

Review



Cite this article: Carouso-Peck S, Goldstein MH, Fitch WT. 2021 The many functions of vocal learning. *Phil. Trans. R. Soc. B* **376**: 20200235.
<https://doi.org/10.1098/rstb.2020.0235>

Accepted: 1 May 2021

One contribution of 21 to a theme issue ‘Vocal learning in animals and humans’.

Subject Areas:

behaviour, evolution, ecology

Keywords:

vocal learning, communication, vocalizations, functions of learned signals, evolution of learning

Author for correspondence:

Michael H. Goldstein
e-mail: mhg26@cornell.edu

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5514758>.

The many functions of vocal learning

Samantha Carouso-Peck¹, Michael H. Goldstein¹ and W. Tecumseh Fitch²

¹Department of Psychology, Cornell University, Ithaca, NY, USA

²University of Vienna, Vienna, Austria

SC-P, 0000-0002-6084-565X; MHG, 0000-0001-6672-3752; WTF, 0000-0003-1830-0928

The capacity to learn novel vocalizations has evolved convergently in a wide range of species. Courtship songs of male birds or whales are often treated as prototypical examples, implying a sexually selected context for the evolution of this ability. However, functions of learned vocalizations in different species are far more diverse than courtship, spanning a range of socio-positive contexts from individual identification, social cohesion, or advertising pair bonds, as well as agonistic contexts such as territorial defence, deceptive alarm calling or luring prey. Here, we survey the diverse usages and proposed functions of learned novel signals, to build a framework for considering the evolution of vocal learning capacities that extends beyond sexual selection. For each function that can be identified for learned signals, we provide examples of species using unlearned signals to accomplish the same goals. We use such comparisons to generate hypotheses concerning when vocal learning is adaptive, given a particular suite of socio-ecological traits. Finally, we identify areas of uncertainty where improved understanding would allow us to better test these hypotheses. Considering the broad range of potential functions of vocal learning will yield a richer appreciation of its evolution than a narrow focus on a few prototypical species.

This article is part of the theme issue ‘Vocal learning in animals and humans’.

1. Introduction

Vocal production learning is a rare trait, known in only seven animal taxa [1]. The range of species studied as model systems is even more restricted. Although researchers often have one or a few particular ‘paragon’ examples in mind when considering the evolution of vocal learning capacity (e.g. zebra finch song), the range of uses for learned novel vocalizations is surprisingly broad and has evolved in a wide variety of contexts: parental care, signalling group membership, territorial defence, mate attraction and mate bonding. However, so strong is this focus on a few model species that certain clades of vocal learners (e.g. seals or elephants) receive far less attention, and certain functions (e.g. deceptive calling) are typically understudied. Interest in vocal learning in some traditionally lesser-studied species has recently expanded, with over 1000 papers published in the last 10 years mentioning vocal learning across all species of bats, up from only about 160 from the 1990s. However, nearly twice as many papers have been published in the last 10 years on vocal learning in just one finch species (the zebra finch). We fear that focusing investigation primarily on songbirds and humans may discourage exploration of non-canonical vocal functions, and we suggest that discussions of the evolution of vocal learning should keep in mind the full range of possible functions, or risk neglecting the many phylogenetic and adaptive contexts in which this capability has arisen. Our goal here is to encourage a broader discussion of such neglected clades and functions.

For example, predatory northern shrikes mimic the songs of prey bird species, and may use these imitations to lure prey to their deaths [2]. Spectacled parrotlets produce short contact calls that indicate both their individual identity and their group membership, and this call rapidly changes when they switch groups [3,4]. Young dolphins learn an individual-specific ‘signature whistle’ in their first year of life. They appear to model their whistles after those of

conspecifics but modify them into a novel form [5,6]. Male humpback whales in a given region all sing the same song, but gradually pick up variants and innovations, so that this shared song constantly changes over the years [7]. Fork-tailed drongos mimic other species' alarm calls, and use these learned alarms deceptively to steal food [8]. As adults, only male sac-winged bats sing, defending their territory, but young of both sexes imitate their local male's song and learn it through a process reminiscent of human babbling [9]. When deprived of conspecific contact, elephants can learn to imitate the vocalizations of other species, including convincing replications of human speech [10]. What all of these examples have in common is that they involve vocal learning, specifically call production learning—but their functions are surprisingly diverse. Such diversity of functions implies a wider diversity of underlying evolutionary pressures and developmental mechanisms that cannot be accounted for by the limited set of species that currently receive the most research attention.

The examples above provide a taste of the many functions that learned vocalizations can play in communication systems, including parent-offspring identification, signalling group membership, deceptive alarm calling, heterospecific recruitment for mobbing, identifying host species among nest parasites and even prey luring. These span a range from strong shared interests (e.g. among parents and offspring), partially shared interests (mated pairs; close kin) to fully opposed interest (predators and prey, feeding competitors, or territorial rivals). These diverse functions for learned vocalizations have emerged as solutions to particular evolutionary challenges faced by a given species. However, the question remains: when should adaptive signals be learned from signals in the environment (including heterospecifics), rather than predetermined by genetic information? Indeed, many bird species solve the issue of species identification and mate attraction with unlearned songs (including subsocial passerines such as flycatchers), fireflies achieve prey luring with innate flashing displays [11], and many species use unlearned alarm calls. Under what circumstances is producing a learned signal more adaptive than when signals are genetically canalized, given that both causal pathways can achieve the same function? Or is it instead the case that vocal learning has emerged in different clades for particular common functions, and then been co-opted for other, diverse functions over the course of evolutionary time? Answering these questions requires a closer examination of the specific socio-ecological circumstances that lead to complex vocal production learning.

