

## Article

# Vocal Learning and Behaviors in Birds and Human Bilinguals: Parallels, Divergences and Directions for Research

Jon T. Sakata <sup>1,\*</sup>  and David Birdsong <sup>2</sup><sup>1</sup> Department of Biology, Faculty of Science, McGill University, Montreal, QC H3A 1B1, Canada<sup>2</sup> Department of French and Italian, College of Liberal Arts, University of Texas at Austin, Austin, TX 78712, USA; birdsong@austin.utexas.edu

\* Correspondence: jon.sakata@mcgill.ca

**Abstract:** Comparisons between the communication systems of humans and animals are instrumental in contextualizing speech and language into an evolutionary and biological framework and for illuminating mechanisms of human communication. As a complement to previous work that compares developmental vocal learning and use among humans and songbirds, in this article we highlight phenomena associated with vocal learning subsequent to the development of primary vocalizations (i.e., the primary language (L1) in humans and the primary song (S1) in songbirds). By framing avian “second-song” (S2) learning and use within the human second-language (L2) context, we lay the groundwork for a scientifically-rich dialogue between disciplines. We begin by summarizing basic birdsong research, focusing on how songs are learned and on constraints on learning. We then consider commonalities in vocal learning across humans and birds, in particular the timing and neural mechanisms of learning, variability of input, and variability of outcomes. For S2 and L2 learning outcomes, we address the respective roles of age, entrenchment, and social interactions. We proceed to orient current and future birdsong inquiry around foundational features of human bilingualism: L1 effects on the L2, L1 attrition, and L1<->L2 switching. Throughout, we highlight characteristics that are shared across species as well as the need for caution in interpreting birdsong research. Thus, from multiple instructive perspectives, our interdisciplinary dialogue sheds light on biological and experiential principles of L2 acquisition that are informed by birdsong research, and leverages well-studied characteristics of bilingualism in order to clarify, contextualize, and further explore S2 learning and use in songbirds.

**Keywords:** birdsong; vocal learning; vocal performance; speech; second language acquisition; bilingualism; age; variability



**Citation:** Sakata, Jon T., and David Birdsong. 2022. Vocal Learning and Behaviors in Birds and Human Bilinguals: Parallels, Divergences and Directions for Research. *Languages* 7: 5. <https://doi.org/10.3390/languages7010005>

Academic Editors: Juana M. Liceras and Raquel Fernández Fuertes

Received: 23 August 2021

Accepted: 14 December 2021

Published: 31 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Defining characteristics of humans may be illuminated by comparisons with animal models. Similarly, our understanding of animals is refined by comparisons with humans. In keeping with traditions of ethological inquiry (see [Gomez-Marin et al. 2014](#) for a synthesis of methods and issues; [Bradbury and Vehrencamp 2011](#) for background in animal communication), the present paper looks at songbirds and humans with respect to commonalities and divergences in vocal learning and behavior. In particular, the paper attempts to contextualize adult vocal learning and use in songbirds with second language (L2) learning and bilingualism in humans.

Much of the essential terrain comparing vocal learning in songbirds with human speech development has been covered by researchers before us (e.g., [Bolhuis et al. 2010](#); [Doupe and Kuhl 1999](#); [Gervain and Mehler 2010](#); [Sakata and Woolley 2020](#)). However, the emphasis in previous reviews has been on the development of what could be called the primary vocalizations or songs (“S1”) of a given songbird species (i.e., the developmental acquisition of vocalizations). In contrast, here we highlight features of vocal learning

and behavior *subsequent to the learning of primary vocalizations* (“S2 learning”) in songbirds, in particular bird species that have extended periods of vocal learning, and relate these features to human L2 learning and bilingual behaviors.

The present article also aims to establish greater dialogue between disciplines. In bilingualism and second language acquisition, much of the disciplinary effort is directed toward examining the relationship of learning outcomes to age of learning and to individual- and group-level differences, which may be of an internal (genetic, cognitive, and motivational, etc.) or external (experiential, input, and interaction, etc.) nature (reviewed in [Birdsong 2018](#); see also Section 5). Songbird research likewise addresses such relationships. However, the two disciplines diverge in terms of their methods and goals of inquiry, and with respect to the types and availability of evidence that could advance knowledge in their respective fields. For example, given the ability to assess and manipulate gene expression and neurochemistry in targeted brain circuits in songbirds, there is considerable emphasis on revealing the cellular and molecular underpinnings of vocal learning and use in songbirds; indeed, while such studies cannot be performed in humans, these studies in songbirds help generate and test models to explain age- or experience-dependent changes in speech acquisition in humans ([Brainard and Doupe 2013](#); [Doupe and Kuhl 1999](#); [Gobes et al. 2019](#); [Prather et al. 2017](#); [Sakata and Yazaki-Sugiyama 2020](#)). The current dialogue between disciplines leverages salient characteristics of human L2 learning and use to clarify and amplify what is already known about birdsong learning and use, and to identify areas of potentially fruitful study in birdsong research. In addition, we highlight areas in which songbird research has been and can be used to deepen our understanding of L2 learning and bilingualism.

In the following main section (Section 2), we summarize the fundamentals of current birdsong research: the scope of inquiry, the characteristics of birdsongs and songbirds, and how songs are learned. We then examine two main categories of songbird species: those whose song learning is open-ended (i.e., relatively unconstrained by age) and those that exhibit closed-ended learning. We conclude this section with a note on the sensory representation of a bird’s target song and the sensorimotor processes involved in the eventual matching of song production to this representation.

Section 3 conceptualizes and delimits our approach to comparing avian S2 learning to human L2 speech learning. We argue that experience-dependent increases and alterations of song repertoires are apt points of entry into such comparisons, allowing for discussion of common foci of timing of learning, input variability and variable outcomes. However, we also describe some experiments in species with limited vocal repertoires that provide useful backdrops to understanding processes in L2 speech acquisition. In this context, we look at vocal learning in birds with respect to the neural correlates of repertoire size and the underlying neural and genetic mechanisms.

Section 4 examines three main factors—age of learning, entrenchment and social interactions—that are shared by songbirds and humans and that are associated with variable learning outcomes. Research indicates that, as in humans’ L2 pronunciation, age of S2 learning is predictive of accuracy of imitation of target song. A high level of entrenchment of previously learned songs appears to be associated with low plasticity for learning new songs, and social interactions are necessary for faithful replication of novel songs.

In Section 5 we formulate questions for birdsong researchers around three essential characteristics of human bilingual learning and use: L1 effects on the L2, L1 attrition, and switching between languages. With respect to switching, partial parallels between humans and birds can be discerned. However, evidence is lacking for S1 effects on the S2 and for S1 attrition, and we suggest possible avenues for future investigation.

In the concluding section (Section 6), we summarize our approach and emphasize the value of discourse between songbird and bilingualism researchers.

## 2. Preliminaries

In examining parallels and dissimilarities between L2 speech acquisition and the learning of birdsong, it is important to accurately represent the behaviors of songbirds, the characteristics of birdsong, and the processes underlying vocal learning in songbirds (and other vocal learning species).

### 2.1. *The Scope of Inquiry and the Boundaries of Our Discussion*

The parallels and contrasts we draw concern the songs of birds insofar as they can be related to human *speech*, not to human *language*. Language is a complex semiotic system that can encode a range of messages that far surpasses that of other animal communication systems, including birdsong (Fitch et al. 2010; Helekar 2013). Speech, one of many distinct and separately analyzable components of language, concerns the articulatory, respiratory, and prosodic components of language. In the same manner as speech, birdsong requires the coordination of muscular and phonatory mechanisms to generate patterned, recognizable sounds (Doupe and Kuhl 1999; Sakata and Woolley 2020). Birdsong and human speech are hierarchically organized, with both segmental and suprasegmental levels, and are learned by listening to and interacting with others (reviewed in Bolhuis et al. 2010; Doupe and Kuhl 1999; Jarvis 2019; Lipkind et al. 2020; Sakata and Woolley 2020). Thus, in this paper we are able to train our lens on elements of speech and song that are similar enough to provide a basis for meaningful comparisons.

To elaborate, the acoustic properties of human speech, at both the segmental and suprasegmental levels, are defined (parameterized) for a given language. Speech sounds in a language are systematically distributed in articulatory and acoustic space to create perceptible contrasts, thereby constituting a phonological basis for distinguishing one language from another. Similarly, in terms of structural and acoustic properties (see below), the songs of different individuals of a songbird species, of different populations within a species, and of different songbird species are likewise distinct from one another. In addition, just as the acoustic dimensions of human speech—in particular, duration, pitch, intensity, and formant frequencies—are modulated by the recruitment and coordination of the larynx, tongue, lips and buccal aperture, the acoustic features of birdsongs are modulated by the shape, tension, and structure of the beak, tongue, and syringeal and respiratory muscles of birds (Elemans et al. 2015; reviewed in Podos and Sung 2020). Such variation can lead to different types of vocalizations, including imitations of the vocalizations of other species (Beckers et al. 2004) or more subtle, prosodic modulations (Mol et al. 2017).

The speech systems of languages can be acquired at different times (sequentially) or simultaneously. With respect to speech systems in simultaneous bilingualism, parallels between birds and humans are difficult to draw because the range of vocalizations a bird produces cannot be grouped into distinct systems (i.e., sets of birdsongs cannot be categorized as a “language” or “speech system”). In addition, parallels between humans and birds are difficult to draw because of the possibility that, among birds, but less so among humans, multiple sets of vocalizations are learned simultaneously but only one set of vocalizations is observed (v. infra). However, as we will argue, sequential bilingualism, wherein a second language (L2) is learned after some degree of native language proficiency is attained, can be related to types of “late” song learning in adult songbirds (i.e., vocal learning after birds demonstrate proficiency in the production of vocalizations learned during development; S2 learning). In this paper, we will focus mainly on sequential bilingualism and draw comparisons between L2 learning in humans and S2 learning in birds. Further, as elaborated below, we conceptualize L2 speech acquisition as reflecting an increase in the repertoire of sounds that individuals can produce, because repertoire building in vocal production is similarly observed in songbirds.

Finally, it is important to establish boundaries on the types of learning to be discussed here. Ethologists distinguish between different types of learning in the context of animal communication, namely vocal production learning, vocal usage learning, and vocal comprehension learning (Janik and Slater 2000; Seyfarth and Cheney 2010; Wirthlin et al. 2019;

Vernes et al. 2021); similar distinctions have been used to describe learning in the context of human communication. Vocal production learning and vocal usage learning are the most appropriate constructs to discuss here (Vocal comprehension learning, which refers to the ability to learn appropriate behavioral responses to communication signals, is not relevant in the present context). Vocal production learning refers to the ability to imitate the acoustic features of a vocalization whereas vocal usage learning concerns the ability to learn the behavioral contexts in which vocalizations (innate or acquired) are to be produced. These abilities roughly correspond to learning “how” and “when” to produce vocalizations. Understanding these various types of learning is important for thinking about the mechanisms underlying vocal learning and variation in the types of sounds animals can learn to produce. Most studies of vocal learning in songbirds focus on vocal production learning, which is most relevant to comparisons with aspects of speech acquisition in humans.

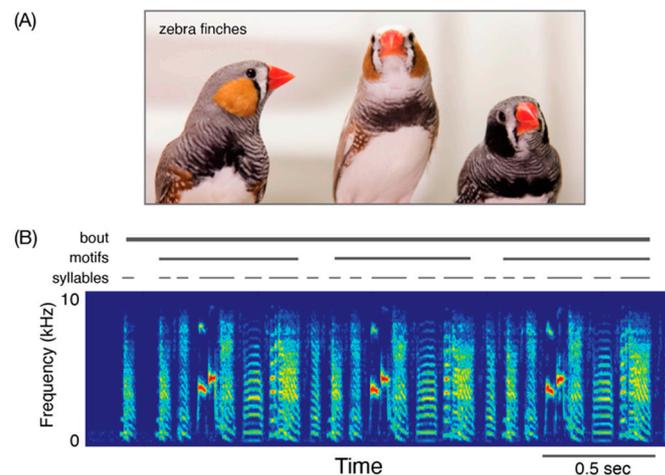
## 2.2. Songbirds and Birdsongs

Songbirds (also known as “oscines”) are a large and diverse group of bird species (~4000–5000 species) that include, for example, cardinals, robins, sparrows, finches and crows (Beecher and Brenowitz 2005; Jarvis 2019; Odom et al. 2014; Petkov and Jarvis 2012; Riebel et al. 2019; Sakata and Woolley 2020). Although birds in general communicate with each other using vocalizations (Catchpole and Slater 2008; Nowicki and Searcy 2014; Podos and Sung 2020), songbirds are distinct from other bird species in that they *learn* their vocalizations, and are broadly categorized as “vocal learners.”<sup>1</sup> Song learning is most often studied with regard to individuals learning the songs of their own species (conspecific song learning), but some songbird species (“mimics”) are capable of learning the songs (or calls) of other species (heterospecific song learning). On the other hand, pigeons, chickens, raptors, various waterfowl<sup>2</sup>, and many other species of birds vocally communicate with each other using innately-specified or unlearned vocalizations; accordingly, these birds are categorized as “vocal non-learners.” (In the relevant literature, bird vocalizations that do not require learning are referred to as *unlearned*. Thus, for bird species, “unlearned” does not mean a failed attempt to learn or a representational deficit, as it can in the human language context). While this nomenclature implies categorical distinctions between vocal learners and vocal non-learners, the differences may in fact be more continuous, and we refer readers to detailed overviews by Janik and Slater (2000); Petkov and Jarvis (2012); and Wirthlin et al. (2019) to further understand the continuum of vocal learning.

Investigations into the mechanisms of song learning have concentrated on a handful of species, and in almost all these species, there exist dramatic sex differences in the extent of vocal learning (reviewed in Ball and Balthazart 2020; Catchpole and Slater 2008). For example, in the zebra finch, the most commonly studied songbird species, only males learn and produce songs (Female zebra finches also vocalize but produce only unlearned vocalizations). The hormonal and genetic factors underlying differences between sexes in vocal learning have been investigated, but the exact nature and mechanisms of these biological factors remain elusive (Ball and Balthazart 2020; Choe and Jarvis 2021). Despite the fact that song production by both males and females is estimated to be the most common pattern across songbird species (Odom et al. 2014; reviewed in Riebel et al. 2019), there is a dearth of studies investigating the parameters and mechanisms of song learning in species in which both sexes learn to produce songs (Riebel et al. 2019) but regardless of the sex of the bird, the general processes of sensory and sensorimotor learning (see below) are central for the development of accurate imitations of song.

The terminology used to describe the complex vocalizations of birds can vary somewhat among researchers. In the present discussion, “birdsong” is understood as a concatenation of hierarchically organized acoustic elements. A “bout” of birdsong consists of sequences (or “phrases”) of “syllables”, which themselves can be composed of multiple “notes” (Figure 1).<sup>3</sup> Birdsong phrases can be consistent (i.e., the same sequence of syllables is produced across song renditions) or variable in sequencing, depending on the species and individual within the species (Murphy et al. 2017), and different songbird species can

produce different numbers of phrases in their song bouts (Beecher and Brenowitz 2005; Robinson et al. 2019; Sakata and Woolley 2020). In addition, some songbird species produce multiple “song types” which are distinguished from each other by syllable composition and sequencing. Because songbirds can produce numerous and acoustically distinct notes, syllables, phrases, and song types, songbird researchers often identify *repertoires* at each of these levels of song organization. Importantly, learning can be observed at each of these levels of structure.



**Figure 1.** Songbirds and the building blocks of their songs. (A) Picture of three adult male zebra finches. Zebra finches are the most extensively studied songbird with regard to vocal learning; photo credit: Raina Fan. (B) Spectrogram (frequency on the y-axis, time on the x-axis, amplitude as differences in color) of a bout of zebra finch song. A “bout” of zebra finch song consists of the repetition of a “motif,” that itself consists of a stereotyped sequence of acoustic elements called “syllables.” Individual syllables and the epochs of silence between them range from tens to hundreds of milliseconds in duration with substantial variation across individuals and across species.

### 2.3. How Are Songs Learned? The Basics

Since as early as the 18th century, researchers have documented how developmental experiences critically shape the vocalizations of certain species of birds. For example, Barrington (1773) noted how cross-fostering songbirds with a different species or raising songbirds without the opportunity to hear their own species song led to the production of songs that deviated from species-typical song. In addition, William Thorpe’s (1958) classic study of song dialects in chaffinches and Peter Marler’s (1970, 1997) seminal studies of song development in sparrows have shaped the contemporary era of birdsong research (Roberts and Mooney 2013). These and other studies revealed that songs are learned, that juvenile songbirds engage in a protracted period of vocal practice and development (“babbling”), and that auditory feedback is critical for the accurate development of song (Catchpole and Slater 2008; Marler 1997). Since those groundbreaking publications, many researchers have provided important insights into various aspects of song acquisition, including the factors that influence the fidelity of song learning, the diversity of learning strategies, and neural mechanisms that underlie song learning (Beecher and Brenowitz 2005; Bolhuis et al. 2010; Brainard and Doupe 2013; Doupe and Kuhl 1999; Sakata and Yazaki-Sugiyama 2020; Tschida and Mooney 2012).

