds and humans: a comparative analysis. *Topics in Cognitive Science*, 12, 894-909. <https://doi.org/10.1111/tops.12414>

Ljubičić, I., Bruno, J. H., & Tchernichovski, O. (2016). Social influences on song learning. *Current Opinion in Behavioral Sciences*, 7, 101-107. <https://doi.org/10.1016/j.cobeha.2015.12.006>

Macedo-Lima, M., & Ramage-Healey, L. (2020). Auditory learning in an operant task with social reinforcement is dependent on neuroestrogen synthesis in the male songbird auditory cortex. *Hormones and behavior*, 121, 104713. <https://doi.org/10.1016/j.yhbeh.2020.104713>

Marler, P. (1970). Birdsong and speech development: Could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man. *American Scientist*, 58, 669-673. <https://www.jstor.org/stable/27829317>

Mennill, D. J., Doucet, S. M., Newman, A. E., Williams, H., Moran, I. G., Thomas, I. P., ... & Norris, D. R. (2018). Wild birds learn songs from experimental vocal tutors. *Current Biology*, 28, 3273-3278. <https://doi.org/10.1016/j.cub.2018.08.011>

Mennill, D. J., Doucet, S. M., Newman, A. E., Williams, H., Moran, I. G., Thomas, I. P., ... & Norris, D. R. (2019). Eavesdropping on adult vocal interactions does not enhance juvenile song learning: an experiment with wild songbirds. *Animal Behaviour*, 155, 67-75. <https://doi.org/10.1016/j.anbehav.2019.05.019>

Mets, D. G., & Brainard, M. S. (2018a). An automated approach to the quantitation of vocalizations and vocal learning in the songbird. *PLoS Computational Biology*, 14, e1006437. <https://doi.org/10.1371/journal.pcbi.1006437>

Mets, D. G., & Brainard, M. S. (2018b). Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences*, 115, 421-426. <https://doi.org/10.1073/pnas.1713031115>

Mets, D. G., & Brainard, M. S. (2019). Learning is enhanced by tailoring instruction to individual genetic differences. *Elife*, 8, e47216. <https://doi.org/10.7554/eLife.47216>

Mulligan, J. A., & Olsen, K. C. (1969). Communication in canary courtship calls. Bird vocalisation. Cambridge University Press, Cambridge, 165-184.

Narula, G., Herbst, J. A., Rychen, J., & Hahnloser, R. H. (2018). Learning auditory discriminations from observation is efficient but less robust than learning from experience. *Nature Communications*, 9, 1-11. <https://doi.org/10.1038/s41467-018-05422-y>

Nicolai, J., Gundacker, C., Teeselink, K., & Güttinger, H. R. (2014). Human melody singing by bullfinches (*Pyrrhula pyrrula*) gives hints about a cognitive note sequence processing. *Animal Cognition*, 17, 143-155. <https://doi.org/10.1007/s10071-013-0647-6>

Nobleville de, L. D. A. (1751). *Aëdologie, ou Traité du rossignol franc ou chanteur*. Paris : Debure l'ainé.

Okanoya, K., & Kimura, T. (1993). A software bird call detector and its application to automated playback experiments. *Bioacoustics*, 5, 117-122. <https://doi.org/10.1080/09524622.1993.9753233>

Patricelli, G. L., Uy, J. A. C., Walsh, G., & Borgia, G. (2002). Male displays adjusted to female's response. *Nature*, 415, 279-280. <https://doi.org/10.1038/415279a>

Patricelli, G. L., Coleman, S. W., & Borgia, G. (2006). Male satin bowerbirds, *Ptilonorhynchus violaceus*, adjust their display intensity in response to female startling: an experiment with robotic females. *Animal Behaviour*, 71, 49-59. <https://doi.org/10.1016/j.anbehav.2005.03.029>

Pepperberg, I. M. (1985). Social modeling theory: A possible framework for understanding avian vocal learning. *The Auk*, 102, 854-864. <https://www.jstor.org/stable/4086722>

Perry, A. C., Krakauer, A. H., McElreath, R., Harris, D. J., & Patricelli, G. L. (2019). Hidden Markov models reveal tactical adjustment of temporally clustered courtship displays in response to the behaviors of a robotic female. *The American Naturalist*, 194, 1-16. <https://doi.org/10.1086/703518>.