In this review, we survey the many proposed functions of vocal production learning in birds and mammals. We will adopt a rather restrictive definition of vocal production learning, namely the capacity to learn vocalizations outside of the reliably developing or 'innate' vocal repertoire ('vocal production learning' or VPL hereafter). Thus, modifications of species-typical calls, as seen in many birds and mammals, will be left aside, as will perceptual and cognitive capacities to learn to recognize or interpret novel sounds or vocalizations. The reason for this 'high bar' is that even given this stringent definition, VPL has evolved convergently in at least seven clades, and appears in thousands of distinct species (most of them little studied).

By contrast, we will be permissive with regard to proposed functions, because, in many cases, only one or a few

species appears to make use of learned vocalizations in a particular manner, and the evidence for a putative function could often use improvement. For example, the imitation of prey species vocalizations by a predator as a 'lure', is firmly known only for humans, and hypothesized in one songbird species (northern shrikes [2]). We hope to inspire further work on some of these less-studied functions and species. Previous reviews of the functions of vocal learning have endeavoured to trace the evolution of VPL in birds [12] or mammals [13], and focused on how learning may affect the generation, usage and comprehension of sounds. By contrast, our goal is to interpret VPL in an ecological and evolutionary framework, across both birds and mammals, identifying putative socio-ecological traits that co-occur in species which employ learned vocalizations.

The paper is structured as follows. We start by briefly summarizing results from a selection of bird and mammal species that, we hope, concisely illustrates the wide range of uses to which vocally learned signals can be put. We also provide examples where this function is achieved *without* vocal learning (figure 1), to illustrate the multiple evolutionary solutions to the demands of various functions, raising the question of under which circumstances is a learned vocalization more adaptive than an innate, less flexible one. We will then explore in more detail the functions of VPL in two well-studied clades, parrots and human infants. Finally, in the third part of the paper, we examine the multiple proposed functions of vocal learning from a sociobiological and ecological viewpoint, hoping to gain insight into the conditions under which particular functions are likely to evolve. We end by proposing several hypotheses, making suggestions for future research, and pointing out lacunae in current knowledge, concluding that a broader appreciation of the many functions of VPL should inform future research into its evolution and mechanistic basis.

2. Selected case studies for vocal learning

We will start with several vignettes illustrating the wide variety of functions known (or proposed) for learned vocalizations, emphasizing those that deviate from the standard male songbird model.

(a) Deception

Human hunters have used vocal mimicry to attract prey for millennia [14,15]. It has long been suggested that some birds might also mimic prey vocalizations to attract them [16–18], but no conclusive evidence demonstrates this (cf. [19]). However, northern shrikes (*Lanius excubitor*) are generalist predators that often prey on smaller birds [20]. Both sexes sing through the year for the functions of attracting mates, facilitating pair formation and defending territory, and (like many shrikes) their songs include clear examples of vocal mimicry (heterospecific imitation). Playback of shrike song attracted small passerine birds significantly more than control song (American robin) or silence [2]. This is consistent with the hypothesis that northern shrikes have evolved a novel function for vocal learning—to lure avian prey via mimicry—but more research is clearly needed to support this hypothesis. By contrast, multiple predatory species use innate anatomical adaptations to lure prey (e.g. anglerfish [21]), and margays are reputed to produce

| | function | subtype | VPL species | non-VPL species | |
|-------------------|------------------------|--------------------------|--|--|--|
| shared interests | caring | parent-offspring contact | greater and lesser spear-nosed bats, bottlenose dolphin, orange-fronted conure | cow, goat, western sandpiper, razorbill | |
| | | begging | bottlenose dolphin, Horsfield's bronze cuckoo, zebra finch | suricate, banded monogoose, pig, pigeon | |
| | | babbling | greater sac-winged bat, green-rumped parrotlet, long-billed hermit, zebra finch, Australian magpie | pygmy marmoset, Japanese quail | |
| | group cohesion | contact/cohesion | bottlenose dolphin, orca, budgerigar, rufous-breasted hermit, yellow-naped amazon, kea, red crossbill, greater spear-nosed bat, green-rumped parrotlet | African wild dog, whooper and Bewick's swans, pygmy marmoset | |
| | | mate attraction | walrus, greater sac-winged bat, budgerigar, long-billed hermit, northern mockingbird | koala, smooth guardian frog, black coucal cicada | |
| | individual recognition | embryonic passwords | superb fairy wren | n/a | |
| | | signature calls | green-rumped parrotlet, orange-fronted conure, bottlenose dolphin, spectacled parrot | white-nosed coati, spotted hyena | |
| | information transfer | object referential | alarm calls | African grey parrot, bottlenose dolphin | vervet monkey, Gunnison's prairie dog |
| | | | mobbing | (only usage/association learning known). Fork-tailed drongo, brown thornbill, Sri lankan magpie, black-capped chickadee, noisy miner | Gunnison's prairie dog, vervet monkey, suricate, Belding's ground squirrel |
| | | lekking | coalition | racket-tailed drongo, black-capped chickadee, phainopepla, noisy miner, Steller's jay, tits | common marmoset, California ground squirrel, terns and gulls, black phoebe |
| opposed interests | territory | duetting | humpback whale, long-billed hermit hummingbird, dusky bush-tanager | turkey, long-tailed manakin | |
| | | advertisement/defence | orange-chinned parakeet, rufous-and-white wren, peach-faced lovebird, yellow-naped amazon | titi monkeys, barred owl, siamang, Alston's singing mouse, barred antshrike | |
| | prey-luring | prey-luring | sac-winged bat, palm cockatoo, sombre hummingbird, vasa parrot, song sparrow, eclectus parrot | rock hyrax, eastern phoebe, red deer, orangutan | |
| | agonistic | prey-luring | northern shrike | margay | |
| | | conflict/aggression | elephant seal, Egyptian fruit bat, sombre hummingbird, humpback whale, many songbirds | red squirrel, neotropical singing mice, antbird, wolf | |