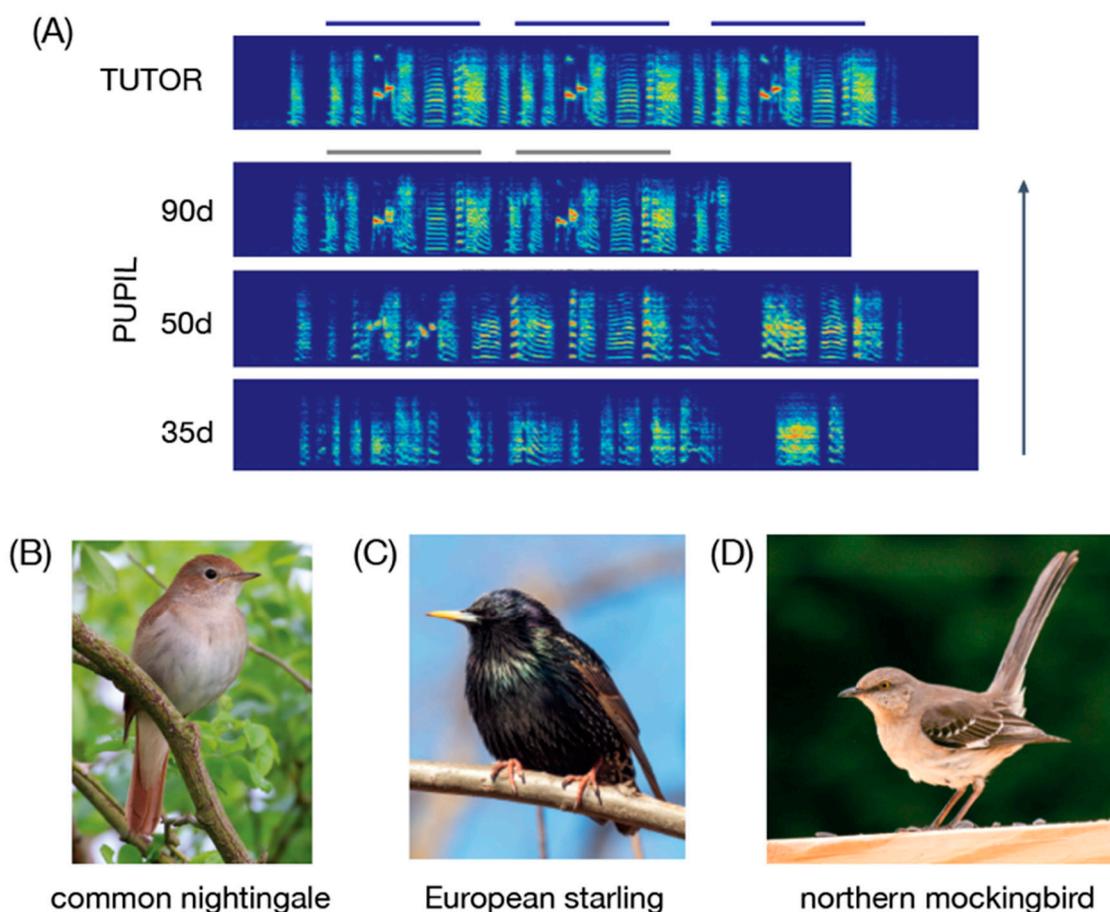
Studies of birdsong learning often analyze relationships between the vocalizations of “tutors” (birds whose vocalizations will be imitated; see below) and “pupils” (birds who are engaged in vocal learning).<sup>4</sup> A primary metric of learning is the degree of similarity in repertoires (e.g., syllable repertoires) between tutors and pupils. These assessments of repertoire overlap (e.g., how many of a tutor’s syllables are incorporated into the song of a pupil) have historically been conducted subjectively, using human classifications of syllable matches across birds, but machine or deep learning approaches have been increasingly

used to quantify repertoires and song similarities (Goffinet et al. 2021; Paul et al. 2021; Sainburg et al. 2020). In addition, similarities in the fine acoustic structure of syllables (e.g., fundamental frequencies of syllables with flat, harmonic structure), syllable sequencing and song tempo have also been evaluated to measure the degree of vocal learning (e.g., James et al. 2020; Mets and Brainard 2018, 2019). Simply put, the fidelity of song learning is generally expressed in terms of imitation at various levels of song organization.

As a side note, one could argue that defining vocal learning simply from the perspective of imitation limits the scope of inquiry. For example, innovation (e.g., to differentiate one's own vocalizations from others) could require individuals to remember the sounds of other individuals and to produce sounds that deviated from the memorized sounds. In addition, studies have systemically described how the degree of deviation between the tutor and pupil's song can depend on features of the tutor model; in general, pupils "innovate" or deviate more from tutor songs that differ from species-typical song (e.g., Fehér et al. 2009; Gardner et al. 2005; James and Sakata 2017; Tchernichovski et al. 2021). Although it could certainly be the case that innovation involves learning, differentiating innovation from "poor" or erroneous learning is very difficult, because poor learning will also lead to disparities in vocal performance between tutors and pupils. Consequently, vocal learning by songbirds is most often viewed through the lens of imitation.

Evidence of vocal learning can be gathered in various ways. For example, cross-fostering experiments, in which offspring of biological parents that produce a particular set of communicative sounds are raised by foster parents that produce a different set of sounds, have been used to document the extent, timing, and modulators of vocal learning (Gobes et al. 2019; Sakata and Woolley 2020). For example, one of the classic studies of birdsong learning used cross-fostering experiments to ask to what degree "dialects" across spatially separated populations of white-crowned sparrows are the consequence of learning or genetic variation (Marler and Tamura 1962). Other insights into vocal learning come from experiments comparing the songs of birds that were raised with or without exposure to the songs of adult tutors (e.g., Doupe and Kuhl 1999; Love et al. 2019; Marler 1997; Soha 2017) or analyzing how feedback (reinforcement) signals shape song development or plasticity (Carouso-Peck et al. 2020; reviewed in Sakata and Yazaki-Sugiyama 2020; West and King 1988).

Songbirds pass through discrete stages of song development, including stepwise acquisition of syllable sequences, before producing an accurate copy of the tutor's song(s) (Lipkind et al. 2013, 2017; Marler 1997; Tchernichovski et al. 2001). As an example of imitation-based song learning, zebra finches first memorize the song of an adult model, and then engage in a protracted period of vocal practice (Figure 2A). Through this process, the developing finch evaluates and fine-tunes its song by comparing its vocalizations to the memorized tutor song (Brainard and Doupe 2000, 2002; Sakata and Yazaki-Sugiyama 2020; Tschida and Mooney 2012). A "crystallized" and stable song is observed when zebra finches are ~90–120 days old, with individual variation in the rate of song development and accuracy of imitation (see below). Although we detail the specifics of song learning in the zebra finch, it should be mentioned that there is considerable variation in the timing and duration of song memorization and vocal practice and in the timing of song crystallization across species of songbirds (Brainard and Doupe 2000; Brenowitz and Beecher 2005; see discussions below including Figure 2B–D for species that can learn songs as adults). Indeed, understanding this evolutionary variation is a central pursuit in birdsong research.



**Figure 2.** Song development and images of songbirds commonly studied with regard to late song learning (S2). (A) As juvenile (<90 d of age) zebra finches mature, their songs become increasingly similar to those of their tutor. In other words, as juveniles engage in a protracted period of vocal practice, their songs become more similar to their tutor's song. In this example, the 35d-old "pupil" (in this case, the offspring of the "tutor") produces an acoustically variable sequence of noisy syllables that bears little resemblance to the tutor's song. By the time he is 50d-old, the syllables have more defined harmonic structure and seem to show some acoustic similarity to the tutor's song. By the time the bird is 90d-old, he produces a song with stereotyped sequences of syllables with distinct acoustic structure and that resembles the song of the tutor. (B–D) Images of three songbird species that have been studied with regard to late song learning; from Wikimedia Commons <https://commons.wikimedia.org/> (accessed on 13 December 2021).

The nature of song learning and species variation in song structure is not simply a consequence of auditory experiences but reflects the interaction of idiosyncratic experiences and biological predispositions. Evidence of predisposition is abundant and comes primarily from laboratory studies. For example, juvenile birds are generally predisposed to learn the songs of their own species; if juveniles are provided with the opportunity to learn the songs of their own species ("conspecific song") and the songs of other species ("heterospecific song"), they will preferentially learn the conspecific song. Although songbirds can learn the acoustic structure of heterospecific syllables (i.e., syllables within songs of a different species), they tend to organize these heterospecific syllables into temporal patterns that match their own species' songs (reviewed in Sakata and Yazaki-Sugiyama 2020). In addition, songbirds raised without exposure to birdsong do not simply produce random patterns of sounds but produce songs that have spectral and temporal features that coarsely resemble species-specific songs (James et al. 2020; reviewed in Love et al. 2019; Marler 1997; Soha 2017).

It is estimated that 15–20% of songbirds worldwide could be considered as “vocal mimics” (Baylis 1982; Dalziell et al. 2015; Kelley et al. 2008), with vocal mimicry differing from canonical song learning in that it is characterized by the imitation of sounds created by other species as well as anthropogenic sounds (Dalziell et al. 2015; Goller and Shizuka 2018).<sup>5</sup> For example, lyrebirds are virtuoso mimics that are able to reproduce construction noises, car sirens, and human voices; streaked bowerbirds have been observed to imitate dogs barking and trees being chopped and crashing to the ground; and magpies have been found to imitate human speech.<sup>6</sup> Although learned mimicry is often considered separately from canonical vocal learning, we consider both variants of vocal learning together because they each rely on the ability to imitate a sound. Indeed, mimicry can simply be viewed as a relaxation of species-specificity of song learning. Another reason to discuss these two forms of learning together is that biases and constraints are also observed for learned vocal mimicry. For example, a survey of mimicked sounds produced by European starlings, which are found around the world, report that starlings mimicked similar types of sounds regardless of their location and that many of these sounds resembled whistles naturally observed in their species-typical song (Hausberger et al. 1991). In addition, the best predictor of the type of heterospecific sounds that northern mockingbirds will imitate is how closely those heterospecific sounds resemble mockingbird-specific song (Gammon 2013).

#### 2.4. Closed-Ended Learning vs. Open-Ended Learning

Birdsong learning is broadly categorized as either “closed-ended learning” or “open-ended learning,” with the primary distinction being the age ranges in which birds can acquire novel vocalizations (e.g., Beecher and Brenowitz 2005; Brainard and Doupe 2002; Robinson et al. 2019). Most studies of birdsong learning have focused on closed-ended learners, species that learn their vocalization during a restricted and limited period in development (i.e., a “critical period” or “sensitive period”). These closed-ended learners, which include species such as zebra finches, Bengalese finches, and canaries, learn their songs during a critical or sensitive period for song learning<sup>7</sup> and do not incorporate new elements into their song after they reach sexual maturity. In contrast, open-ended learners (e.g., European starlings) are species that can learn new vocalizations throughout their lifetime. Historically, some species have been classified as open-ended learners because their repertoire size increases over time (Kipper and Kiefer 2010; Robinson et al. 2019), but, as we discuss below, increases in repertoire size and changes to repertoire composition in adulthood could be traced to factors orthogonal to adult vocal learning. Finally, it should be noted that the distinction between closed- or open-ended learners can be unclear because, despite the inability to incorporate new elements into their songs after sexual maturity, closed-ended learners have been shown to be able to alter some aspects of their song over time and because the capacity for vocal learning can wane over time even in open-ended learners (Brainard and Doupe 2013; Kipper and Kiefer 2010; Sakata and Woolley 2020).

#### 2.5. Sensory Learning and Sensorimotor Learning

In the course of song learning, accurate perception of the features of tutors’ song is a precondition for learners matching the features of that song in eventual production; thus, song learning (similar to human vocal learning) involves acquiring a sensory representation of a “target song” (sensory learning: Brainard and Doupe 2013; Doupe and Kuhl 1999; Sakata and Woolley 2020; Woolley and Woolley 2020). In addition, the development of an accurate imitation of a tutor’s song requires learning the motor commands to generate that target song (sensorimotor learning). Under sensorimotor learning, birds compare the performance of their current song to the sensory representation of the target song and gradually adjust the motor commands for song production to minimize the deviations between the bird’s current song and the target song (Doya and Sejnowski 1995; Fee and Goldberg 2011; Murphy et al. 2017, 2020; Sakata and Yazaki-Sugiyama 2020). The comparisons between current and target songs appear to be conducted online (i.e., in real-time), and neural activity in various parts of the songbird brain have been found to be sensitive

to perturbations of auditory feedback, in much the same way that areas in the human brain are sensitive to auditory feedback (Keller and Hahnloser 2009; Sakata and Brainard 2006, 2008; Tschida and Mooney 2012). However, it is worth noting that some studies highlight the importance of offline processes (e.g., during sleep) for vocal learning (reviewed in Margoliash and Schmidt 2010).

### 2.6. Concluding Remarks

In summary, vocal learning in songbirds (i.e., vocal production learning) entails the memorization of a tutor's song(s) (sensory learning) followed by vocal practice to learn how to produce the memorized sounds (sensorimotor learning). In this respect, vocal learning in songbirds resembles speech acquisition in humans (as well as many other forms of sensorimotor learning). Different species of birds engage in vocal learning during a restricted period in development or for more extended periods throughout their lives, including in adulthood. Our subsequent sections focus on songbird species that demonstrate song learning both early and later in development (Figure 2B–D) but also highlight experiments in species with more limited periods of vocal learning to emphasize particular aspects relevant to L2 speech acquisition.

## 3. Comparing Late Birdsong Learning and L2 Speech Learning

Numerous reviews discuss parallels between human speech and birdsong learning (e.g., Bolhuis et al. 2010; Brainard and Doupe 2013; Doupe and Kuhl 1999; Gervain and Mehler 2010; Jarvis 2019; Marler 1997; Sakata and Woolley 2020). These reviews have benefited from a rich body of literature in comparing the development of birdsong in juvenile songbirds to speech development in toddlers and children. However, whereas speech acquisition in adult humans has been extensively studied, the amount of experimental research in birdsong learning in later life stages (S2 learning) is relatively limited. This lack of research limits the degree to which insightful parallels and contrasts to L2 speech acquisition can be drawn. Moreover, as we will see and have alluded to above, there are challenges in interpreting data in adult learning among both birds and humans.

The challenges to comparison as well as to interpretation often relate to the level at which observations are addressed and at which explanations are attempted. In particular, to an extent that diverges from many studies of S2 learning in birds, the emphasis in the L2 pronunciation literature is on individual differences, which may be conditioned by idiosyncratic learner-external and learner-internal factors, and which have effects (singly, cumulatively or interactively) of greater or lesser magnitude depending upon the individual.

The external factors may involve, for example, experience (e.g., musical training, duration of L2 immersion, prior training or experience with other non-native languages, pronunciation training, and perceptual training), level of prior development and entrenchment of the L1, and proximity of the L1 and L2 sound systems. The internal factors include, for example, psychoacoustic or sensory abilities (e.g., general hearing acuity, acuity in perceiving pitch (height and excursions), and amplitude and overtones), cognition (e.g., working memory, ability to suppress competing information, and ability to extract signal from noise), singing and music processing ability, musical expertise, sensorimotor skills (e.g., ability to imitate sounds with fine-grained neuro-muscular coordination of the vocal apparatus, and ability to compare self-produced and target pronunciations), motivation to identify with or be taken for a native speaker of the L2 (including will or choice to achieve pronunciation accuracy<sup>8</sup>), and self-regulation of learning (for overviews as well as specific findings, see Birdsong 2007, 2018; Bongaerts 1999; Coumel et al. 2019; Guion et al. 2000; Hu et al. 2012; Ingvalson et al. 2017; Jenkins and Setter 2005; Mora et al. 2013; Moyer 2004, 2014; Piske et al. 2001; Reiterer et al. 2011, 2013; Rota and Reiterer 2009; Slevc and Miyake 2006; Wong et al. 2012; Yeni-Komshian et al. 2000).

Further, in certain ways that differ from S2 learning (see Sections 3.3 and 4.1), neurological maturation, which would presumably circumscribe a critical period for learning, does not map neatly onto observed age-related effects in L2 vocal learning. Typically,

the age at which L2 learning begins in earnest (Age of Acquisition or AoA) negatively correlates with accuracy of L2 accent. However, the overall AoA–accent relationship is roughly linear, i.e., not suggestive of a period during which a monolingual-like target accent can be attained. Rather, and crucially, because in bilingualism the L1 and the L2 *affect one another* at all levels of language production and processing, from-birth simultaneous bilinguals under close examination do not display monolingual-like accents in *either their first or second language*. In this respect, departures from the presumed target of L2 speech learning cannot be ascribed to neurological features of biological age (see [Birdsong 2018](#)). Moreover, according to [Flege and Bohn \(2021, p. 64\)](#), under the revised Speech Learning Model (SLM-r), the ability to form new phonetic categories does not entirely disappear with aging. Rather, “phonetic category formation is possible regardless of age of first exposure to an L2 and is crucial for phonetic organization and reorganization across the life-span.”

The preceding observations converge on a picture of L2 speech learning whereby: (1) the potential for acquiring (or at least accurately imitating) L2 accent is maintained to some degree over time; (2) the range of possible outcomes is conditioned by a suite of learner-internal and learner-external factors (some of which are not at play in S2 learning in songbirds); and (3) the facts of L2 speech learning are best understood in terms of within-species individual differences.