Poirier, C., Henry, L., Mathelier, M., Lumineau, S., Cousillas, H., & Hausberger, M. (2004). Direct social contacts override auditory information in the song-learning



process in starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology*, 118, 179.  
<https://doi.org/10.1037/0735-7036.118.2.179>

Romano, D., Donati, E., Benelli, G., & Stefanini, C. (2019). A review on animal–robot interaction: from bio-hybrid organisms to mixed societies. *Biological cybernetics*, 113, 201-225. <https://doi.org/10.1007/s00422-018-0787-5>

Rychen, J., Rodrigues, D. I., Tomka, T., Rüttimann, L., Yamahachi, H., & Hahnloser, R. H. (2021). A system for controlling vocal communication networks. *Scientific Reports*, 11, 1-15. <https://doi.org/10.1038/s41598-021-90549-0>

Stowell, D., & Sueur, J. (2020). Ecoacoustics: acoustic sensing for biodiversity monitoring at scale. *Remote Sensing in Ecology and Conservation*, 6, 217-219.

Shank, S. S., & Margoliash, D. (2009). Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature*, 458, 73-77.  
<https://doi.org/10.1038/nature07615>

Stresemann, E. (1947). Baron von Pernau, pioneer student of bird behavior. *The Auk*, 64, 35-52. <https://doi.org/10.1002/rse2.174>

Tchernichovski, O., Lints, T., Mitra, P. P., & Nottebohm, F. (1999). Vocal imitation in zebra finches is inversely related to model abundance. *Proceedings of the National Academy of Sciences*, 96, 12901-12904. <https://doi.org/10.1073/pnas.96.22.12901>.

Tchernichovski, O., Lints, T. J., Derégnaucourt, S., Cimenser, A., & Mitra, P. P. (2004). Studying the song development process: rationale and methods. *Annals of the New York Academy of Sciences*, 1016, 348-363.  
<https://doi.org/10.1196/annals.1298.031>.

Ten Cate, C. (2021). Re-evaluating vocal production learning in non-oscine birds. *Philosophical Transactions of the Royal Society B*, 376, 20200249.  
<https://doi.org/10.1098/rstb.2020.0249>

Todt, D. (1975). Social learning of vocal patterns and modes of their application in grey parrots (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie*, 39, 178-188. <https://doi.org/10.1111/j.1439-0310.1975.tb00907.x>

Tokarev, K. & Tchernichovski, O. (2014). A novel paradigm for auditory discrimination training with social reinforcement in songbirds. *BioRxiv* : 004176. <https://doi.org/10.1101/004176>

Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B*, 375, 20180406. <https://doi.org/10.1098/rstb.2018.0406>

Varkevisser, J. M., Simon, R., Mendoza, E., How, M., van Hijkema, I., Jin, R., ... & Riebel, K. (2021). Adding colour-realistic video images to audio playbacks increases stimulus engagement but does not enhance vocal learning in zebra finches. *Animal cognition*, 1-26. <https://doi.org/10.1007/s10071-021-01547-8>

Vehrencamp, S. L. (2001). Is song-type matching a conventional signal of aggressive intentions? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268, 1637-1642. <https://doi.org/10.1098/rspb.2001.1714>

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., ... & Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, 25, 495-499. <https://doi.org/10.1016/j.cub.2014.12.032>

747        **Acknowledgements**

748        We would like to thank Irene Pepperberg for giving us to opportunity to write this  
749        review. We thank Kazuo Okanoya and two anonymous reviewers for their valuable  
750        comments and the Institut Universitaire de France for financial support.

751