Figure 1. Examples of varying functions for learned and unlearned vocal signals across species. Functions are divided into shared interests, in which signaller and receiver mutually benefit from the signal, and opposed interests, in which the signaller and receiver are in conflict. These functions are further subdivided into subtypes. For each subtype, we provide examples of vocal production learning (VPL) species which employ vocal signals for that function, as well as non-VPL species as examples of achieving the same function without learning. For references and scientific species names, see the electronic supplementary material, table S1.

calls that lure in tamarin prey [22], but are not known to be vocal learners.

Another antagonistic use of vocal mimicry is provided by the fork-tailed drongo (*Dicrurus adsimilis*). Several species respond to drongos' alarm calls, which often warn of the

presence of a predator. However, drongos in the Kalahari desert use alarm calls in the absence of a predator to cause competitors to flee and abandon food, which they eat, obtaining up to 23% of their food this way, and using both their own species-typical alarms and mimicking alarm calls

of 51 other species [8]. Although other species use their species-typical alarm calls deceptively to obtain food [23], competitors tend to habituate to the same call, while varied alarms maintain their salience [8]. However, the mimicry abilities of drongos can also be used for non-deceptive functions. Racket-tailed drongos (*Dicrurus paradiseus*) mimic both mobbing and alarm calls of other species regularly [24,25]. They produce mobbing calls honestly, in the presence of predators, and incorporation of heterospecific mobbing calls induces more intense mobbing by the mimicked species. Although it is not uncommon for a species to respond to heterospecific mobbing calls [26,27], playbacks showed that drongo-specific mobbing calls were less effective at recruiting heterospecific individuals than calls including mimicry [25]. Because both callers and mobbing individuals jointly benefit from this predator-deterrent behaviour, such mobbing recruitment involves shared interests (unlike deceptive fork-tailed drongo alarm calling). However, deterrence functions can also be accomplished by unlearned sounds, e.g. nestling burrowing owls produce an unlearned call that mimics rattlesnakes and may deter nest-burrow predators or competitors [28]. Similarly, nestling northern flickers (*Colpates auratus*) produce a buzzing sound, reminiscent of a beehive, potentially deterring predatory squirrels [29].

Another fascinating case of learning heterospecific vocalizations is provided by the indigobird *Vidua chalybeata*. Indigobirds belong to a nest-parasitic genus, in which each of the visually similar species has a preferred host [30,31]. Male indigobirds mimic their host father's song as adults, and females imprint on their host father's song, allowing the females to selectively mate with males that sing the song of their own host species, despite the close visual resemblance among host *Vidua* species [32,33]. Although many species use unlearned calls or songs for species identification, this unusual usage of vocal learning allows assortative mating, benefiting both male singers and female listeners, but may also pave the way for flexible exploitation of new host species [31].

(b) Individual and group identification

Vocal learning occurs in multiple bat species, where learned calls often seem to play a role in individual or group identification. The screech calls of adult greater spear-nosed bats *Phyllostomus hastatus* are socially modified to encode a group signature [34], while in the lesser spear-nosed bat *Phyllostomus discolor*, isolation calls of pups converge to resemble calls of their mothers but pups raised in isolation do not acquire normal calls [35,36]. There also appear to be group dialects in this species [37], which has recently shown to have volitional control over vocalizations, changing specific acoustic parameters in an operant situation [38]. Finally, in the sac-winged bat *Saccopteryx bilineata*, pup calls from both sexes converge to match those of their resident male's territorial 'songs' [9], again suggesting a group identification function.

Similarly, orcas (*Orcinus orca*) live in tight-knit groups called 'pods' which associate with higher-level 'clans'. Orcas off the coast of British Columbia are often highly vocal, and within each pod, there is a substantial sharing of the entire vocal repertoire; furthermore, there is less but still significant sharing at the clan level [39–41]. Both pods and clans represent matrilineal and thus kin groups and the

shared repertoire may serve to behaviourally identify the pod and clan affiliation of the caller [39]. Although evidence for vocal learning of these wild dialects remains circumstantial, recent experiments clearly demonstrate a well-developed capacity for vocal learning in orcas [42]. Together, experimental and field data suggest that, as for bottlenose dolphins [5]—but contrasting sharply with baleen whales—odontocete vocal learning is well-developed in both sexes.