As we will elaborate in later parts of this article, because the features of L2 speech learning in (2) align only partially with what is known (or knowable) about S2 learning in songbirds, we are limited in the extent to which we can directly compare humans with birds. Nevertheless, in this section we attempt to contextualize and synthesize studies of S2 learning in songbirds, and to compare and contrast factors that modulate the extent of S2 learning and L2 speech acquisition in humans. We also argue that, to the extent that processes regulating L1 speech and S1 acquisition also modulate L2 speech and S2 learning, it is informative to draw parallels between L1 speech and S1 acquisition to understand mechanisms underlying variability in L2 speech and S2 learning. In addition, in [Section 4.1](#) we go on to more closely examine three conditioning factors in L2 speech learning—the age at which L2 learning is begun, L1 entrenchment, and interactions with speakers of the L2—which lend themselves to meaningful comparisons with S2 learning.

### 3.1. Repertoire Size, Timing of Learning, and Variability

As stated above, we conceptualize L2 speech acquisition in humans as an increase in the repertoire of sounds that individuals can produce and use for communication. In this respect, songbirds that increase or alter their vocal repertoires in adulthood are most informative in our examination of parallels with humans.<sup>9</sup> To this end, we highlight studies describing experience-dependent repertoire changes in birds that have acquired a mature set of vocalizations and relate variation in this type of learning to variation in L2 speech acquisition in individuals who are already proficient at their native language (“sequential bilingualism”).

A central focus of bilingualism research is revealing the factors that contribute to variation in L2 speech acquisition (see [Section 4](#)). Whereas, by definition, studies of human speech acquisition focus on within-species (i.e., between-individual) variation in learning, much of the research on adult birdsong learning has examined between-species variation in the extent of adult song learning or in repertoire size (but see below for related discussions of individual variability in learning; [Beecher and Brenowitz 2005](#); [Brainard and Doupe 2002](#); [Brenowitz and Beecher 2005](#); [Robinson et al. 2019](#); [Sakata and Woolley 2020](#)). The questions most frequently posed relate to how and why some songbird species can acquire novel vocalizations as adults (e.g., European starlings and nightingales) whereas others cannot (e.g., zebra finches and white-crowned sparrows). With regard to “how,” a number of neurobiological factors have been proposed to relate to this species variation (see discussion below), and with regard to “why,” scientists have speculated on the importance of sound diversity to increase the attractiveness of an individual’s song (because song is often used in sexual contexts: [Beecher and Brenowitz 2005](#); [Podos and Sung 2020](#); [Robinson and Creanza](#)

2019; Robinson et al. 2019). A related question posed by biologists is how repertoire size is related to the duration of learning. Recent broad-scale surveys of birdsong reveal that, unsurprisingly, species in which individual birds produce larger vocal repertoires (e.g., greater diversity of syllables, phrases, and song types, etc.) tend to have longer periods of vocal learning, including learning periods that extend into adulthood (reviewed in Robinson et al. 2019).

However, it is important to highlight challenges in the interpretation of findings about vocal change in adulthood as well as differences in the nature and depth of studies into late learning in songbirds. First, vocal change in adulthood does not necessarily require adult vocal learning. A number of longitudinal studies document changes to the range and types of sounds produced by adult songbirds over time, but it is not clear that the production of a “novel” vocalization (i.e., a vocalization that an individual bird is observed to produce at a later time but not observed at an earlier time point) reflects the acquisition of a new vocalization (Brenowitz and Beecher 2005; Marler 1997; Robinson et al. 2019). For one thing, it is possible that “novel” vocalizations represent vocalizations learned during development but only expressed later in life (e.g., perhaps because of a change in the social environment: Geberzahn et al. 2002; Marler 1997). Second, it is possible that “novel” vocalizations represent vocal plasticity and not vocal learning. A novel song type, for example, can emerge from an individual recombining syllables or phrases in different ways; this has been termed vocal “innovation” by numerous researchers and is distinct from vocal learning per se (i.e., imitating a novel vocalization; Catchpole and Slater 2008) (This is a challenge for studies of animal communication but not necessarily for L1–L2 speech acquisition: because linguists, anthropologists and psychologists have parameterized the boundaries of languages, the production of a unique word in a different language (and in an appropriate semantic context) is almost always a product of learning (vs. random vocal exploration or innovation)). In addition, the acoustic properties of one’s vocalizations can also change due to changes to vocal muscles, body size, and brain physiology instead of as a consequence of learning. Third, from a methodological standpoint, it is also possible that “novel” vocalizations were in fact produced by focal individuals at an earlier time point but that researchers had not recorded the bird producing those vocalizations at an earlier time point.<sup>10</sup> Thus, as is the case with humans, quantifying the full repertoire of individual birds is challenging, especially for species with relatively large vocal repertoires, so it can be very difficult to determine the novelty of vocalizations.

Finally, it is tempting to compare impressive feats of mimicry (i.e., the ability to imitate the sounds of other species) to “successful” or “target-like” L2 speech acquisition in humans; indeed, various popular press articles draw parallels between mimicry in birds and multilingualism; see Section 5. However, the same cautions noted above about adult song learning should be applied to mimicry, along with an additional caveat. Acoustic similarities across species can arise in the absence of learning because many innate, unlearned vocalizations (e.g., various calls) are acoustically similar across species (“convergence”; Dalziell et al. 2015; Kelley et al. 2008). Consequently, while the acoustic properties and functionalities could be sufficiently similar to qualify vocalizations as mimicry, vocal similarities across species should not be related to L2 speech acquisition (or any other form of vocal learning) without demonstration of learning (i.e., the extent to which experiences lead to vocal similarities between mimics and heterospecifics). In cases in which mimicry involves learning (e.g., imitation of anthropogenic sounds), the timing of learning generally remains unclear from natural observations. Given that the species composition of habitats occupied by avian mimics can be stable across the lifespan, one cannot differentiate imitations learned during development or as adults without proper experiments. Taken together, while existing and ongoing field studies of vocal change and plasticity in songbirds expand and deepen our understanding of vocal plasticity and can help identify bird species to model L2 speech acquisition, various factors need to be taken into consideration before making conclusions about human–animal parallels in adult vocal learning.

Despite the challenges of effectively demonstrating adult vocal learning in songbirds, some researchers have managed to raise some songbird species from an early age in the lab, regulate (to the extent possible) the sensory environment of birds throughout development, and demonstrate the ability to acquire novel vocalizations as adults (Figure 2B,C). For example, after experimentally tutoring European starlings with distinct sets of song types at various time points in development and adulthood, it was observed that adults learned to produce not only songs heard during development but also some song types that they were exposed to after sexual maturation (Chaiken et al. 1994). Similarly, experimentally tutoring common nightingales with distinct song types during development and again around the time of sexual maturation (~9–12 months of age) led to nightingales imitating song types that they heard during development as well as song types that they only heard around the time of sexual maturation (Todt and Geberzahn 2003). Interestingly, song types learned around the time of maturation were produced in the nightingale's second but not first year of life, thus highlighting the complexities of documenting adult learning among songbirds (see also related data on delayed expression of learning in European starlings (Chaiken et al. 1994) and brown-headed cowbirds (O'Loghlen and Rothstein 2010<sup>11</sup>). In another experiment, adult pied flycatchers were exposed to unfamiliar syllable types (i.e., syllable types that were determined not to be produced by birds in the study area, based on ~15 years of song surveys) for 4 h a day for a week, and birds demonstrated some (albeit limited) song learning (Eriksen et al. 2011); in particular, of the 20 birds studied, one 1-year old male and two older males (>2 yrs old) imitated one of the novel syllables. Additional studies reported song learning in adult canaries (1–2 yrs old: Lehongre et al. 2009), indigo buntings (<1–2 yrs old; Margoliash et al. 1994) and cardinals (9–12 months old; Yamaguchi 2001); however, the lack of information about the distinctiveness of songs heard early vs. later in life prevents firm conclusions about adult vocal learning in these birds.<sup>12</sup>

In some instances, experimental tutoring has not supported previous conclusions about adult vocal learning. For example, northern mockingbirds (Figure 2D) have been classically considered as capable of adult vocal learning and mimicry because their vocal repertoire increases with age (Derrickson 1987). However, a longitudinal field study that exposed adult mockingbirds to unfamiliar heterospecific songs (i.e., songs from species that lived hundreds of kilometers away and were not found in the study area) for 6–7 months failed to find evidence of adult vocal learning (Gammon 2020). Although various factors could have contributed to this lack of learning (e.g., no social interactions with unfamiliar heterospecifics, and the ages of study populations were unknown), these data raise the possibility that northern mockingbirds might not be capable of adult vocal learning.

Although we argue that studies of S2 learning in adult songbirds represent the most compelling parallels with sequential L2 speech acquisition, studies of sequential song tutoring in developing songbirds could also resemble L2 speech acquisition during development in humans (e.g., L2 learning in children (e.g., 6–10 yrs old)). Sequential tutoring paradigms in closed-ended learners such as the zebra finch (reviewed in Gobes et al. 2019) highlight tutoring-dependent changes in repertoire composition, including increases in repertoire size, or in the temporal organization of song (e.g., Chen and Sakata 2021; Eales 1985; Gobes et al. 2019; Lipkind et al. 2013, 2017; Olson et al. 2016; Yazaki-Sugiyama and Mooney 2004). For example, a number of studies have found that juvenile zebra finches will incorporate syllables from multiple tutors in response to sequential tutoring, with the magnitude of intermixing of tutors' syllables depending on the timing and nature of sequential tutoring (reviewed in Gobes et al. 2019). These behavioral studies serve as a foundation for exploring the brain mechanisms underlying S2 learning.

### 3.2. Neural Representation of Speech and Song

Within the context of well-known individual-level differences in L2 neural representation, considerable neuroimaging evidence suggests that individuals (including late L2 learners) who acquire L2 grammar to high levels of proficiency and who are proficient in L2 lexico-semantic processing employ the same neural structures that are responsible

for L1 learning and processing (e.g., [Abutalebi 2008](#); [Klein et al. 1995](#); [Ripollés et al. 2016](#); [Steinhauer et al. 2009](#); [Tagarelli et al. 2019](#)). In the specific domain of L2 speech, a similar association is found. Although neural representations of L2 speech learning can vary across individuals (e.g., [Díaz et al. 2008, 2016](#); [Golestani and Zatorre 2004](#); [Golestani et al. 2007](#)), those who succeed in L2 speech learning have been shown to display emergent neural representations that are similar to those involved in primary (native) speech learning in early infancy, particularly in the superior temporal gyrus and right precentral gyrus ([Feng et al. 2021](#); see also [Díaz et al. 2008](#)).

To elaborate, [Golestani and Zatorre \(2004\)](#) examined L2 learning of the Hindi-like dental-retroflex contrast for occlusives /t/ and /d/ in adult English natives and found that successful learning was associated with recruitment of the same brain areas that are involved in processing native phonetic contrasts, specifically the left superior temporal gyrus, the insula-frontal operculum and the inferior frontal gyrus (In previous work, [Golestani et al. \(2002\)](#) report that adult English native speakers who were successful in learning the dental-retroflex contrast exhibited more white matter than gray matter in the parietal lobes). In a training study on the imitation of the novel vowel /y/ (the rounded counterpart of high-front rounded /i/) in nonce words by native English-speaking adults, [Carey et al. \(2017\)](#) found that imitation improvement was associated with increased activation in the insular cortex, pre-supplementary motor area (pre-SMA), cerebellum, cingulate gyrus and sulcus, whereas worsening performance was associated with reduced activation in these same brain regions. Many of these areas, particularly within the left hemisphere, are involved in the perception and production of features of speech by adults at higher levels of L2 proficiency ([Minagawa-Kawai et al. 2011](#)).

In conceptualizing parallels in the neural representations of L2 speech with S2 vocalizations, we believe that the most apposite point of comparison is with the neural correlates of vocal repertoire sizes in adult songbirds (Section 3.1). Accordingly, we examine the relationship between the size of focal areas in the songbird brain and measures of repertoire size. For example, the size of a sensorimotor brain area called HVC (acronym used as the proper name) is positively related to the number of syllable or song types in a bird's song (reviewed in [DeVoogd 2004](#)). This relationship has been observed when examining between-individual variation within songbird species such as zebra finches ([Airey and DeVoogd 2000](#); [Moore et al. 2011](#)), marsh wrens ([Airey et al. 2000](#)), and sedge warblers ([Pfaff et al. 2007](#)), and is also characteristic of between-species variation in syllable repertoire sizes (e.g., [Moore et al. 2011](#); [Székely et al. 1996](#)). These findings have led to the proposal that more HVC neurons are required to encode and produce a greater diversity of sounds.<sup>13</sup> HVC has been demonstrated to be important for S1 learning (reviewed in [Sakata and Yazaki-Sugiyama 2020](#)) but its contribution to S2 learning has not been examined.

It is important to emphasize that brain organization and response not only reflect learned sounds that individuals can produce but can also reflect learned vocalizations that are not produced by individuals. In the literature on human language learning, studies of international adoptees who are adopted into a different language environment generally point to a neural maintenance of phonetic features of the first-learned language (e.g., [Au et al. 2002](#); [Choi et al. 2017](#); [Oh et al. 2010](#); [Pallier et al. 2003](#); [Pierce et al. 2014](#); [Ventureyra et al. 2004](#)). For example, [Pierce et al. \(2014\)](#) showed that native Chinese children exposed exclusively to French after adoption (N = 23; mean age of adoption 12.8 mo.), and who had no use or conscious recollection of their L1, maintained neural representations in the left temporal lobe from the posterior to anterior superior temporal gyrus and in the left planum temporale when processing lexical tone—a feature of Chinese that is absent in French. Importantly, the representations for tone processing exhibited by the adoptees were identical to those of a group of Chinese-French bilinguals (N = 12) who had started learning French at avg. 16.9 months.

Similarly, neural correlates of unproduced (or at least unobserved) vocalizations have been documented in the songbird brain. A number of songbird species, including various species of sparrows and nightingales, learn multiple and different song types during

development but only produce a subset of these song types as adults. From a behavioral perspective, it appears that memory traces of various song types persist, as some learned songs not produced during the first year have been found to be produced in subsequent years (Geberzahn and Hultsch 2003; Hough et al. 2000; Nelson and Marler 1994). From a neurobiological perspective, neurons in sensorimotor and auditory processing areas of the songbird brain have been found to encode representations of multiple song types that are or are not produced by the bird. For example, Prather et al. (2010) demonstrated that HVC neurons of adult swamp sparrows were activated by most of the song types that birds learned during development, regardless of whether the bird produced that song in adulthood or not. The caudomedial nidopallium (NCM) is a part of the songbird auditory system that has been argued to represent the associative auditory cortex in mammals as well as Wernicke's area in humans and found to be important for sensory learning in the service of vocal learning (e.g., Bolhuis et al. 2010; Chen and Sakata 2021; Gobes and Bolhuis 2007; London and Clayton 2008; Olson et al. 2016; Woolley and Woolley 2020). Different neurons in the NCM respond preferentially to different songs heard by individual birds; for example, some neurons in the NCM are activated most when a zebra finch hears its tutor's song (the song that it imitated) whereas other neurons fire most when a bird hears the song of familiar bird that it does not imitate (e.g., Yanagihara and Yazaki-Sugiyama 2016). Collectively, these data suggest that the songbird brain retains a sensory representation of song types that were learned early in development, regardless of whether or not birds actively produce those song types.

The lateralization of representations of learned vocalizations has been actively studied in the L1 and L2 literature. In general, among monolinguals, the brain regions activated during speech perception become more left-lateralized with increasing language proficiency (Dehaene-Lambertz et al. 2002; but see Friederici 2011 for evidence of right-hemisphere involvement in the processing of prosody). In L2, brain activation in response to speech perception is highly variable across individuals, and tends to be right-dominant or bilateral among those who start L2 learning after 7 years of age and reach only moderate levels of proficiency in that language (Dehaene et al. 1997). As mentioned above, at higher levels of L2 proficiency, both production and perception in the L2 involve activity in the same left-hemisphere areas that are associated with the processing of native-language phonetic features and contrasts (Minagawa-Kawai et al. 2011; see also Carey et al. 2017; Feng et al. 2021; Garcia-Sierra et al. 2011; Golestani and Zatorre 2004).

Interestingly, variation in the lateralization of neural activity in response to song perception has been found to relate to proficiency of vocal performance or the degree of song imitation in zebra finches (Gobes et al. 2019; Moorman et al. 2012). Specifically, among zebra finches that are sequentially tutored with two different songs (Tutor 1 vs. Tutor 2), song playback-evoked brain activity in the auditory area NCM is left lateralized in birds that produce a more accurate imitation of the first tutor's song but right lateralized in birds that produce a more accurate imitation of the second tutor's song (Olson et al. 2016).

### 3.3. Mechanisms Underlying Age-Dependent Changes in Vocal Learning Abilities

Birdsong (2018) discusses how developmental neurobiology and experience interact to contribute to variable timing and outcomes in L1 and L2 speech learning (see also Werker and Hensch 2015). For example, compared to monolingual acquisition, the period of sensitivity for discrimination of speech sounds (which is regulated by various biological processes: Werker and Hensch 2015) is extended by simultaneous and early-sequential bilingual acquisition (e.g., Petitto et al. 2012), and the production of fine-grained acoustic features of L1 and L2 speech may vary depending on exposure to speakers who have accents in one or both languages and on the amount of exposure to, and use of, each language (e.g., Mack 1989).

Among animals, various types of behaviors and cognitive processes appear to be regulated by neural processes that change over age. For example, the developmental emergence of "molecular brakes" over time (i.e., cellular products that limit plasticity) in various

sensory systems of the mammalian brain have been proposed to constrain experience-dependent plasticity in sensory processing (reviewed in [Takesian and Hensch 2013](#)). Importantly, the emergence of these “brakes” in sensorimotor brain areas interacts with experiential variables to regulate the closure of sensitive periods for vocal learning in songbirds ([Balmer et al. 2009](#); [Cornez et al. 2018](#); reviewed in [Takesian and Hensch 2013](#)). Similar phenomena have been proposed to occur in the human brain ([Werker and Hensch 2015](#)).