(c) Attraction and contact

Like most baleen whales, male humpback whales (*Megaptera novaeangliae*) produce display vocalizations, both on their breeding grounds and while migrating to and from them. Mysticete vocal displays range from simple and highly repetitive calls in fin or blue whales [43,44] to quite complex songs in bowheads [45] and humpbacks [46,47]. Migrating male humpbacks typically begin to sing when they have encountered a female and sing for longer durations when a female is nearby, suggesting a courtship or mate-attraction function of song [48]. Singing humpback males may, therefore, constitute a 'floating lek' that attracts females to the mating area, potentially stimulating female receptivity [49]. However, males are aggressive towards other singing males and will avoid producing song if another competitive male is close by, suggesting an additional function for intra-sexual selection [50].

Within mated pairs, many taxa engage in vocal duetting, but the function of duets remains debated, and probably varies both between and within species [51]. Duetting can be directed at conspecifics, to declare joint territory, or used for mate-guarding by making the pair bond evident to third parties, as in red-backed fairy wren males that duet to deter extra-pair copulations [52]. Duetting can also function within the pair, to keep mates coordinated or in contact, as with black-bellied wrens that duet during territorial disputes to prevent inadvertent intrapair aggression [53]. Female happy wrens (*Pheugopedius felix*) duet with their mate to signal pair commitment, but sing overlapping songs to signal aggression towards intruders [54]. Finally, male slate-colored boubous (*Laniarius funebris*) use multiple distinct duet song types, each of which serves a distinct function: reaching breeding synchrony with their mates, defending territory or mate-guarding [55]. However, species lacking VPL achieve similar functions by duetting, including titi monkeys (Pitheciidae: Callicebinae; [56,57]) and Alston's singing mice (*Scotinomys teguina*; [58]), demonstrating that VPL is not necessary for vocal synchronization, and again raising questions about the circumstances under which the flexibility of learning is adaptive.

3. Models of vocal learning

Birdsong is a useful model for vocal learning and language development in humans owing to similarities at the neurological level and similar developmental trajectories (including a 'babbling' subsong phase; [59,60]). A recent survey found that roughly half of songbirds are closed-ended learners, only able to learn new vocal forms for a brief developmental period [61], while the other half continue to show vocal learning into adulthood (like humans). In contrast to the enormous learned vocabularies of adult humans, many songbirds acquire only a small repertoire: about 80% of

songbirds learn fewer than six songs in their lifetime [62]. In roughly 30% of songbirds, learning to sing is a male-only ability [63], and in species in which both sexes sing, females typically have less expansive or complex repertoires than males [64]. Thus, songbirds range from species where only males sing and learn a fixed song (or a few songs) early in life, to those with large repertoires, such as the brown thrasher (which sing an excess of 1100 song types [65,66]), or where both sexes retain the VPL capacity into adulthood. In some songbirds, but not all, vocal learning and production does not generally resemble that of humans, which have open-ended, lifelong vocal learning in both sexes, leading to vast vocal repertoires [67].

We will now discuss two VPL taxa whose mechanisms of learning have received less attention than songbirds: parrots and prelinguistic human infants. These taxa have similar socio-ecological traits, and a similarly diverse suite of functions for learned signals. Along with other complex vocal learners such as elephants and toothed whales, both of these groups (i) are highly social, with fission–fusion grouping patterns, and use learned vocalizations in a wide range of social contexts, (ii) have extended juvenile periods and long life-spans, characterized by prolonged associations between the offspring and at least one parent, (iii) use vocal learning in a sexually monomorphic fashion, with both males and females learning and vocalizing similarly, (iv) show some evidence of offspring ‘babbling’ behaviour, and facilitation of vocal learning by parents, (v) are open-ended vocal learners, continuing to acquire and modify vocalizations throughout their lives, and (vi) some show evidence of using some form of signature call that identifies individuals or groups.

4. The case of parrots

In contrast to many passerines, whose songs are limited in function to mate attraction and territory defence, most parrots use their learned vocalizations for a wide range of functions across multiple social contexts. Parrots also, of course, use these vocalizations for the purposes of sexual selection, but given that the predominant mating systems of songbirds and parrots are highly similar (tending towards solitary nesting and social monogamy [68]), the function of sexual selection alone cannot explain the highly flexible vocal learning capabilities of parrots. While most songbirds defend large breeding territories using their song, no parrot species is known to do so [69,70]. Because most parrots defend only the area immediately around the nest, as a pair, the function of territory defence also seems a poor explanation for their sophisticated vocal learning capacities. Suboscine songbirds achieve both mate attraction/pair maintenance and territory defence without the need for learning (e.g. eastern phoebe (*Sayornis phoebe*) [71]). The complex social system of parrots may play a strong role in the functions for which parrots employ learned signals.