Although a suite of molecules and gene products (e.g., *lynx1*, brain-derived neurotrophic factor (BDNF), myelin and myelin-associated inhibitors) have been found to affect plasticity in brain circuitry and proposed to regulate the timing of critical periods (reviewed in [Reichelt et al. 2019](#); [Takesian and Hensch 2013](#)), the molecular brake most extensively studied in the context of vocal learning in songbirds is the perineuronal net (PNN). Perineuronal nets are lattice-like structures of macromolecules that form around neurons when the critical periods close and are hypothesized to constrain the plasticity of neurons they ensheath (Takesian and Hensch 2013; Reichelt et al. 2019; Wang and Fawcett 2012). For example, the organization of the visual system is shaped by visual experiences during a critical period in development, and PNNs emerge in the primary visual cortex as the critical period for visual plasticity closes. Moreover, experimentally degrading PNNs after the closure of the critical period reinstates experience-dependent plasticity of visual processing (e.g., [Beurdeley et al. 2012](#); [Pizzorusso et al. 2002](#)).

In songbirds that are considered closed-ended learners (e.g., zebra finches), PNNs are found throughout brain circuits regulating the production and acquisition of birdsong ([Balmer et al. 2009](#); [Cornez et al. 2017, 2018](#)). Moreover, PNN expression in these areas changes over development, with PNNs increasing in relevant brain areas (e.g., HVC) as the critical period for song learning closes. Furthermore, species variation in PNN expression in the adult songbird brain inversely covaries with species variation in adult song learning: PNN expression is reduced in species that demonstrate more extensive adult vocal plasticity (European starlings) compared to species that do not learn songs as adults (zebra finches; [Cornez et al. 2017](#)). These findings suggest that the expression of PNNs in key brain areas could constrain the acquisition of new vocalizations in adulthood. Although PNNs are found in cortical areas for speech and language ([Werker and Hensch 2015](#)), little is known about how their emergence in relevant brain circuits could shape L2 speech acquisition.

As indicated above, vocal learning and imitation require the transformation of an acquired sensory representation into a behavioral output (vocal performance) through sensorimotor learning. In this respect, it is important to understand how age-dependent changes in vocal learning are due to age-dependent changes in sensory learning, sensorimotor learning, and the transformation of sensory learning into sensorimotor learning. For example, although humans are able to associate new speech sounds with new words (and their meanings) throughout their lifespan (e.g., [Gaskell and Ellis 2009](#); [Park et al. 2001](#); [Singleton 1995](#); see also [Yeung and Werker 2009](#) for the relationship between nonnative phonetic distinctions and word learning), humans exhibit a general decline in pronunciation accuracy with increasing age (see Section 4.1).

Age-dependent changes in sensory and sensorimotor learning have been most extensively addressed in the zebra finch, a closed-ended learner. Overall, the capacity for sensory learning and subtle forms of sensorimotor learning are retained even well past the critical period for song learning. For instance, adult songbirds can memorize the many different sounds of other birds that they hear in adulthood (reviewed in [Dai et al. 2018](#); [Elie and Theunissen 2020](#); [Woolley and Woolley 2020](#); [Yu et al. 2020](#)). In addition, manipulations of sensory feedback can drive adaptive changes to the acoustic structure, timing, and sequencing in adult songbirds ([Tumer and Brainard 2007](#); [Brainard and Doupe 2013](#); [Sakata and Yazaki-Sugiyama 2020](#)). Consequently, sensory and sensorimotor learning appear to both be intact in adult songbirds that cannot acquire new songs as adults. In this respect, it is hypothesized that the inability for closed-ended birds to imitate novel sounds after the closure of the critical period reflects an inability to “translate” sensory learning of new sounds into sensorimotor learning of new sounds (This is analogous to adult humans being

able to understand elements of a novel language but limited in their ability to produce extensively). Discerning the neural and genetic mechanisms that regulate age-dependent changes in the ability to translate sensory learning into sensorimotor learning is an active area of research in songbirds.

### 3.4. Concluding Remarks

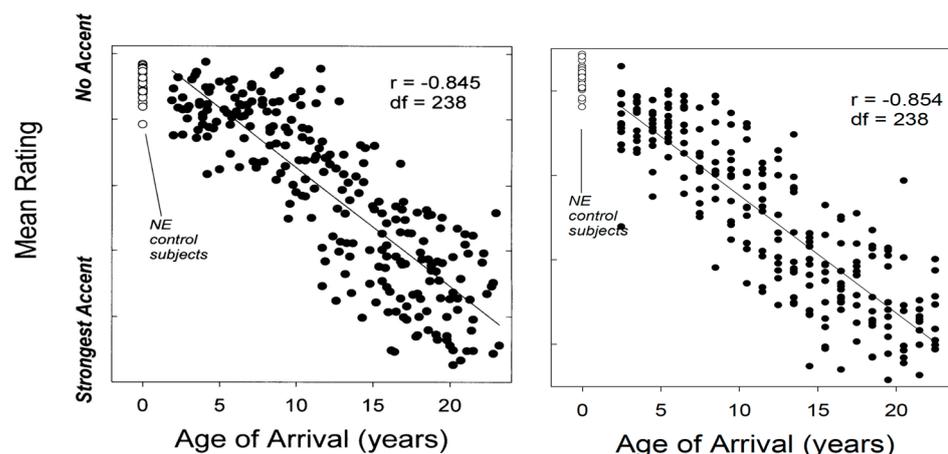
In summary, some studies of S2 learning in songbirds provide useful parallels to L2 speech acquisition in humans. Little is known about the degree to which S1 and S2 learning recruits overlapping or distinct neural populations, but further research on sequential song learning in more extensively studied songbirds (e.g., zebra finches) could be directed toward establishing similarities (and differences) across species.

## 4. Factors in Variable Outcomes

It is well known that individuals vary in the degree to which their L2 speech—be it measured on fine-grained acoustic features or on strings of connected words—resembles that of native monolingual speakers. As mentioned at the beginning of Section 3, there are many factors that condition individual-level accentedness, including: the amount of vocal, motor, and perceptual training; age at the initial state of L2 learning; the amount and quality of input and interaction in the L2 vs. L1; the degree of L1 vs. L2 dominance (which is relatable to the degree of L1 entrenchment, L1 attrition and L1–L2 similarity, see Köpke 2021); genetic markers; cognitive style; motivation; education; and perceptual acuity; etc. (e.g., Birdsong 2007, 2018; Bongaerts 1999; Guion et al. 2000; Ingvalson et al. 2017; Jenkins and Setter 2005; Moyer 2014; Wong et al. 2012; chapters in Watkins et al. 2009; Yeni-Komshian et al. 2000). Of these factors, three are most apposite for consideration alongside S2 learning in birds: age at which L2 learning begins, degree of entrenchment of the L1, and interactions with L2 speakers.

### 4.1. Age of Learning, Accentedness and Variability

The age at which L2 learning begins, often referred to as Age of Acquisition (AoA), is generally associated with immersion in the L2 and with the beginning of frequent use of the L2. As previously noted, later AoAs are generally associated with greater accentedness (i.e., less native-like pronunciation; Figure 3). In addition, the variability or range of accent scores is larger among later learners than among earlier learners. This AoA-related variability is also illustrated in Figure 3, where we observe modest dispersion of L2 English accent scores among groups of Italian natives (left) and Korean natives (right) with AoA < 6–8 y, and considerably greater dispersion among participants with later AoA.<sup>14</sup>



**Figure 3.** The relationship between AoA and L2 accentedness. Left image: L2 English accent ratings for 240 Italian native speakers (dark circles), and ratings for 24 native English controls (open circles); all participants were residents of Ottawa (adapted by Jim Flege from Flege 1999, fig. 5.1; data originally from Flege et al. 1995, republished with permission from Taylor & Francis Group). Right image: L2 English accent ratings for 240 Korean native speakers (dark circles), and accent ratings for 24 native English-speaking controls (open circles); Korean participants resided in the Washington, DC, area and English natives were born in the US and lived in Mid-Atlantic states (adapted by Jim Flege from Yeni-Komshian et al. 2000, fig. 1, republished with permission from Cambridge University Press). In both instances, AoA > ~6–10 y is associated with a stronger (less English-like) accent and with greater dispersion of accent ratings.

AoA is an individual variable that is often predictive of outcomes of L2 learning, which some researchers have attributed to the maturational state of neurobiological mechanisms involved in L2 learning (for overviews, see Birdsong 1999, 2017, 2018; Hyltenstam and Abrahamsson 2003). However, as many researchers have pointed out (e.g., Birdsong 2009, 2018; Flege 1998, 2018; Hartshorne et al. 2018), the meta-variable of AoA subsumes, or directly or indirectly conditions, several individual variables that are unrelated to nervous system maturation. These variables include L1 entrenchment, amount and quality of L2/L1 input and interaction, education, attitudes, goal orientation, and domain-general cognitive styles, etc. We note as well that, due to scaling effects, whereby the likelihood for range differences in at least some of these domains increases with age (and thus with the corresponding multiplication of experiences, diversification of learning goals, and modulation of attitudes), greater variability in L2 outcomes (in all domains, including pronunciation) can be expected in later AoA (Birdsong 2018). In a similar manner, the age at which individual songbirds are exposed to different songs is fundamentally related to the degree to which individuals can learn to produce accurate imitations of those songs. For closed-ended learners, the age of exposure to song (S1) constrains the degree of vocal imitation (see above for discussion of sensitive or critical periods in birdsong learning). Similarly, song learning can be less robust after the first or second year of life in songbird species that demonstrate adult vocal learning (i.e., learning S2 subsequent to learning an S1; Chaiken et al. 1994; Robinson et al. 2019). Thus, the research indicates that declines in production accuracy of S1 and S2 songs are related to the age of learning. However, songbird research has generally not addressed the possibility (and potential sources) of age-related variability in S2 outcomes; see the following section.

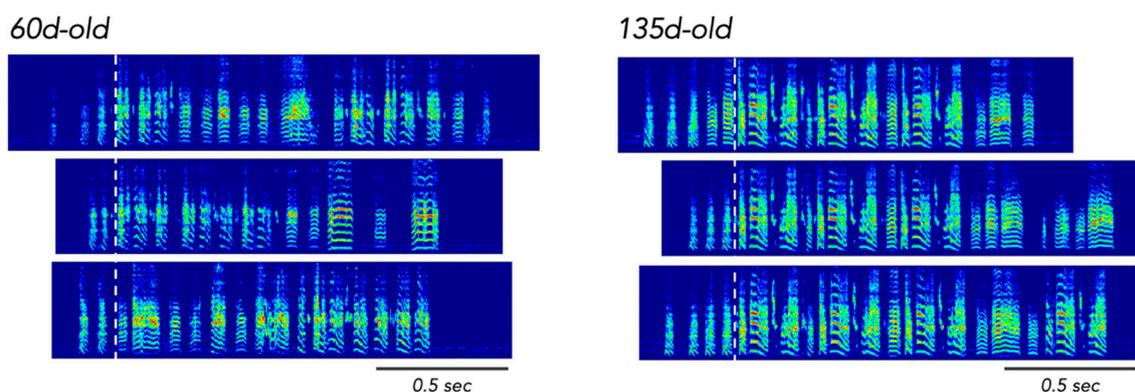
#### 4.2. The Role of Entrenchment

One of the factors that is associated with inter-individual variability in L2 speech acquisition is the degree of L1 entrenchment (e.g., Birdsong 2018; Flege 1999; Flege and Bohn 2021; MacWhinney 2005; Marchman 1993; Simmonds 2015). On this view, individuals

with highly-entrenched L1 speech sound repertoires are less likely to produce L2 speech with an accent resembling that of monolingual native speakers of that language.

In attempting to find animal parallels with this phenomenon, we first conceptualize entrenchment in a manner that permits comparison. In humans, L1 entrenchment is understood in reference to the developmental state of the L1 (i.e., the degree to which it is a mature or “crystallized” system), and in terms of the repeated exposure and cumulative use, which reinforces the linguistic system in representational, neurological and behavioral (processing and production) terms (e.g., Theakston 2017; Tomasello 2005; Zhang and Mai 2018). Because the expression of learned behaviors (including speech) becomes more stereotypic and consistent as one repeats and masters the behavior (e.g., MacWhinney 2005; reviewed in Dhawale et al. 2017), one way to frame L1 entrenchment that allows for cross-species comparison could be in terms of the stereotypy of L1 speech production (which would correspond to consistency in articulation across renditions). In other words, individuals who produce the compositional units (phonetic segments, syllables, and lexical collocations) of the L1 with more consistency are those whose L1 is more entrenched.

Given this definition of entrenchment, interesting parallels can be observed in songbirds. For example, adult zebra finches (a closed-ended learner species) produce a single song type that is learned from a tutor during development, and as a juvenile zebra finch gradually masters the imitation of his tutor’s song, his song becomes increasingly stereotyped Figure 4 (Tchernichovski et al. 2001; reviewed in Dhawale et al. 2017; Sakata and Woolley 2020). Moreover, in parallel with this developmental change in vocal stereotypy, one observes a decrease in the ability of zebra finches to acquire novel vocalizations. Previous studies have documented that some young adult zebra finches alter their vocalizations when exposed to conspecifics that sing a different song (e.g., when placed in an aviary with other birds that produce distinct songs) and that the degree of vocal plasticity is inversely related to the stereotypy of the young adult’s song at the time they are placed in the aviary. In other words, compared to birds with more acoustically variable songs, young adult zebra finches that produce more acoustically stereotyped songs (i.e., the production of their song is consistent from rendition to rendition) demonstrate smaller changes to their song in response to cohabitation with other birds (Derégnaucourt et al. 2013; Eales 1985; Jones et al. 1996; but see Morrison and Nottebohm 1993).



**Figure 4.** Developmental changes in the stereotypy with which song bouts are produced. On the left are spectrograms of three song bouts of a developing zebra finch when he is 60 days old, and on the right are spectrograms of three song bouts from the same bird when he is 135 days old. All song bouts are aligned by the first syllable of the song (white, dashed vertical line; i.e., after introductory elements of the song bout). The syllable composition and sequencing of syllables within the bout are much more consistent across renditions when the bird is 135 days old.

Although factors unrelated to the biology of vocal stereotypy could contribute to this individual variability in learning (e.g., birds that produce more stereotyped S1 could interact less or differently with S2 tutors), these data suggest that neural “commitment” to S1 could constrain the ability of birds to acquire novel elements from an S2 and that

the stabilization of the vocal motor program is intimately related to changes to plasticity mechanisms in the brain (see above for discussion of brain mechanisms). For instance, as a bird's song becomes more stereotyped, the songbird brain becomes more "tuned" to the bird's own song (i.e., more neurons are preferentially activated by hearing the sound of the bird's own song compared to hearing the songs of other birds: e.g., [Doupe 1997](#); [Nick and Konishi 2005a, 2005b](#)), and the rate of neurogenesis (argued as a metric of brain plasticity) decreases ([Pytte et al. 2007](#)).

As mentioned above, studies of species with vocal repertoires (i.e., species that produce multiple song types) are useful in comparisons with L2 speech acquisition. Unfortunately, the relationship between vocal stereotypy and the capacity for adult vocal learning has not been examined in the experiments previously described. Given how L1 entrenchment influences L2 speech acquisition in humans (and late song learning in closed-ended learners), it would be informative to investigate how individual variation in vocal stereotypy at the time of S2 tutoring predicts the extent of adult song learning in response to tutoring (see [Section 3.1](#)).

Finally, although we discuss stereotypy at a broad scale, the emergence of acoustic stereotypy is regulated at the local (fine-scale) level and related to local variation in plasticity. For example, within individual songbirds, different syllables can become stereotyped at slower or faster rates ([Ravbar et al. 2012](#); [Vallentin et al. 2016](#)), and local variation in the stereotypy of acoustic structure, sequencing, or timing of syllables within a bird's song inversely relates to local variation in experimentally-driven plasticity (e.g., [Tachibana et al. 2017](#); [Vallentin et al. 2016](#); [Warren et al. 2012](#)). A similar set of conditions is associated with L2 speech learning. According to [Flege and Bohn \(2021, p. 66\)](#), L1 entrenchment, L2 input, perception-based production tuning, and feedback looping at a local level could correlate with L2 speech acquisition at this same local level: "Individuals differ in terms of how accurately they produce and perceive L2 sounds. By hypothesis, inter-subject phonetic variability can be explained, at least in part, by knowing how individual learners' L1 phonetic categories were specified when they were first exposed to an L2, how they perceptually linked L2 sounds to L1 sounds via the mechanism of interlingual identification, how dissimilar they perceived an L2 sound to be from the closest L1 sound, and the quantity and quality of L2 phonetic input they have received." Thus, for example, individual-level variability in the enunciation and production of particular words in their L1 (e.g., "information" spoken in L1 English), along with variations in input, perception and feedback, could predict the degree to which their production of similar words in their L2 is accented (e.g., "information" in L2 French).

#### *4.3. The Importance of Social Interactions for Vocal Learning*

It is to be expected that the amount of L2 use is predictive of the accuracy of L2 pronunciation (e.g., [Flege 2018](#)). Underlying the simple dimension of quantity of L2 use are factors of a conative, instrumental, or identifying nature that further condition the degree of native-like speech production. For example, L2 learners who strive to sound similar to native speakers may purposefully engage and interact socially with natives on a frequent basis, thereby assuring naturalistic input as well as richer opportunities for speaking practice, feedback on pronunciation and fine-tuning of pronunciation; see [Li and Jeong \(2020\)](#), and references therein, for a neurological model of L2 learning based in social interaction.

Social interactions profoundly affect the rate, fidelity, and trajectory of vocal learning not only in humans but also in a variety of non-human animal species, including songbirds ([Kuhl 2007](#); [Ljubičić et al. 2016](#); [Sakata and Yazaki-Sugiyama 2020](#)). Most of the studies dealing with social influences on birdsong learning reveal how social interactions promote the developmental acquisition of song (S1 learning). For example, juvenile birds that are socially tutored produce more accurate imitations of the tutor's song than juveniles that are passively tutored with audio playbacks of a tutor's song, and visual and acoustic interactions with the tutor are thought to be critical for this social potentiation of learning (e.g., [Baptista and Petrinovich 1986](#); [Chaiken et al. 1994](#); [Chen et al. 2016](#); [Houx and Cate](#)

1998; Todt et al. 1979). In addition, visual and acoustic reinforcement signals provided by adults have been found to shape vocal development in juvenile songbirds (Carouso-Peck et al. 2020; West and King 1988).

However, to our knowledge only one study has assessed the role of social interactions on the acquisition of novel vocal elements in songbirds that have previously been tutored (i.e., social influences on S2 learning). Chaiken et al. (1994) either socially tutored adult European starlings with a live tutor or passively tutored age-matched starlings with playbacks of starling songs. They reported learning of novel song types in both groups of tutored starlings and reported a trend for socially tutored birds to imitate a larger repertoire of song types than passively tutored birds. As such, these data are consistent with the social potentiation of vocal learning during development of the S1 (It should be noted that this study lacks appropriate controls for the type and amount of song between socially and passively tutored adults (i.e., socially tutored pupils could have been exposed to more songs from their tutor compared to passively tutored pupils), which limits the degree to which group differences can be attributed to social interactions per se). It is useful to point out that studies that report minimal evidence of adult vocal learning in putatively open-ended learners (e.g., northern mockingbirds: Gammon 2020; pied flycatchers: Eriksen et al. 2011) relied on passive methods of song tutoring, and the authors proposed that social interactions might be required for more substantial adult S2 learning.

#### 4.4. Concluding Remarks

Understanding the factors that similarly regulate variation in L2 speech acquisition and S2 learning is important for revealing the biological foundations of L2 speech acquisition. Although studies of variation in S1 song learning in songbirds abound (e.g., Brainard and Doupe 2002, 2013; Doupe and Kuhl 1999; Gobes et al. 2019; Sakata and Woolley 2020), relatively little is known about processes that regulate S2 song learning, and we propose that vocal entrenchment and social interactivity represent two important lines of inquiry to discern human-songbird parallels. Conversely, given AoA-dependent variation in the attainment of L2 proficiency in humans, age-dependent changes to brain mechanisms in songbirds (see Section 3.3) provide a useful roadmap to understand how AoA modulates L2 speech acquisition.

### 5. In What Ways Are “Bilingual Birds” Comparable to Bilingual Humans?

In this section, we explore how specific dimensions of song production, learning and use can be illuminated by comparison with facets of human bilingualism. To this end, we identify three foundational characteristics of human bilingualism—interference of the L1 in the L2, L1 attrition, and code-switching—and ask whether these characteristics are shared by songbirds. We will show that comparisons along these lines must be carefully nuanced, and that relevant evidence is sometimes lacking. We go on to suggest ways that the study of birdsong may benefit from future research that is responsive to our inquiry.

#### 5.1. Among Bird Species That Learn an S2, Is There Evidence of Interference from S1 in S2 Production?

In its most rudimentary formulation, the S1-to-S2 interference question asks: If features of the S2 diverge from native singers’ song features, are the divergences associable to the first song of the S2 bird? In other words, does their S2 have an “S1 accent” in the same way that the L2 of humans could have an L1 accent i.e., comparable to L1 German speakers’ devoicing of word-final voiced consonants (e.g., “dog” as “dock”) in their L2 English? (For the contribution of the L1 to features of L2 accent, see, e.g., Kartushina and Frauenfelder 2014; Schepens et al. 2020)

To our knowledge, there have been no direct investigations of the possible influence of previously learned songs on later song learning. However, it would be compelling to investigate this phenomenon in species such as European starlings or common nightingales (see above). To this end, one could tutor birds on different sets of songs early in development

(e.g., S1a and S1b as distinct sets of S1 songs) and then tutor each group of birds on the same S2s in adulthood (e.g., S1a->S2 for one group and S1b->S2 for another group). As a control, one could expose a group of birds to the same sets of songs during development and in adulthood (S2->S2). Thereafter, one would investigate systematic differences in S2 production across groups to reveal the impact of different forms of S1 learning on S2 acquisition and production.

Broadly speaking, interference of this sort could be construed as early experiences shaping vocal imitation abilities. Because vocal imitation requires sensory learning, we would point out that previous studies have investigated how early vocal learning can shape sensory processing in songbirds (reviewed in [Woolley and Woolley 2020](#)). For example, neurons of the auditory system of male zebra finches demonstrate distinct response profiles depending on whether the birds are tutored by a conspecific zebra finch, by a different songbird species (e.g., Bengalese finch), or are prevented from learning song ([Amin et al. 2013](#); [Moore and Woolley 2019](#); [Woolley and Woolley 2020](#)). Furthermore, species variations in the “innate” organization of sensory and motor systems have been found to sculpt the nature of song learning. The songs of zebra finches and Bengalese finches differ in various ways, including the spectral and temporal organization of song syllables (reviewed in [Murphy et al. 2017](#); [Sakata and Yazaki-Sugiyama 2020](#)). Although each species has been found to be able to imitate the syllables of the other species, there is a tendency to organize heterospecific syllables into sequences typical of their own species and to produce them at a tempo that is characteristic of their own species (e.g., [Araki et al. 2016](#); [Clayton 1989](#)). As such, both experiential and innate factors shape how the avian nervous system processes and reproduces sounds.

*5.2. Among Bird Species That Learn an S2, Is There Evidence for S1 Attrition? And If So, Is S1 Attrition Age-Dependent? Relatedly, among Bird Species That Learn an S2, Is There Evidence of Effects of S2 on S1 Production?*

Attrition of L1 is a well-known and well-studied phenomenon (for collections on attrition, see [Köpke et al. 2007](#); [Schmid et al. 2004](#); also overview in [Schmid and Köpke 2017](#); for attrition in L1 speech, see [Mayr et al. 2012](#); for evidence of L2 effects on pronunciation in the L1, see [Kornder and Mennen 2021](#)). The depth and breadth to which L1 attrition occurs are dependent on the degree to which the L2 is dominant as well as the degree to which the L1 is entrenched, both of which are conditioned by the age at which routine use of the L2 begins (e.g., [Birdsong 2014, 2018](#); [Hopp and Schmid 2013](#); [Karayayla and Schmid 2019](#)). In attrition of pronunciation, the acoustic qualities of L1 sounds tend to drift in the direction of L2 sounds (e.g., [Bergmann et al. 2016](#); [de Leeuw et al. 2010](#); [Mayr et al. 2012](#)).<sup>15</sup>

In songbirds, S1 attrition subsequent to learning an S2 represents another phenomenon that has not been extensively studied. In the published experiments described above involving European starlings and common nightingales, researchers have noted that some S1s are no longer produced by adult songbirds, or are modified in some way, after being tutored with S2s in adulthood. However, it is unclear whether these features reflect some form of attrition, “overwriting,” or mixing of previously acquired vocalizations. In human pronunciation, the direction of drift of acoustic properties may change from L2 -> L1, then L1->L2, and back again in instances of change of residence or change of dominant language (cf. François Grosjean’s notion of “the wax and wane of languages” among bilinguals: [Grosjean 2010](#), chp. 8).

A similar phenomenon could occur in birdsong, albeit at a different time scale. For example, birds that produce song repertoires tend to repeat one song type over and over before switching to another song type (reviewed in [Byers and Kroodsmas 2009](#); [Sakata and Vehrencamp 2012](#); [Wiley 2000](#)), and one could examine how repeating one song type could influence the acoustic properties of a subsequent song type. To make such research relatable to work in humans, a key element would be to quantify the degree to which such influence is modulated by age.

### 5.3. Is There Evidence of “Code-Switching” or “Code-Mixing” in Songbirds?

Among bilingual humans, switching and mixing of languages in everyday conversation is common. It has been extensively studied in naturalistic contexts (e.g., [Muysken 2000](#); [Myers-Scotton 1993](#); [Poplack 1980](#); chapters in [Bullock and Toribio 2009](#)), as well as in laboratory contexts in both forced and voluntary paradigms (e.g., [de Bruin et al. 2018](#); [Heredia and Altarriba 2001](#); [Hernandez 2009](#)).

In considering avian parallels, we may conceptualize human bilinguals’ language switching and mixing behaviors in terms of their deployment of two learned sets of vocalizations. In this sense, avian parallels of two types can be posited: the use of conspecific vs. heterospecific vocalizations in vocal mimics, and the use of songs that are learned early vs. late in a bird’s life (heuristically defined as S1 vs. S2 here).

With respect to the first type of parallel, greater racket-tailed drongos produce both conspecific and mimicked (heterospecific) alarm calls in threatening contexts ([Goodale and Kotagama 2006](#)), and brown thornbills use both conspecific and mimetic versions of functionally similar calls in behaviorally appropriate contexts (e.g., they use conspecific aerial alarm calls in addition to mimetic versions of these calls in response to flying predators ([Igic and Magrath 2014](#))). In this respect, the interleaving of conspecific and acquired heterospecific vocalizations seems to resemble code switching in humans (As one caveat, however, many of the conspecific calls that are studied in vocal mimics (e.g., alarm calls) are innate (i.e., do not require learning); accordingly, these examples of mixing could simply reflect the interleaving of innate and learned calls). As an example of the second analog to bilingual human interaction, some species of vocal mimics are known to integrate (putatively) learned heterospecific vocalizations (S2) into their learned S1 (see [Kelley et al. 2008](#)).

In addition, some types of *vocal exchanges* in songbirds could be related to code-switching or code-mixing. Male song sparrows defend their territories from other males of the same species, including other neighboring males (reviewed in [Beecher 2017](#)). Song sparrows generally produce between 5–13 song types in their repertoire, with some number of these song types being shared between conspecific neighbors. With neighbors, song sparrows use shared vs. non-shared song types depending on the social context and “intent” of communication. If a territory holder initiates a vocal interaction with a neighbor using a shared song type, and the neighbor replies by producing the same song type (“type matching”), this signals an escalation of aggression by the neighbor. If the neighbor replies with a different song type but one that the two birds share (“repertoire matching”), this suggests a less aggressive response by the neighbor. Moreover, if the neighbor replies with a non-shared song type, this signals a termination of the interaction. In this respect, the switching between shared and non-shared song types resembles switching between L1 and L2 speech systems between interlocutors.

## 6. Conclusions

Humans are not the only animals that use learned vocal signals for communication. Songbirds are among the most abundant and well-studied non-human animals that also learn their vocalizations through experience, thereby representing a potentially powerful model for exploring the biological and experiential foundations of speech and language in humans. By the same token, and as we have shown here, the human model can be a rich source of new questions about birdsong learning and use.

Although previous works have generally examined parallels between the acquisition of “primary” languages and songs (L1 and S1, respectively), here we have exposed potential areas of overlap between humans and songbirds in vocal learning following the initial acquisition. We have discussed the diversity of song learning strategies in songbirds and the types of learning that most closely resemble L2 speech acquisition; we have defined the parameters of birdsong and how they relate to speech; we have provided an overview of social, age, and behavioral contributions to variability in learning outcomes; and we have brought to light the need for caution in the interpretation of data in songbirds. We have

also pointed to research areas in which birdsong studies historically diverge from studies in bilingualism, and have suggested ways that future experiments can lend insight into avian behavior and learning. In summarizing some of our main points (Figure 5), we hope that our work has illustrated, and will promote, meaningful inter-disciplinary discourse in the domains of communicative learning and behaviors.



- Songbirds learn their vocalizations in ways that resemble how humans acquire speech.
- Some songbird species can learn novel vocalizations (S2) as adults, after they acquired previous sets of vocalizations (S1), resembling L2 speech acquisition in adult humans.
- Areas where the two disciplines converge, and allow for interdisciplinary dialogue, include biological foundations of learning, age-related changes in learning ability, and variability in learning outcomes.
- Features of human L2 speech acquisition and bilingual speech practices can guide birdsong research. Future studies might examine S1-S2 / S2-S1 influences, S1 attrition following S2 learning, S1-S2 switching, S1-entrenchment effects on S2 learning, and social influences on S2 learning.

**Figure 5.** Overview of some main points of the article. (Silhouettes: human; <https://pxhere.com/en/photo/1583761> (accessed on 13 December 2021); bird <https://pixabay.com/vectors/bird-stand-silhouette-nature-2803709> (accessed on 13 December 2021)).

**Author Contributions:** All the authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council (05016-2016) to J.T.S.

**Institutional Review Board Statement:** Procedures performed were in accordance with McGill University Animal Care and Use Committee protocols (2012-7149, 17 June 2021), as well as guidelines from the Canadian Council on Animal Care.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. Data are not publicly available because they are part of ongoing studies.

**Acknowledgments:** We are grateful to two reviewers for their helpful comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Notes

- <sup>1</sup> Parrots and hummingbirds are other species that learn their vocalizations but will not be the focus here. As an additional clarification, to be amplified later, we note that songbirds, in the same manner as other birds, use their vocalizations during social interactions (e.g., males courting females or defending their territory from other males: [Catchpole and Slater 2008](#); [Nowicki and Searcy 2014](#); [Podos and Sung 2020](#)).
- <sup>2</sup> However, a recent paper documents evidence of human speech imitation by musk ducks ([ten Cate and Fullagar 2021](#)).
- <sup>3</sup> In the interest of clarity, we highlight a distinction between songs and calls. Whereas songs consist of multiple, learned vocal elements that are strung together in rapid succession, calls are most often produced as single vocalizations (i.e., calls are generally not concatenated together to form stereotyped sequences of sounds) and often do not require learning (but, for discussions of calls that do require learning: e.g., [Elie and Theunissen 2020](#)). In addition, songs and calls are used in different social contexts: songs are used primarily during mating or aggressive interactions, whereas diverse types of calls are used in different contexts including feeding, defense, and bonding contexts ([Elie and Theunissen 2020](#); [Marler 2004](#)).
- <sup>4</sup> Although the terms “tutor” and “pupil” could suggest some degree of active teaching, there are insufficient data to indicate that directed instruction takes place in the context of vocal learning by songbirds; see [Caro and Hauser \(1992\)](#) for their description of teaching.
- <sup>5</sup> Although we, similarly to other researchers, define mimicry as the “imitation of all types of non-conspecific sounds: other species, anthropogenic (e.g., dog whistle, chainsaw), and environmental noises (e.g., water drip, leaves rustling)” ([Goller and Shizuka 2018](#)), others define mimicry based on the functional consequences of the vocalizations (e.g., whether individuals of the different species respond in the “appropriate” way to the mimic’s imitation of that species vocalization; [Dalziell et al. 2015](#)). We align our discussions of imitation fidelity to the degree to which the speech sounds of an L2 speaker acoustically resemble those same

sounds used by L1 speakers of that language. That being said, a number of L2 researchers have emphasized the extent to which L2 speech serves the social function of acceptance (e.g., “passing for” a native speaker, Piller 2002), a reminder that function, purpose and intent are not irrelevant in this context.