Parrots’ highly social, fission–fusion foraging culture is a result of their uniquely challenging diet of toxic, unripe seed, driven by four strategies: (i) physiological detoxification by gut microbiota, necessitating rest in communal roosts after each foraging period [72]; (ii) geophagy (consuming clay-rich soil to absorb toxins), making songbird-like territories unfeasible, and requiring information transfer between individuals with knowledge of clay sources and naive birds

[69,73]; (iii) eating a mixed diet of varying toxicity types and levels, again requiring information transfer of foraging sites between individuals; and (iv) neophobia, whereby cautious individuals learn safe food choices from experienced individuals [68]. Social learning, thus, mitigates the challenges of geophagy, mixed diet and neophobia, and is partially synergistic with the obligate social roosting time required by detoxification. This social-learning-dependent system favours not only high sociality within a flock, but interacting with and gaining foraging information from as many individuals as possible, leading to flocks with overlapping foraging ranges and frequent, selective exchange of individuals with differing information. Such social complexity requires a robust system for individual recognition, which parrots appear to accomplish largely through distinctive calls [3,74,75]. Note that a similar mixed toxic diet strategy is used by non-VPL birds such as the hoatzin (*Opisthocomus hoazin*), which has small, stable social groups [76].

What level of referential specificity characterizes parrot communicative systems? In laboratory settings, parrots represent one of the few taxa in which referential signalling has been demonstrated. Alex, the famous African grey parrot (*Psittacus erithacus*), was capable of using learned English words to identify and request objects of particular colours, numbers, shapes and materials [77]. This shows that at least some parrots are capable of referential signalling, but it remains unclear whether wild parrots use such signals to exchange foraging information. Functionally referential alarm calls have been documented in a number of species with complex social structures (electronic supplementary material, table S1), but there is no clear evidence that parrot alarm calls are referential, or even that they are learned [78]. The laboratory tasks that demonstrate this surprising cognitive sophistication may exploit capacities underlying the use of flexible vocal signals to keep track of individuals in a complex and dynamic social environment.

Budgerigars show call convergence in the laboratory [75,79,80], but there is surprisingly limited evidence of flock-level vocal convergence in the wild [81], and convergent versus divergent responses may vary among the sexes [82]. With unstable flocks using fission–fusion dynamics, vocal convergence at the level of the flock may not be advantageous for many parrots. However, convergence at the individual level is adaptive for several species, where individuals temporarily converge on the call of another individual in order to selectively attract its attention, or ‘address’ it directly [68]. Playback experiments in wild orange-fronted conures found that individuals modify their contact calls to be more similar to playbacks [83], and that birds played calls similar to their own signature call responded at higher rates [84]. When playbacks did not converge with their own, subjects responded with agonistic calls [83]. These findings are consistent with the hypothesis that call convergence is used to address individuals affiliatively, a function particularly valuable in species with complex fission–fusion social structures. Whether vocal learning evolved to enable call convergence, or more likely existed previously for another function and was co-opted for this purpose, requires further investigation.

Two further potential examples of referential signals for individuals in parrots include captive spectacled parrotlets (*Forpus conspicillatus*), which produce different contact calls depending on the individual with whom they are interacting [4]. Although this could indicate referential ‘naming’ of

individuals or simply the morphing or converging of call structures, parrots appear to possess and use the ability to dynamically alter their learned vocalizations to address specific individuals, either by 'name' or via call structure matching, in order to coordinate group activities such as foraging. Furthermore, wild Venezuelan green-rumped parrotlet nestlings (*Forpus passerines*) develop individual-specific contact calls, but with certain signature attributes learned from their parents [85]. These calls develop during ontogeny, beginning with short, non-frequency modulated utterances, and shifting over time into mature, adult-like, modulated contact calls, influenced by the vocalizations of the parents [86], in a manner reminiscent of human vocal development from babbling to speech.

5. The case of human infants

The prolonged period of immaturity that characterizes vocal development in human infants affords unique opportunities for vocal learning and gives rise to a diverse array of functions. Over the first year, the sound patterns of the ambient language are incorporated into infants' prelinguistic vocal repertoire [87–89] and large individual differences emerge in the rate of vocal development. Data from deaf and hearing-impaired infants reveal that their babbling is acoustically different from that of hearing infants from six months onwards, with abnormalities in early syllables and a marked delay in onset of acoustically mature (canonical) syllables [90–95]. These studies suggest that typical vocal development requires auditory input. Indeed, infants can rapidly learn new prelinguistic vocal forms based on social feedback contingent on their babbling [96,97].

Human vocal learning clearly functions eventually to acquire the words and sound patterns of the child's native language. But what functions are served by prelinguistic vocal learning? In the most general terms, the plasticity inherent in early vocal development allows for the production of a large and flexible repertoire of sounds that serve as the foundation of words. Prelinguistic vocalizations are highly plastic, often including energetically demanding, loud sounds, such as squeals and growls [98] that do not play a role in later language but are a form of vocal play and exploration. Such plasticity allows infants to explore the acoustic space of vocalizing. Babbling is contextually free, in that the sounds are not bound to specific states or objects [98–100]. Although prelinguistic vocalizations share few features of adult language, their contextual freedom allows for large modifications in form and function that facilitates the later emergence of spoken language. These modifications occur to a large degree as a result of environmental input, including social feedback, so that the pluripotency of early vocalizing becomes canalized to the sounds of the ambient language.