- 6 A video of birds’ imitations of human speech: <https://www.youtube.com/watch?v=N5YbWHrnjrg> (accessed on 13 December 2021).
- 7 The existence and nature of periods for optimal speech acquisition continue to be debated, but generally speaking, it is agreed that humans are able to learn new languages and sounds throughout their lives, that attainment of new language and pronunciation features declines but does not cease over age of learning, and that factors unrelated to neurological maturation, detailed in later sections of this review, can condition the outcomes of L2 speech learning (e.g., Birdsong 2017, 2018; Flege et al. 1999; Flege and Bohn 2021; Werker and Hensch 2015).
- 8 An example of how will and identity determine specific features of L2 pronunciation is noted by Walters (2011). By choice, Tunisian women speakers of L2 French pronounce French /r/ as the uvular [ʁ], just as native French speakers do. Tunisian men, on the other hand, choose to pronounce their L2 French /r/ as apical [r]. In this way they deliberately distinguish themselves from Tunisian women *and* from native speakers of standard French. Importantly, in Tunisian Arabic (the L1 of Tunisian men and women), both /r/ and /ʁ/ are produced, and in fact both are phonemic. Thus, articulating the French-like [ʁ] among Tunisian speakers of L2 French is not a question of having to master a new speech sound, but a matter of assertion of identity.
- 9 Estimates of individual or species variation in repertoire size can dramatically vary depending on the unit of measurement: e.g., repertoire size as the number of song types, phrase types (strophes or motifs), syllable types, or note types (Creanza et al. 2016). (As an analogy, one might imagine quantifying the number of words vs. syllables in lexicons). Rather than limit ourselves to a single definition of repertoire size, in relevant contexts we will consider a range of “units” of birdsong and will explicitly indicate the unit of analysis.
- 10 This could be likened to discovering that a friend of yours is able to speak a different language, one that they learned while growing up.
- 11 Given the emphasis on experimental design for providing compelling evidence of adult vocal learning, it should be mentioned that previous studies would be strengthened with the inclusion of a control group in which birds are not exposed to tutor songs in adulthood. For example, the European starlings that demonstrated vocal changes in response to late tutoring (Chaiken et al. 1994) were exposed to tutor songs during both development (“early tutors”) and in adulthood (“late tutors”), but it is possible that starlings that were exposed to tutor songs only during development could have “improvised” the early tutor songs in a way that caused their songs to coincidentally match late tutor songs. The degree to which this alternative explanation accounts for published observations remains unclear (because the full range of songs used for tutoring is generally not published along with these papers), but it stands as an important control to consider for future experiments.
- 12 There are a number of papers describing adult vocal learning in birds raised without exposure to adult song during development (e.g., in zebra finches, canaries, and brown-headed cowbirds: Gobes et al. 2019; Lehongre et al. 2009; Leitner and Catchpole 2007; O’Loughlen and Rothstein 2010; Sakata and Yazaki-Sugiyama 2020). Although orthogonal to our highlighting of parallels between S2 learning and L2 speech acquisition (i.e., acquiring new vocalizations after learning a previous set of vocalizations), these are compelling and useful examples because adult learning in these cases cannot be attributed to developmental song learning. We refer readers interested in this topic to various reviews and primary research articles cited in this paragraph.
- 13 A number of studies have found that individual neurons in HVC are “tuned” to multiple song types. For example, adult male swamp sparrows produce 2–5 different song types that consist of distinct sounds. Although various neurons in the HVC of male song sparrows are activated when the bird hears only one of the bird’s song types (i.e., neurons are “tuned” to one song type), many HVC neurons are activated in response to hearing multiple song types that are produced by the bird (Mooney et al. 2001). Similar findings have been observed in white-crowned sparrows (Prather et al. 2010). These findings suggest that there is not a simple one-to-one correspondence between the size of a brain area (or number of neurons in a brain area) and the repertoire of sounds that an individual songbird uses for communication.
- 14 Studies of AoA-related morphosyntactic attainment, which are more numerous than those for pronunciation, reveal a similar pattern of decline with greater dispersion at later AoA (e.g., DeKeyser et al. 2010; Chen and Hartshorne 2021; Flege et al. 1999; Hartshorne et al. 2018).
- 15 In the area of morphosyntax, L2 effects on the L1 are seen in online processing and in judgments of grammaticality (Kasparian and Steinhauer 2017), and have been attributed by different researchers to changes in representational knowledge and to non-pathological neurocognitive changes traceable to dominance factors (e.g., Steinhauer and Kasparian 2020 and references therein).

## References

- Abutalebi, Jubin. 2008. Neural aspects of second language representation and language control. *Acta Psychologica* 128: 466–78. [CrossRef] [PubMed]
- Airey, David C., and Timothy J. DeVoogd. 2000. Greater song complexity is associated with augmented song system anatomy in zebra finches. *Neuroreport* 11: 1749–54. [CrossRef]

- Airey, David C., Donald E. Kroodsmas, and Timothy J. DeVoogd. 2000. Differences in the complexity of song tutoring cause differences in the amount learned and in dendritic spine density in a songbird telencephalic song control nucleus. *Neurobiology of Learning and Memory* 73: 274–81. [\[CrossRef\]](#)
- Amin, Noopur, Michael Gastpar, and Frédéric E. Theunissen. 2013. Selective and efficient neural coding of communication signals depends on early acoustic and social environment. *PLoS ONE* 8: e61417. [\[CrossRef\]](#)
- Araki, Makoto, Manesh M. Bandi, and Yoko Yazaki-Sugiyama. 2016. Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354: 1282–87. [\[CrossRef\]](#) [\[PubMed\]](#)
- Au, Terry Kit-Fong, Leah M. Knightly, Sun-Ah Jun, and Janet S. Oh. 2002. Overhearing a language during childhood. *Psychological Science* 13: 238–43. [\[CrossRef\]](#)
- Ball, Gregory F., and Jacques Balthazard. 2020. Sex differences and similarities in the neural circuit regulating song and other reproductive behaviors in songbirds. *Neuroscience & Biobehavioral Reviews* 118: 258–69. [\[CrossRef\]](#)
- Balmer, Timothy S., Vanessa M. Carels, Jillian L. Frisch, and Teresa A. Nick. 2009. Modulation of perineuronal nets and parvalbumin with developmental song learning. *Journal of Neuroscience* 29: 12878–85. [\[CrossRef\]](#)
- Baptista, Luis F., and Lewis Petrinovich. 1986. Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behaviour* 34: 1359–71. [\[CrossRef\]](#)
- Barrington, Daines. 1773. Experiments and observations on the singing of birds. *Philosophical Transactions of the Royal Society of London* 63: 249–91.
- Beckers, Gabriël J. L., Brian S. Nelson, and Roderick A. Suthers. 2004. Vocal-tract filtering by lingual articulation in a parrot. *Current Biology* 14: 1592–97. [\[CrossRef\]](#)
- Beecher, Michael D. 2017. Birdsong learning as a social process. *Animal Behaviour* 124: 233–46. [\[CrossRef\]](#)
- Beecher, Michael D., and Eliot A. Brenowitz. 2005. Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution* 20: 143–49. [\[CrossRef\]](#)
- Bergmann, Christopher, Amber Nota, Simone A. Sprenger, and Monika S. Schmid. 2016. L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics* 58: 71–86. [\[CrossRef\]](#)
- Beurdeley, Marine, Julien Spatazza, Henry H. C. Lee, Sayaka Sugiyama, Clémence Bernard, Ariel A. Di Nardo, Takao K. Hensch, and Alain Prochiantz. 2012. Otx2 binding to perineuronal nets persistently regulates plasticity in the mature visual cortex. *Journal of Neuroscience* 32: 9429–37. [\[CrossRef\]](#)
- Birdsong, David. 1999. Introduction: Whys and why not of the Critical Period Hypothesis for second language acquisition. In *Second Language Acquisition and the Critical Period Hypothesis*. Edited by David Birdsong. Mahwah: Lawrence Erlbaum, pp. 1–22.
- Birdsong, David. 2007. Nativelike pronunciation among late learners of French as a second language. In *Language Experience in Second Language Speech Learning, in Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Amsterdam: John Benjamins, pp. 99–116.
- Birdsong, David. 2009. Age and the end state of second language acquisition. In *The New Handbook of Second Language Acquisition*, 2nd ed. Edited by William C. Ritchie and Tej K. Bhatia. Bingley: Emerald, pp. 401–24.
- Birdsong, David. 2014. Dominance and age in bilingualism. *Applied Linguistics* 35: 374–92. [\[CrossRef\]](#)
- Birdsong, David. 2017. Critical periods. In *Oxford Bibliographies in Linguistics*. Edited by Mark Aronoff. New York: Oxford University Press.
- Birdsong, David. 2018. Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology* 9: 81. [\[CrossRef\]](#)
- Bolhuis, Johan J., Kazuo Okanoya, and Constance Scharff. 2010. Twitter evolution: Converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience* 11: 747–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bongaerts, Theo. 1999. Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In *Second Language Acquisition and the Critical Period Hypothesis*. Edited by David Birdsong. Mahwah: Lawrence Erlbaum, pp. 133–59.
- Bradbury, Jack W., and Sandra L. Vehrencamp. 2011. *Principles of Animal Communication*, 2nd ed. Massachusetts: Sinauer Associates.
- Brainard, Michael S., and Allison J. Doupe. 2000. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404: 762–66. [\[CrossRef\]](#) [\[PubMed\]](#)
- Brainard, Michael S., and Allison J. Doupe. 2002. What songbirds teach us about learning. *Nature* 417: 351–58. [\[CrossRef\]](#)
- Brainard, Michael S., and Allison J. Doupe. 2013. Translating birdsong: Songbirds as a model for basic and applied medical research. *Annual Review of Neuroscience* 36: 489–517. [\[CrossRef\]](#)
- Brenowitz, Eliot A., and Michael D. Beecher. 2005. Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neurosciences* 28: 127–32. [\[CrossRef\]](#)
- Bullock, Barbara E., and Almeida Jacqueline Toribio, eds. 2009. *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge: Cambridge University Press. [\[CrossRef\]](#)
- Byers, Bruce E., and Donald E. Kroodsmas. 2009. Female mate choice and songbird song repertoires. *Animal Behaviour* 77: 13–22. [\[CrossRef\]](#)
- Carey, Daniel, Marc E. Miquel, Bronwen G. Evans, Patti Adank, and Carolyn McGettigan. 2017. Functional brain outcomes of L2 speech learning emerge during sensorimotor transformation. *NeuroImage* 159: 18–31. [\[CrossRef\]](#) [\[PubMed\]](#)
- Caro, Tim M., and Marc D. Hauser. 1992. Is there teaching in nonhuman animals? *Quarterly Review of Biology* 67: 151–74. [\[CrossRef\]](#)

- Carouso-Peck, Samantha, Otilia Menyhart, Timothy J. DeVoogd, and Michael H. Goldstein. 2020. Contingent parental responses are naturally associated with zebra finch song learning. *Animal Behaviour* 165: 123–32. [\[CrossRef\]](#)
- Catchpole, Clive K., and Peter J. B. Slater. 2008. *Bird Song: Biological Themes and Variations*, 2nd ed. Cambridge: Cambridge University Press.
- Chaiken, Marthaleah, Jörg Böhner, and Peter Marler. 1994. Repertoire turnover and the timing of song acquisition in European starlings. *Behaviour* 128: 25–39. [\[CrossRef\]](#)
- Chen, Tony, and Joshua K. Hartshorne. 2021. More evidence from over 1.1 million subjects that the critical period for syntax closes in late adolescence. *Cognition* 214: 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chen, Yining, and Jon T. Sakata. 2021. Norepinephrine in the avian auditory cortex enhances developmental song learning. *Journal of Neurophysiology* 125: 2397–407. [\[CrossRef\]](#)
- Chen, Yining, Laura E. Matheson, and Jon T. Sakata. 2016. Mechanisms underlying the social enhancement of vocal learning. *Proceedings of the National Academy of Sciences of the United States of America* 113: 6641–46. [\[CrossRef\]](#)
- Choe, Ha Na, and Erich D. Jarvis. 2021. The role of sex chromosomes and sex hormones in vocal learning systems. *Hormones and Behavior* 132: 104978. [\[CrossRef\]](#) [\[PubMed\]](#)
- Choi, Jiyoun, Mirjam Broersma, and Anne Cutler. 2017. Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences of the United States of America* 114: 7307–12. [\[CrossRef\]](#) [\[PubMed\]](#)
- Clayton, Nicky S. 1989. The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour* 109: 163–75. [\[CrossRef\]](#)
- Cornez, Gilles, Farrah N. Madison, Annemie Van der Linden, Charlotte Cornil, Kathleen M. Yoder, Gregory F. Ball, and Jacques Balthazart. 2017. Perineuronal nets and vocal plasticity in songbirds: A proposed mechanism to explain the difference between closed-ended and open-ended learning. *Developmental Neurobiology* 77: 975–94. [\[CrossRef\]](#)
- Cornez, Gilles, Elisabeth Jonckers, Sita M. Ter Haar, Annemie Van der Linden, Charlotte A. Cornil, and Jacques Balthazart. 2018. Timing of perineuronal net development in the zebra finch song control system correlates with developmental song learning. *Proceedings of the Royal Society B: Biological Sciences* 285: 20180849. [\[CrossRef\]](#) [\[PubMed\]](#)
- Coumel, Marion, Markus Christiner, and Susanne Maria Reiterer. 2019. Second language accent faking ability depends on musical abilities, not on working memory. *Frontiers in Psychology* 10: 257. [\[CrossRef\]](#) [\[PubMed\]](#)
- Creanza, Nicole, Laurel Fogarty, and Marcus W. Feldman. 2016. Cultural niche construction of repertoire size and learning strategies in songbirds. *Evolutionary Ecology* 30: 285–305. [\[CrossRef\]](#)
- Dai, Jennifer B., Yining Chen, and Jon T. Sakata. 2018. EGR-1 expression in catecholamine-synthesizing neurons reflects auditory learning and correlates with responses in auditory processing areas. *Neuroscience* 379: 415–27. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dalziell, Anastasia H., Justin A. Welbergen, Branislav Ijic, and Robert D. Magrath. 2015. Avian vocal mimicry: A unified conceptual framework. *Biological Reviews of the Cambridge Philosophical Society* 90: 643–68. [\[CrossRef\]](#) [\[PubMed\]](#)
- de Bruin, Angela, Arthur G. Samuel, and Jon Anoni Duñabeitia. 2018. Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language* 103: 28–43. [\[CrossRef\]](#)
- de Leeuw, Esther, Monika S. Schmid, and Ineke Mennen. 2010. The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition* 13: 33–40. [\[CrossRef\]](#)
- Dehaene, Stanislas, Emmanuel Dupoux, Jacques Mehler, Laurent Cohen, Eraldo Paulesu, Daniela Perani, Pierre-François van de Moortele, Stéphane Lehericy, and Denis Le Bihan. 1997. Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8: 3809–15. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dehaene-Lambertz, Ghislaine, Stanislas Dehaene, and Lucie Hertz-Pannier. 2002. Functional neuroimaging of speech perception in infants. *Science* 298: 2013–15. [\[CrossRef\]](#)
- DeKeyser, Robert, Iris Alfi-Shabtay, and Dorit Ravid. 2010. Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31: 413–38. [\[CrossRef\]](#)
- Derégnaucourt, Sébastien, Colline Poirier, Anne Van der Kant, Annemie Van der Linden, and Manfred Gahr. 2013. Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *Journal of Physiology-Paris* 107: 210–18. [\[CrossRef\]](#) [\[PubMed\]](#)
- Derrickson, Kim C. 1987. Yearly and situational changes in the estimate of repertoire size in Northern Mockingbirds (*Mimus polyglottos*). *The Auk* 104: 198–207. [\[CrossRef\]](#)
- DeVoogd, Timothy J. 2004. Neural constraints on the complexity of avian song. *Brain, Behavior and Evolution* 63: 221–32. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dhawale, Ashesh K., Maurice A. Smith, and Bence P. Ölveczky. 2017. The role of variability in motor learning. *Annual Review of Neuroscience* 40: 479–98. [\[CrossRef\]](#)
- Díaz, Begoña, Cristina Baus, Carles Escera, Albert Costa, and Núria Sebastián-Gallés. 2008. Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America* 105: 16083–88. [\[CrossRef\]](#)
- Díaz, Begoña, Holger Mitterer, Mirjam Broersma, Carles Escera, and Núria Sebastián-Gallés. 2016. Variability in L2 phonemic learning originates from speech-specific capabilities: An MMN study on late bilinguals. *Bilingualism: Language and Cognition* 19: 955–70. [\[CrossRef\]](#)
- Doupe, Allison J. 1997. Song-and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *Journal of Neuroscience* 17: 1147–67. [\[CrossRef\]](#) [\[PubMed\]](#)

- Doupe, Allison J., and Patricia K. Kuhl. 1999. Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 22: 567–631. [CrossRef] [PubMed]
- Doya, Kenji, and Terrence J. Sejnowski. 1995. A novel reinforcement model of birdsong vocalization learning. In *Advances in Neural Information Processing Systems*. Edited by Gerald Tesauro, David S. Touretzky and Todd K. Leen. Cambridge: MIT Press, pp. 101–8.
- Eales, Lucy A. 1985. Song learning in zebra finches: Some effects of song model availability on what is learnt and when. *Animal Behaviour* 33: 1293–300. [CrossRef]
- Elemans, Coen P. H., Jeppe Have Rasmussen, Christian T. Herbst, Daniel Normen Düring, Sue Anne Zollinger, Henrik Brumm, Krishna C. Srivastava, Niels Svane, Ming Ding, Ole Naesbye Larsen, and et al. 2015. Universal mechanisms of sound production and control in birds and mammals. *Nature Communications* 6: 8978. [CrossRef] [PubMed]
- Elie, Julie E., and Frédéric Theunissen. 2020. The neuroethology of vocal communication in songbirds: Production and perception of a call repertoire. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Cham: Springer Nature, pp. 175–209. [CrossRef]
- Eriksen, Ane, Tore Slagsvold, and Helene Marie Lampe. 2011. Vocal plasticity—are pied flycatchers, *Ficedula hypoleuca*, open-ended learners? *Ethology* 117: 188–98. [CrossRef]
- Fee, Michale S., and Jesse H. Goldberg. 2011. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* 198: 152–70. [CrossRef]
- Fehér, Olga, Haibin Wang, Sigal Saar, Partha P. Mitra, and Ofer Tchernichovski. 2009. *De novo* establishment of wild-type song culture in the zebra finch. *Nature* 459: 564–68. [CrossRef] [PubMed]
- Feng, Gangyi, Yu Li, Shen-Mou Hsu, Patrick C. M. Wong, Tai-Li Chou, and Bharath Chandrasekaran. 2021. Emerging native-similar neural representations underlie non-native speech category learning success. *Neurobiology of Language* 2: 280–307. [CrossRef] [PubMed]
- Fitch, W. Tecumseh, Ludwig Huber, and Thomas Bugnyar. 2010. Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron* 65: 795–814. [CrossRef]
- Flege, James Emil. 1998. The role of subject and phonetic variables in second-language learning. In *CLS 34: The Panels*. Edited by M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson and Tamra Wysocki. Chicago: Chicago Linguistic Society, pp. 213–32.
- Flege, James E. 1999. Age of learning and second language speech. In *Second Language Acquisition and the Critical Period Hypothesis*. Edited by David Birdsong. Mahwah: Lawrence Erlbaum, pp. 101–31.
- Flege, James Emil. 2018. A non-critical period for second language. In *A Sound Approach to Language Matters: In Honor of Ocke-Schwen Bohn*. Edited by Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen and Mette Hjortshøj Sørensen. Aarhus: Aarhus University Library Scholarly Publishing Services, pp. 501–41.
- Flege, James Emil, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Rtree Wayland. Cambridge: Cambridge University Press, pp. 3–83. [CrossRef]
- Flege, James Emil, Murray J. Munro, and Ian R. A. MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustic Society of America* 97: 3125–34. [CrossRef] [PubMed]
- Flege, James Emil, Grace H. Yeni-Komshian, and Serena Liu. 1999. Age constraints on second-language acquisition. *Journal of Memory and Language* 41: 78–104. [CrossRef]
- Friederici, Angela. 2011. The brain basis of language processing: From structure to function. *Physiological Reviews* 91: 1357–92. [CrossRef] [PubMed]
- Gammon, David E. 2013. How is model selection determined in a vocal mimic?: Tests of five hypotheses. *Behaviour* 150: 1375–97. [CrossRef]
- Gammon, David E. 2020. Are northern mockingbirds classic open-ended song learners? *Ethology* 126: 1038–47. [CrossRef]
- Garcia-Sierra, Adrian, Maritza Rivera-Gaxiola, Cherie R. Percaccio, Barbara T. Conboy, Harriett Romo, Lindsay Klarman, Sophia Ortiz, and Patricia K. Kuhl. 2011. Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word recognition. *Journal of Phonetics* 39: 546–57. [CrossRef]
- Gardner, Timothy J., Felix Naef, and Fernando Nottebohm. 2005. Freedom and rules: The acquisition and reprogramming of a bird's learned song. *Science* 308: 1046–49. Available online: <https://www.science.org/doi/10.1126/science.1108214> (accessed on 13 December 2021). [CrossRef]
- Gaskell, M. Gareth, and Andrew W. Ellis. 2009. Word learning and lexical development across the lifespan. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364: 3607–15. [CrossRef]
- Geberzahn, Nicole, and Henrike Hultsch. 2003. Long-time storage of song types in birds: Evidence from interactive playbacks. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: 1085–90. [CrossRef]
- Geberzahn, Nicole, Henrike Hultsch, and Dietmar Todt. 2002. Latent song type memories are accessible through auditory stimulation in a hand-reared songbird. *Animal Behaviour* 64: 783–90. [CrossRef]
- Gervain, Judit, and Jacques Mehler. 2010. Speech perception and language acquisition in the first year of life. *Annual Review of Psychology* 61: 191–218. [CrossRef]
- Gobes, Sharon MH, and Johan J. Bolhuis. 2007. Birdsong memory: A neural dissociation between song recognition and production. *Current Biology* 17: 789–93. [CrossRef]

- Gobes, Sharon M. H., Rebecca B. Jennings, and Rie K. Maeda. 2019. The sensitive period for auditory-vocal learning in the zebra finch: Consequences of limited-model availability and multiple-tutor paradigms on song imitation. *Behavioural Processes* 163: 5–12. [CrossRef]
- Goffinet, Jack, Samuel Brudner, Richard Mooney, and John Pearson. 2021. Low-dimensional learned feature spaces quantify individual and group differences in vocal repertoires. *eLife* 10: e67855. [CrossRef]
- Golestani, Narly, and Robert J. Zatorre. 2004. Learning new sounds of speech: Reallocation of neural substrates. *NeuroImage* 21: 494–506. [CrossRef]
- Golestani, Narly, Tomás Paus, and Robert J. Zatorre. 2002. Anatomical correlates of learning novel speech sounds. *Neuron* 35: 997–1010. [CrossRef]
- Golestani, Narly, Nicolas Molko, Stanislas DeHaene, Denis Le Bihan, and Christophe Pallier. 2007. Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex* 17: 575–82. [CrossRef]
- Goller, Maria, and Daizaburo Shizuka. 2018. Evolutionary origins of vocal mimicry in songbirds. *Evolution Letters* 2: 417–26. [CrossRef] [PubMed]
- Gomez-Marin, Alex, Joseph J. Patton, Adam R. Kampff, Rui M. Costa, and Zachary F. Mainen. 2014. Big behavioral data: Psychology, ethology and the foundations of neuroscience. *Nature Neuroscience* 17: 1455–62. [CrossRef]
- Goodale, Eben, and Sarath W. Kotagama. 2006. Context-dependent vocal mimicry in a passerine bird. *Proceedings of the Royal Society B: Biological Sciences* 273: 875–80. [CrossRef] [PubMed]
- Grosjean, François. 2010. *Bilingual: Life and Reality*. Cambridge: Harvard University Press.
- Guion, Susan G., James E. Flege, and Jonathan D. Loftin. 2000. The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics* 28: 27–42. [CrossRef]
- Hartshorne, Joshua K., Joshua B. Tenenbaum, and Steven Pinker. 2018. A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition* 177: 263–77. [CrossRef]
- Hausberger, Martine, Peter F. Jenkins, and Jeremy Keene. 1991. Species-specificity and mimicry in bird song: Are they paradoxes? A reevaluation of song mimicry in the European starling. *Behaviour* 117: 53–81. [CrossRef]
- Helekar, Santosh, ed. 2013. *Animal Models of Speech and Language Disorders*. New York: Springer.
- Heredia, Roberto R., and Jeanette Altarriba. 2001. Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science* 10: 164–68. [CrossRef]
- Hernandez, Arturo. 2009. Language switching in the bilingual brain: What's next? *Brain and Language* 108: 133–40. [CrossRef] [PubMed]
- Hopp, Holger, and Monika S. Schmid. 2013. Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics* 34: 361–94. [CrossRef]
- Hough, Gerald E. II, Douglas A. Nelson, and Susan F. Volman. 2000. Re-expression of songs deleted during vocal development in white-crowned sparrows, *Zonotrichia leucophrys*. *Animal Behaviour* 60: 279–87. [CrossRef] [PubMed]
- Houx, Bart B., and Carel ten Cate. 1998. Do contingencies with tutor behaviour influence song learning in zebra finches? *Behaviour* 135: 599–614. Available online: <https://www.jstor.org/stable/4535547> (accessed on 13 December 2021). [CrossRef]
- Hu, Xiaochen, Hermann Ackermann, Jason A. Martin, Michael Erb, Susanne Winkler, and Susanne M. Reiterer. 2012. Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain and Language* 127: 366–76. [CrossRef] [PubMed]
- Hyltenstam, Kenneth, and Niclas Abrahamsson. 2003. Maturation constraints in SLA. In *The Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Malden: Blackwell, pp. 539–88.
- Igic, Branislav, and Robert D. Magrath. 2014. A songbird mimics different heterospecific alarm calls in response to different types of threat. *Behavioral Ecology* 25: 538–48. [CrossRef]
- Ingvalson, Erin M., Casandra Nowicki, Audrey Zong, and Patrick C. M. Wong. 2017. Non-native speech learning in older adults. *Frontiers in Psychology* 8: 148. [CrossRef] [PubMed]
- James, Logan S., and Jon T. Sakata. 2017. Learning biases underlie “universals” in avian vocal sequencing. *Current Biology* 27: 3676–82. [CrossRef]
- James, Logan S., Ronald Davies Jr, Chihiro Mori, Kazuhiro Wada, and Jon T. Sakata. 2020. Manipulations of sensory experiences during development reveal mechanisms underlying vocal learning biases in zebra finches. *Developmental Neurobiology* 80: 132–46. [CrossRef] [PubMed]
- Janik, Vincent M., and Peter J. Slater. 2000. The different roles of social learning in vocal communication. *Animal Behaviour* 60: 1–11. [CrossRef] [PubMed]
- Jarvis, Erich D. 2019. Evolution of vocal learning and spoken language. *Science* 366: 50–54. [CrossRef] [PubMed]
- Baylis, Jeffrey R. 1982. Avian vocal mimicry: Its function and evolution. In *Acoustic Communication in Birds*. Edited by Donald E. Kroodsma, Edward H. Miller and Henri Ouellet. New York: Academic Press, vol. 2, pp. 51–83.
- Jenkins, Jane, and Jennifer Setter. 2005. State of the art review article: Pronunciation. *Language Teaching* 38: 1–17. [CrossRef]
- Jones, Alex E., Carel ten Cate, and Peter J. B. Slater. 1996. Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology* 110: 354–69. [CrossRef]
- Karayayla, Tugba, and Monika S. Schmid. 2019. First language attrition as a function of age of onset of bilingualism: First language attainment of Turkish-English bilinguals in the United Kingdom. *Language Learning* 69: 106–42. [CrossRef]

- Kartushina, Natalia, and Ulrich H. Frauenfelder. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology* 5: 1246. [CrossRef]
- Kasparian, Kristina, and Karsten Steinhauer. 2017. When the second language takes the lead: Neurocognitive processing changes in the first language of adult attriters. *Frontiers in Psychology* 8: 389. [CrossRef]
- Keller, Georg B., and Richard H. R. Hahnloser. 2009. Neural processing of auditory feedback during vocal practice in a songbird. *Nature* 457: 187–90. [CrossRef] [PubMed]
- Kelley, Laura A., Rebecca L. Coe, Joah R. Madden, and Susan D. Healy. 2008. Vocal mimicry in songbirds. *Animal Behaviour* 76: 521–28. [CrossRef]
- Kipper, Silke, and Sarah Kiefer. 2010. Age-related changes in birds' singing styles: On fresh tunes and fading voices? *Advances in the Study of Behavior* 41: 77–118. [CrossRef]
- Klein, Denise, Brenda Milner, Robert J. Zatorre, Ernst Meyer, and Alan C. Evans. 1995. The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences of the United States of America* 92: 2899–903. [CrossRef] [PubMed]
- Köpke, Barbara. 2021. Language attrition: A matter of brain plasticity? Some preliminary thoughts. *Language, Interaction and Acquisition* 12: 110–32. [CrossRef]
- Köpke, Barbara, Monika S. Schmid, Merel Keijzer, and Susan Dostert, eds. 2007. *Language Attrition: Theoretical Perspectives*. Amsterdam: John Benjamins.
- Kornder, Lisa, and Ineke Mennen. 2021. Longitudinal developments in bilingual second language acquisition and first language attrition of speech: The case of Arnold Schwarzenegger. *Languages* 6: 61. [CrossRef]
- Kuhl, Patricia K. 2007. Is speech learning 'gated' by the social brain? *Developmental Science* 10: 110–20. [CrossRef]
- Lehongre, Katia, Thierry Aubin, and Catherine Del Negro. 2009. Influence of social conditions in song sharing in the adult canary. *Animal Cognition* 12: 823–32. [CrossRef]
- Leitner, Stefan, and Clive K. Catchpole. 2007. Song and brain development in canaries raised under different conditions of acoustic and social isolation over two years. *Developmental Neurobiology* 67: 1478–87. [CrossRef] [PubMed]
- Li, Ping, and Hyeonjeong Jeong. 2020. The social brain of language: Grounding second language learning in social interaction. *NPJ Science of Learning* 5: 1–9. [CrossRef]
- Lipkind, Dina, Gary F. Marcus, Douglas K. Bemis, Kazutoshi Sasahara, Nori Jacoby, Miki Takahasi, Kenta Suzuki, Olga Feher, Primoz Ravbar, Kazuo Okanoya, and et al. 2013. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498: 104–8. [CrossRef]
- Lipkind, Dina, Anja T. Zai, Alexander Hanuschkin, Gary F. Marcus, Ofer Tchernichovski, and Richard H. R. Hahnloser. 2017. Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nature Communications* 8: 1247. [CrossRef]
- Lipkind, Dina, Andreea Geambasu, and Clara C. Levelt. 2020. The development of structured vocalizations in songbirds and humans: A comparative analysis. *Topics in Cognitive Science* 12: 894–909. [CrossRef] [PubMed]
- Ljubičić, Iva, Julia Hyland Bruno, and Ofer Tchernichovski. 2016. Social influences on song learning. *Current Opinion in Behavioral Sciences* 7: 101–7. [CrossRef]
- London, Sarah E., and David F. Clayton. 2008. Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience* 11: 579–86. [CrossRef]
- Love, Jay, Amanda Hoepfner, and Franz Goller. 2019. Song feature specific analysis of isolate song reveals interspecific variation in learned components. *Developmental Neurobiology* 79: 350–69. [CrossRef]
- Mack, Molly. 1989. Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perceptual Psychophysics* 46: 187–200. [CrossRef] [PubMed]
- MacWhinney, Brian. 2005. A unified model of language acquisition. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Annette M. B. de Groot. New York: Oxford University Press, pp. 49–67.
- Marchman, Virginia A. 1993. Constraints on plasticity in a connectionist model of the English Past tense. *Journal of Cognitive Neuroscience* 5: 215–34. [CrossRef]
- Margoliash, Daniel, and Marc F. Schmidt. 2010. Sleep, off-line processing, and vocal learning. *Brain and Language* 115: 45–58. [CrossRef]
- Margoliash, Daniel, Cynthia Staicer, and Sue A. Inoue. 1994. The process of syllable acquisition in adult indigo buntings (*Passerina cyanea*). *Behaviour* 131: 39–64. Available online: <https://www.jstor.org/stable/4535227> (accessed on 13 December 2021).
- Marler, Peter. 1970. A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology* 71: 1–25. [CrossRef]
- Marler, Peter. 1997. Three models of song learning: Evidence from behavior. *Journal of Neurobiology* 33: 501–16. [CrossRef]
- Marler, Peter. 2004. Bird calls: Their potential for neurobiology. *Annals of the New York Academy of Sciences* 1016: 31–44. [CrossRef]
- Marler, Peter, and Miwako Tamura. 1962. Song "dialects" in three populations of white-crowned sparrows. *Condor* 64: 368–77. [CrossRef]
- Mayr, Robert, Sacha Price, and Ineke Mennen. 2012. First language attrition in the case of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition* 15: 687–700. [CrossRef]
- Mets, David G., and Michael S. Brainard. 2018. Genetic variation interacts with experience to determine interindividual differences in learned song. *Proceedings of the National Academy of Sciences of the United States of America* 115: 421–26. [CrossRef]