Specific functions of prelinguistic vocalizing emerge at different times over the first year. Infants first learn about the range of sounds they can produce, as well as their social effects, and later learn how to produce specific sounds in ways that are tied to specific social interactions and objects. Early vocal development (zero to six months) is first characterized by *contextual learning* [101], in which babbling becomes associated with contingent and positive responses from the social environment. As a result, infants build an expectation that their vocalizations will get social attention. Such learning

is evidenced by the reaction of infants when social responses are withheld. In face-to-face interactions with adults, five-month-olds temporarily increase their amount of babbling during a 'still-face' manipulation, when adult social responses cease [102,103]. This increase owing to the absence of adult reactivity indicates an 'extinction burst' [104], the result of a learned association between vocalizing and social responding.

Learning the social functions of vocalizing sets the stage for more advanced vocal learning in the second half of the first year. Vocal development from 6 to 12 months is now characterized by *production learning* [101], in which infants modify the acoustic structure of their babbling as a function of social feedback. Infants rapidly learn new patterns of vocal production from carer speech that is temporally contingent on their sounds [96,97]. When mothers were instructed to respond to their infant's vocalizations with specific speech patterns, infants learned to produce phonological patterns that significantly resembled those of their mothers' vocal responses. However, infant phonemes were not the same as those of their mothers' utterances, suggesting that phonological learning was not imitation, but rather a form of phonological pattern matching, facilitated by comparing their own to their mother's speech when these occurred in close temporal proximity [97]. Infants who receive adult responses for their vocalizing tend to continue to produce speech-like vocalizations [105].

Such a strong role of social interaction illustrates another function of prelinguistic vocal learning, that of facilitating attention, caring and feedback from carers [106]. Carers responses to immature speech tend to be appropriately structured and temporally coordinated with the infants' utterances [107,108]. Vocalizations that are more speech-like are more likely to receive responses from adults [105,109]. These responses contain information that is useful for the learning of more advanced forms of vocal communication. For example, adults simplify the linguistic structure of their speech when responding to babbling, creating a favourable environment for learning language [110]. Infants who hear object names in response to their object-directed vocalizing rapidly learned the names for the objects [111]. Thus, the social environment of a vocalizing infant is a source of rich structure that can guide advances in vocal learning. Immature human vocalizing thus functions to elicit carers social responses, creating opportunities for language learning.

6. Discussion

As we have shown, VPL has many functions and can serve as a solution to diverse communicative challenges, but researchers typically study only a handful. Biologists should, thus, cast a wider net when considering the adaptive functions of vocal learning abilities. We have reviewed a number of communicative functions used by VPL species, but many of these functions are also accomplished by non-learners. While parrots are capable of object-referential vocalizations using VPL, vervet monkeys produce referential alarm calls without it. Oscine songbirds learn a complex song to attract a mate, but the unlearned songs of frogs and cicadas accomplish the same function. For the majority of uses for vocalizations in VPL species we could identify, we also found examples of non-VPL species accomplishing the same vocal functions, as illustrated in figure 1.

Organizing VPL and non-VPL species according to the functions of their vocalizations provides insight into the ecological factors giving rise to VPL. This comparative approach creates a richer dataset than considering vocal production learners alone, and allows us to ask more powerful and precise questions about how and why VPL evolved, and/or was co-opted for new functions once it emerged. Under what circumstances is learning to produce a vocal signal more advantageous for achieving its adaptive function(s)? We suggest that for VPL to evolve in a given species requires as prerequisites both the necessary cognitive capacity, and the developmental opportunity (often, prolonged access to conspecific ‘tutors’ granted either by an extended childhood with parental care or high sociality throughout the lifespan), but also some adaptive advantage owing to vocal learning for enhancing communicative functions.

Vocal learning may confer an adaptive advantage in multiple ways. We propose three specific ways in which vocal learning can provide an adaptive advantage. First, the ‘more is better’ hypothesis posits that vocal learning increases repertoire size for a particular class of signals (songs, alarm calls, etc.), therefore increasing the efficacy of that signal type [112,113]. The mechanism underlying this posited increase in efficacy could be avoidance of habituation in listeners and/or increased signal range to reach a wider variety of perceivers. This hypothesis seems applicable to many songbird song repertoires, predicting that mimics (who imitate heterospecific song) and open-ended learners should have larger song repertoires. It applies to human language, because a large shared vocabulary of words enables enhanced communication. It also applies to the drongo species that mimic alarm and mobbing calls discussed above, or the shrike ‘prey luring’ example, as the putative goal of these calls is to frighten or attract a large number of heterospecific individuals. Finally, in dolphins or parrots that use calls as vocal labels, a larger call repertoire could target a larger group of recipients. The prediction that vocal learning increases repertoire size has recently been verified in a phylogenetically controlled analysis in open-versus closed-ended vocal learning songbirds [61], consistent with the ‘more is better’ hypothesis for open-ended learners. The ecological pressure for vocal learning of complex repertoires may originate in the demands of complex social structures [114]. For example, chickadees housed in large social groups acquired more complex repertoires, carrying more information owing to a greater diversity of note types and combinations than those housed in smaller groups [115].