- Mets, David G., and Michael S. Brainard. 2019. Learning is enhanced by tailoring instruction to individual genetic differences. *eLife* 8: e47216. [\[CrossRef\]](#) [\[PubMed\]](#)
- Minagawa-Kawai, Yasuyo, Cristià Alejandrina, and Emmanuel Dupoux. 2011. Cerebral lateralization and early speech acquisition: A developmental scenario. *Developmental Cognitive Neuroscience* 1: 217–32. [\[CrossRef\]](#)
- Mol, Carien, Aoju Chen, René W. J. Kager, and Sita M. ter Haar. 2017. Prosody in birdsong: A review and perspective. *Neuroscience & Biobehavioral Reviews* 81: 167–80. [\[CrossRef\]](#)
- Mooney, Richard, William Hoese, and Stephen Nowicki. 2001. Auditory representation of the vocal repertoire in a songbird with multiple song types. *Proceedings of the National Academy of Sciences of the United States of America* 98: 12778–83. [\[CrossRef\]](#) [\[PubMed\]](#)
- Moore, Jordan M., and Sarah M. N. Woolley. 2019. Emergent tuning for learned vocalizations in auditory cortex. *Nature Neuroscience* 22: 1469–76. [\[CrossRef\]](#) [\[PubMed\]](#)
- Moore, Jordan M., Tamás Székely, József Büki, and Timothy J. DeVoogd. 2011. Motor pathway convergence predicts syllable repertoire size in oscine birds. *Proceedings of the National Academy of Sciences of the United States of America* 108: 16440–45. [\[CrossRef\]](#)
- Moorman, Sanne, Sharon MH Gobes, Maaïke Kuijpers, Amber Kerkhofs, Matthijs A. Zandbergen, and Johan J. Bolhuis. 2012. Human-like brain hemispheric dominance in birdsong learning. *Proceedings of the National Academy of Sciences* 109: 12782–87. [\[CrossRef\]](#)
- Mora, Joan C., Youssef Rochdi, and Hanna Kivistö-de-Souza. 2013. Mimicking accented speech as L2 phonological awareness. *Language Awareness* 23: 57–75. [\[CrossRef\]](#)
- Morrison, Robert G., and Fernando Nottebohm. 1993. Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *Journal of Neurobiology* 24: 1045–64. [\[CrossRef\]](#)
- Moyer, Alene. 2004. *Age, Accent and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Clevedon: Multilingual Matters.
- Moyer, Alene. 2014. Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics* 35: 418–40. [\[CrossRef\]](#)
- Murphy, Karagh, Logan S. James, Jon T. Sakata, and Jonathan F. Prather. 2017. Advantages of comparative studies to understand the neural basis of sensorimotor integration. *Journal of Neurophysiology* 118: 800–16. [\[CrossRef\]](#) [\[PubMed\]](#)
- Murphy, Karagh, Koedi Lawley, Perry Smith, and Jonathan F. Prather. 2020. New insights into the avian song system and neuronal control of learned vocalizations. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Switzerland: Springer Nature, pp. 65–92. [\[CrossRef\]](#)
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press. [\[CrossRef\]](#)
- Myers-Scotton, Carol. 1993. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Clarendon.
- Nelson, Douglas A., and Peter Marler. 1994. Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences of the United States of America* 91: 10498–501. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nick, Teresa A., and Masakazu Konishi. 2005a. Neural auditory selectivity develops in parallel with song. *Journal of Neurobiology* 62: 469–81. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nick, Teresa A., and Masakazu Konishi. 2005b. Neural song preference during vocal learning in the zebra finch depends on age and state. *Journal of Neurobiology* 62: 231–42. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nowicki, Stephen, and William A. Searcy. 2014. The evolution of vocal learning. *Current Opinion in Neurobiology* 28: 48–53. [\[CrossRef\]](#) [\[PubMed\]](#)
- Odom, Karan J., Michelle L. Hall, Katharina Riebel, Kevin E. Omland, and Naomi E. Langmore. 2014. Female song is widespread and ancestral in songbirds. *Nature Communications* 5: 1–6. [\[CrossRef\]](#)
- Oh, Janet S., Terry Kit-Fong Au, and Sun-Ah Jun. 2010. Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language* 37: 1123–32. [\[CrossRef\]](#)
- O’Loughlen, Adrian L., and Stephen I. Rothstein. 2010. Delayed sensory learning and development of dialect songs in brown-headed cowbirds, *Molothrus ater*. *Animal Behaviour* 79: 299–311. [\[CrossRef\]](#)
- Olson, Elizabeth M., Rie K. Maeda, and Sharon M. H. Gobes. 2016. Mirrored patterns of lateralized neuronal activation reflect old and new memories in the avian auditory cortex. *Neuroscience* 330: 395–402. [\[CrossRef\]](#)
- Pallier, Christophe, Stanislas Dehaene, Jean-Baptiste Poline, Denis Le Bihan, A.-M. Argenti, Emmanuel Dupoux, and Jacques Mehler. 2003. Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex* 13: 155–61. [\[CrossRef\]](#)
- Park, Denise C., Thad A. Polk, Joseph A. Mikels, Stephan F. Taylor, and Christy Marshuetz. 2001. Cerebral aging: Integration of brain and behavioral models of cognitive function. *Dialogues in Clinical Neuroscience* 3: 151–65. [\[CrossRef\]](#) [\[PubMed\]](#)
- Paul, Avishek, Helen McLendon, Veronica Rally, Jon T. Sakata, and Sarah C. Woolley. 2021. Behavioral discrimination and time-series phenotyping of birdsong performance. *PLoS Computational Biology* 17: e1008820. [\[CrossRef\]](#)
- Petitto, Laura-Ann, Melody S. Berens, Ioulia Kovelman, M. H. Dubins, Kaja K. Jasinska, and M. Shalinsky. 2012. The “perceptual wedge hypothesis” as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language* 121: 130–43. [\[CrossRef\]](#)
- Petkov, Christopher I., and Erich Jarvis. 2012. Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience* 4: 12. [\[CrossRef\]](#)

- Pfaff, Jeremy A., Liana Zanette, Scott A. MacDougall-Shackleton, and Elizabeth A. MacDougall-Shackleton. 2007. Song repertoire size varies with HVC volume and is indicative of male quality in song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B: Biological Sciences* 274: 2035–40. [\[CrossRef\]](#)
- Pierce, Lara J., Denise Klein, Jen-Kai Chen, Audrey Delcenserie, and Fred Genesee. 2014. Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences of the United States of America* 111: 17314–19. [\[CrossRef\]](#)
- Piller, Ingrid. 2002. Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics* 6: 179–206. [\[CrossRef\]](#)
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [\[CrossRef\]](#)
- Pizzorusso, Tommaso, Paolo Medini, Nicoletta Berardi, Sabrina Chierzi, James W. Fawcett, and Lamberto Maffei. 2002. Reactivation of ocular dominance plasticity in the adult visual cortex. *Science* 298: 1248–51. [\[CrossRef\]](#)
- Podos, Jeffrey, and Ha-Cheol Sung. 2020. Vocal performance in songbirds: From mechanisms to evolution. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Cham: Springer Nature, pp. 245–68. [\[CrossRef\]](#)
- Poplack, Shana. 1980. “Sometimes I’ll start a sentence in Spanish y termino en español”: Toward a typology of code-switching. *Linguistics* 18: 581–618. [\[CrossRef\]](#)
- Prather, Jonathan F., Susan Peters, Stephen Nowicki, and Richard Mooney. 2010. Persistent representation of juvenile experience in the adult songbird brain. *Journal of Neuroscience* 30: 10586–98. [\[CrossRef\]](#)
- Prather, Jonathan F., Kazuo Okanoya, and Johan J. Bolhuis. 2017. Brains for birds and babies: Neural parallels between birdsong and speech acquisition. *Neuroscience & Biobehavioral Reviews* 81: 225–37. [\[CrossRef\]](#)
- Pytte, Carolyn L., Miles Gerson, Janet Miller, and John R. Kirn. 2007. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Developmental Neurobiology* 67: 1699–720. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ravbar, Primož, Dina Lipkind, Lucas C. Parra, and Ofer Tchernichovski. 2012. Vocal exploration is locally regulated during song learning. *Journal of Neuroscience* 32: 3422–32. [\[CrossRef\]](#)
- Reichelt, Amy C., Dominic J. Hare, Timothy J. Bussey, and Lisa M. Saksida. 2019. Perineuronal nets: Plasticity, protection, and therapeutic potential. *Trends in Neurosciences* 42: 458–70. [\[CrossRef\]](#) [\[PubMed\]](#)
- Reiterer, Susanne Maria, Xiaochen Hu, Michael Erb, Guiseppina Rota, Davide Nardo, Wolfgang Grodd, Susanne Winkler, and Hermann Ackermann. 2011. Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology* 2: 271. [\[CrossRef\]](#) [\[PubMed\]](#)
- Reiterer, Susanne M., Xiaochen Hu, T. A. Sumathi, and Nandini C. Singh. 2013. Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology* 4: 782. [\[CrossRef\]](#) [\[PubMed\]](#)
- Riebel, Katharina, Karan J. Odom, Naomi E. Langmore, and Michelle L. Hall. 2019. New insights from female bird song: Towards an integrated approach to studying male and female communication roles. *Biology Letters* 15: 20190059. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ripollés, Pablo, Josep Marco-Pallarés, Helena Alicart, Claus Tempelmann, Antonio Rodríguez-Fornells, and Toemme Noesselt. 2016. Intrinsic monitoring of learning success facilitates memory encoding via the activation of the SN/VTA-Hippocampal loop. *eLife* 5: e17441. [\[CrossRef\]](#)
- Roberts, Todd F., and Richard Mooney. 2013. Motor circuits help encode auditory memories of vocal models used to guide vocal learning. *Hearing Research* 303: 48–57. [\[CrossRef\]](#) [\[PubMed\]](#)
- Robinson, Cristina M., and Nicole Creanza. 2019. Species-level repertoire size predicts a correlation between individual song elaboration and reproductive success. *Ecology and Evolution* 9: 8362–77. [\[CrossRef\]](#)
- Robinson, Cristina M., Kate T. Snyder, and Nicole Creanza. 2019. Correlated evolution between repertoire size and song plasticity predicts that sexual selection on song promotes open-ended learning. *eLife* 8: e44454. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rota, Guiseppina, and Susanne Maria Reiterer. 2009. Cognitive aspects of pronunciation talent. In *Language Talent and Brain Activity*. Edited by Grzegorz Dogil and Susanne Maria Reiterer. Berlin: Mouton de Gruyter, pp. 67–96. [\[CrossRef\]](#)
- Sainburg, Tim, Marvin Thielk, and Timothy Q. Gentner. 2020. Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLoS Computational Biology* 16: e1008228. [\[CrossRef\]](#)
- Sakata, Jon T., and Michael S. Brainard. 2006. Real-time contributions of auditory feedback to avian vocal motor control. *Journal of Neuroscience* 26: 9619–28. [\[CrossRef\]](#)
- Sakata, Jon T., and Michael S. Brainard. 2008. Online contributions of auditory feedback to neural activity in avian song control circuitry. *Journal of Neuroscience* 28: 11378–90. [\[CrossRef\]](#)
- Sakata, Jon T., and Sandra L. Vehrencamp. 2012. Integrating perspectives on vocal performance and consistency. *Journal of Experimental Biology* 215: 201–9. [\[CrossRef\]](#)
- Sakata, Jon T., and Sarah C. Woolley. 2020. Scaling the levels of birdsong analysis. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Cham: Springer Nature, pp. 1–27. [\[CrossRef\]](#)
- Sakata, Jon T., and Yoko Yazaki-Sugiyama. 2020. Neural circuits underlying vocal learning in songbirds. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Cham: Springer Nature, pp. 29–63. [\[CrossRef\]](#)
- Schepens, Job Johannes, Roeland van Hout, and T. Florian Jaeger. 2020. Big data suggest strong constraints of linguistic similarity on adult language learning. *Cognition* 194: 104056. [\[CrossRef\]](#)

- Schmid, Monika S., and Barbara Köpke. 2017. The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism* 7: 637–67. [CrossRef]
- Schmid, Monika S., Barbara Köpke, Merel Keijzer, and Lina Weilemar, eds. 2004. *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam: John Benjamins.
- Seyfarth, Robert M., and Dorothy L. Cheney. 2010. Production, usage, and comprehension in animal vocalizations. *Brain and Language* 115: 92–100. [CrossRef]
- Simmonds, Anna J. 2015. A hypothesis on improving foreign accents by optimizing variability in vocal learning brain circuits. *Frontiers in Human Neuroscience* 9: 606. [CrossRef]
- Singleton, David. 1995. Introduction: A critical look at the Critical Period Hypothesis in Second Language Acquisition Research. In *The Age Factor in Second Language Acquisition: A Critical Look at the Critical Period Hypothesis*. Edited by David Singleton and Zsolt Lengyel. Clevedon: Multilingual Matters, pp. 1–29.
- Slevc, L. Robert, and Akira Miyake. 2006. Individual differences in second-language proficiency. Does musical ability matter? *Psychological Science* 17: 675–81. [CrossRef] [PubMed]
- Soha, Jill. 2017. The auditory template hypothesis: A review and comparative perspective. *Animal Behaviour* 124: 247–54. [CrossRef]
- Steinhauer, Karsten, and Kristina Kasparian. 2020. Brain plasticity in adulthood: ERP evidence for L1-attrition in lexicon and morphosyntax after predominant L2 use. *Language Learning* 70: 171–93. [CrossRef]
- Steinhauer, Karsten, Erin J. White, and John E. Drury. 2009. Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research* 25: 13–41. [CrossRef]
- Székel, Tamás, Clive K. Catchpole, Albert DeVoogd, Zsuzsa Marchl, and Timothy J. DeVoogd. 1996. Evolutionary changes in a song control area of the brain (HVC) are associated with evolutionary changes in song repertoire among European warblers (*Sylviidae*). *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263: 607–10. [CrossRef]
- Tachibana, Ryosuke O., Miki Takahashi, Neal A. Hessler, and Kazuo Okanoya. 2017. Maturation-dependent control of vocal temporal plasticity in a songbird. *Developmental Neurobiology* 77: 995–1006. [CrossRef]
- Tagarelli, Kaitlyn M., Kyle F. Shattuck, Peter E. Turkeltaub, and Michael T. Ullman. 2019. Language learning in the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *Neuroimage* 193: 178–200. [CrossRef]
- Takesian, Anne E., and Takao K. Hensch. 2013. Balancing plasticity/stability across brain development. *Progress in Brain Research* 207: 3–34. [CrossRef]
- Tchernichovski, Ofer, Partha P. Mitra, Thierry Lints, and Fernando Nottebohm. 2001. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* 291: 2564–69. [CrossRef] [PubMed]
- Tchernichovski, Ofer, Sophie Eisenberg-Edidin, and Erich D. Jarvis. 2021. Balanced imitation sustains song culture in zebra finches. *Nature Communications* 12: 1–14. [CrossRef]
- ten Cate, Carel, and Peter J. Fullagar. 2021. Vocal imitations and production learning by Australian musk ducks (*Biziura lobata*). *Philosophical Transactions of the Royal Society B* 376: 20200243. [CrossRef]
- Theakston, Anna L. 2017. Entrenchment in first language learning. In *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*. Edited by Hans-Jörg Schmid. Washington, DC: American Psychological Association, pp. 315–41. [CrossRef]
- Thorpe, William H. 1958. The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs*. *Ibis* 100: 535–70. [CrossRef]
- Todt, Dietmar, and Nicole Geberzahn. 2003. Age-dependent effects of song exposure: Song crystallization sets a boundary between fast and delayed vocal imitation. *Animal Behaviour* 65: 971–79. [CrossRef]
- Todt, Dietmar, Henrike Hultsch, and Dietmar Heike. 1979. Conditions affecting song acquisition in nightingales (*Luscinia megarhynchos* L.). *Ethology* 51: 23–35. [CrossRef]
- Tomasello, Michael. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Tschida, Katherine, and Richard Mooney. 2012. The role of auditory feedback in vocal learning and maintenance. *Current Opinion in Neurobiology* 22: 320–27. [CrossRef] [PubMed]
- Tumer, Evren C., and Michael S. Brainard. 2007. Performance variability enables adaptive plasticity of ‘crystallized’ adult birdsong. *Nature* 450: 1240–44. [CrossRef] [PubMed]
- Vallentin, Daniela, Georg Kosche, Dina Lipkind, and Michael A. Long. 2016. Inhibition protects acquired song segments during vocal learning in zebra finches. *Science* 351: 267–71. [CrossRef]
- Ventureyra, Valerie A. G., Christophe Pallier, and Hi-Yon Yoo. 2004. The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* 17: 79–91. [CrossRef]
- Vernes, Sonja C., Buddhamas Pralle Kriengwatana, Veronika C. Beeck, Julia Fischer, Peter L. Tyack, Carel Ten Cate, and Vincent M. Janik. 2021. The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B* 376: 20200236. [CrossRef]
- Walters, Keith. 2011. Gendering French in Tunisia: Language ideologies and nationalism. *International Journal of The Sociology of Language* 211: 83–111. [CrossRef]
- Wang, Difei, and James Fawcett. 2012. The perineuronal net and the control of CNS plasticity. *Cell and Tissue Research* 349: 147–60. [CrossRef]

- Warren, Timothy L., Jonathan D. Charlesworth, Evren C. Tumer, and Michael S. Brainard. 2012. Variable sequencing is actively maintained in a well learned motor skill. *Journal of Neuroscience* 32: 15414–25. [[CrossRef](#)]
- Watkins, Michael, Andreia S. Rauber, and Barbara O. Baptista. 2009. *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Cambridge: Cambridge Scholars Publishing.
- Werker, Janet F., and Takao K. Hensch. 2015. Critical periods in speech perception: New directions. *Annual Review of Psychology* 66: 173–96. [[CrossRef](#)]
- West, Meredith J., and Andrew P. King. 1988. Female visual displays affect the development of male song in the cowbird. *Nature* 334: 244–46. [[CrossRef](#)]
- Wiley, R. Haven. 2000. A new sense of the complexities of bird song. *The Auk* 117: 861–68. [[CrossRef](#)]
- Wirthlin, Morgan, Edward F. Chang, Mirjam Knörnschild, Leah A. Krubitzer, Claudio V. Mello, Cory T. Miller, Andreas R. Pfenning, Sonja C. Vernes, Ofer Tchernichovski, and Michael M. Yartsev. 2019. A modular approach to vocal learning: Disentangling the diversity of a complex behavioral trait. *Neuron* 104: 87–99. [[CrossRef](#)]
- Wong, Patrick C. M., Kara Morgan-Short, Marc Ettliger, and Jing Zheng. 2012. Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex* 48: 1091–102. [[CrossRef](#)]
- Woolley, Sarah C., and Sarah M. N. Woolley. 2020. Integrating form and function in the songbird auditory forebrain. In *The Neuroethology of Birdsong*. Edited by Jon T. Sakata, Sarah C. Woolley, Richard R. Fay and Arthur N. Popper. Cham: Springer Nature, pp. 127–55. [[CrossRef](#)]
- Yamaguchi, Ayako. 2001. Sex differences in vocal learning in birds. *Nature* 411: 257–58. [[CrossRef](#)]
- Yanagihara, Shin, and Yoko Yazaki-Sugiyama. 2016. Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. *Nature Communications* 7: 1–11. [[CrossRef](#)] [[PubMed](#)]
- Yazaki-Sugiyama, Yoko, and Richard Mooney. 2004. Sequential learning from multiple tutors and serial retuning of auditory neurons in a brain area important to birdsong learning. *Journal of Neurophysiology* 92: 2771–88. [[CrossRef](#)]
- Yeni-Komshian, Grace H., James E. Flege, and Serena Liu. 2000. Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition* 3: 131–49. [[CrossRef](#)]
- Yeung, H. Henry, and Janet F. Werker. 2009. Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition* 113: 234–43. [[CrossRef](#)]
- Yu, K., W. E. Wood, and F. E. Theunissen. 2020. High-capacity auditory memory for vocal communication in a social songbird. *Science Advances* 6: eabe0440. [[CrossRef](#)]
- Zhang, Xiaopeng, and Chunping Mai. 2018. Effects of entrenchment and preemption in second language learners' acceptance of English denominal verbs. *Applied Psycholinguistics* 39: 413–36. [[CrossRef](#)]