However, some vocal learners do not seem to fulfil these predictions. Budgerigars, caciques, humpbacks and orcas all tend to produce distinctive group signatures or repertoires rather than using vocal learning to increase repertoire size. These examples suggest a second hypothesis, the ‘flexible pruning’ hypothesis, whereby vocal learning provides an enlarged *potential* repertoire, which is then ‘pruned’ down to a much smaller set of the most appropriate or efficacious signal(s), as in action-based learning in songbirds [116]. By this hypothesis, the goal is not simply more vocalizations, but better (or more flexible) vocalizations, and vocal learning combined with some selective process provides a means to this end. This hypothesis predicts a large potential vocabulary (which can be tested with tape tutoring or cross-fostering), followed by selective attrition down to a smaller set (observable during development). This smaller set may be determined by feedback from listeners (e.g. female cowbirds [117], or bronze

cuckoo host parents [118]), be copied from successful individuals (e.g. humpback whales [119]), or be ‘negotiated’ via mutual imitation (as for budgerigars [75]). In open-ended learners such as canaries, flexible pruning may occur seasonally as the birds continually refine their repertoires [120]. Humpback whales also continually refine their repertoires, occasionally in dramatic fashion as eastern Australian humpbacks completely replace their song every few years [121]. Flexibility via over-generation and pruning also applies to human language, because the phonological rules underlying word formation generate vastly more potential words than are actually used in the language (cf. [122]).

Finally, the ability of species to refine vocal repertoires via social learning creates an opportunity for ‘cultural refinement’ to arise, potentially accelerating vocal adaptation to local environmental conditions by superimposing the possibility for ‘cultural evolution’ over ordinary natural and sexual selection and individual learning. Our third hypothesis, the ‘cultural refinement’ hypothesis, is that vocal learning specifically evolves to create the potential for cumulative vocal culture. To the extent that vocal learning enables individuals to copy call variants that are better suited to their current context (whether owing to environmental transmission, differing local competitors or changing mate preferences), individuals capable of vocal learning would show an advantage over those who rely solely on inborn vocalizations, or who must generate and select suitable vocal variants themselves (cf. [123]). This hypothesis treats vocal learning as a subtype of social learning [124,125], and sees vocally learned population-level vocal repertoires (dialects) as prototypical examples of ‘animal culture’ [119,126]. Clear examples of rapid, and apparently functionally adaptive, cultural transmission are of course rampant in humans [127], but cultural evolution across time is also well-attested in humpback whales [7], sperm whales [128] and in the songs of savannah sparrows [129] and white-throated sparrows [130]. For species like drongos which live in mixed flocks and extensively use heterospecific calls [8,25], the value of learning over inheritance seems clear. Testing this hypothesis is challenging owing to difficulties quantifying ‘well-suitedness’ and of measuring the speed of vocal adaptation by natural selection versus cultural evolution. However, the proposal could be tested in particular domains where some independent measure of goodness exists; for example, in the domain of environmental transmission, it predicts that vocal learner’s songs should be better suited to their local environments (e.g. by propagating further or resisting noise better) than unlearned calls or songs (cf. [131]). Evidence of rapid changes in songs in response to urban noise [132,133] are consistent with this hypothesis. The local social environment can similarly influence and be influenced by cultural refinement, as is the case of sperm whales, whose culturally transmitted dialects are a marker of social group and serve to segregate individuals into social units [134,135].

These three hypotheses are not mutually exclusive. For example, in songbirds, if longer or more elaborate songs more effectively attract mates, this can result in the evolution of highly complex displays under either the ‘more is better’ or the ‘flexible pruning’ hypotheses. Our hypotheses and case studies suggest a common theme: vocal learning is most likely, and most useful, when the function of the learned signal requires plasticity and social input. This is especially

the case in species with complex social systems, in which an individual may need to navigate a constantly changing group of social partners, individually recognize and locate familiar and unfamiliar conspecifics, form coalitions and coordinate activities with them, and exchange information about patchy resources. If so, why has VPL not evolved in other highly social species, such as ungulates, seabirds and galliformes, canines, mongooses and non-human primates? Learning vocalizations can be a risky strategy: along with the potential to learn to produce one's signals comes the potential to learn them incorrectly, jeopardizing chances of reproduction and survival. An innately specified, unalterable vocalization seems a safer prospect in many cases.

The capacity for vocal learning also comes with developmental costs owing to the need for increased cognitive capacity, and opportunity costs owing to the time and social feedback required to learn. Many vocal learners share lengthy childhoods, which provides the time needed to learn to vocalize from adults; however, prolonging the time devoted to offspring care is costly to parents [136,137]. Thus, one important potential constraint on VPL may be the requirement for a prolonged learning period, and a concomitant delay in becoming communicatively and reproductively competent. This may help explain why roughly half of songbird species show close-ended learning during a short sensitive period, after which they are relatively inflexible (thus minimizing this cost [61]). Furthermore, it may restrict the evolution of open-ended vocal learning to relatively long-lived species. However, this still does not fully explain the lack of vocal learning in certain species: orangutans have the longest period to parental independence of any non-human animal (up to 8 years), but evidence for VPL in orangutans remains very limited [138,139]. Orangutans are mostly solitary, suggesting that large social groups may also be crucial. Environmental factors may also play an important role in the adaptive advantage of vocal flexibility. For example, flexibly learned individual identity calls may be a necessity for parrots and cetaceans as a result of the habitats in which they live. Using innate physiological voice cues for individual identification, as some social species do, may be unreliable when sound is affected by distance and density of intervening foliage, or by water depth and pressure. In a complex social system or environment where vocal cues of individual or group identity need to be maximized, VPL may be highly advantageous [140,141].

References

1. Jarvis ED. 2007 Neural systems for vocal learning in birds and humans: a synopsis. *J. Ornithol.* **148**, S35–S44. (doi:10.1007/s10336-007-0243-0)
2. Atkinson EC. 1997 Singing for your supper: acoustic luring of avian prey by northern shrikes. *Condor* **99**, 203–206. (doi:10.2307/1370239)
3. Wanker R, Fischer J. 2001 Intra- and interindividual variation in the contact calls of spectacled parrotlets (*Forpus conspicillatus*). *Behavior* **138**, 709–726. (doi:10.1163/156853901752233361)
4. Wanker R, Sugama Y, Prinaige S. 2005 Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. *Anim. Behav.* **70**, 111–118. (doi:10.1016/j.anbehav.2004.09.022)
5. Sayigh LS, Tyack PL, Wells RS, Scott MD. 1990 Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: stability and mother-offspring comparisons. *Behav. Ecol. Sociobiol.* **26**, 247–260. (doi:10.1007/BF00178318)
6. Janik VM, Sayigh LS. 2013 Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **199**, 479–489. (doi:10.1007/s00359-013-0817-7)
7. Garland EC, Gedamke J, Rekdahl ML, Noad MJ, Garrigue C, Gales N. 2013 Humpback whale song on the southern ocean feeding grounds: implications for cultural transmission. *PLoS ONE* **8**, e79422. (doi:10.1371/journal.pone.0079422)
8. Flower TP, Gribble M, Ridley AR. 2014 Deception by flexible alarm mimicry in an African bird. *Science* **344**, 513–516. (doi:10.1126/science.1249723)
9. Knörnschild M, Nagy M, Metz M, Mayer F, von Helversen O. 2010 Complex vocal imitation during ontogeny in a bat. *Biol. Lett.* **6**, 156–159. (doi:10.1098/rsbl.2009.0685)
10. Stoeger AS, Mietchen D, Oh S, de Silva S, Herbst CT, Kwon S, Fitch WT. 2012 An Asian elephant imitates human speech. *Curr. Biol.* **22**, 2144–2148. (doi:10.1016/j.cub.2012.09.022)

7. Conclusion

What have we learned about the evolution of vocal learning, and its phylogenetic occurrence, from our focus on diversity of functions? First, we gain greater insight into the multiple functions of learned signals. In many cases, the same vocalization may serve multiple functions simultaneously: birdsong often functions in both territorial defence and mate attraction and also plays a role in individual recognition. But learned songs can also serve as a 'password' for group membership (caciques [142]) or an indicator of family membership (grosbeaks [143]), or play a role in creating and maintaining pair bonds (in many duetting species). It is probably futile to seek 'the' function of song, and more useful to explore the range of contexts and putative functions of song, whether learned or not. Second, we can better generate and test evolutionary hypotheses by comparing known vocal learners (e.g. northern shrikes) with specific non-vocal-learning clades (e.g. margay), where we often have limited understanding of the function of vocal signals. Investigating the evolutionary circumstances by which signals accomplish similar functions with or without learning (e.g. deceptive vocal mimicry to deter heterospecifics is learned in drongos, but not nestling burrowing owls), would help shed light on the adaptive value of vocal plasticity. Finally, a broader focus beyond a few prototypical species and the typically studied communicative functions—mate attraction and territory defence—may reveal unexpected species as vocal learners. Elephants, seals and killer whales were only discovered to be vocal learners after mimicking heterospecifics, following rare instances of a prolonged period in captivity and isolation from conspecifics [10,144–146]. A broader focus on the many functions of, and socio-ecological traits associated with, vocal learning, may spur the discovery of new species of vocal learners, and help us gain a richer and more complete understanding of the evolutionary forces underlying this unique capacity.

Data accessibility. This article has no additional data.

Authors' contributions. All authors contributed equally to conceiving of, drafting and editing this manuscript. All authors gave final approval for publication.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

11. Lloyd JE. 1986 Firefly communication and deception: 'Oh, what a tangled web'. In *Deception: perspectives on human and nonhuman deceit* (eds RW Mitchell, NS Thompson), pp. 113–128. New York, NY: SUNY Press.
12. Nottebohm F. 1972 The origins of vocal learning. *Am. Nat.* **106**, 116–140. (doi:10.1086/282756)
13. Janik V, Slater P. 1997 Vocal learning in mammals. *Adv. Study Behav.* **26**, 59–99. (doi:10.1016/S0065-3454(08)60377-0)
14. Frederick II HRE. 1194–1250. 1943 (c. 1248) *The art of falconry, being the De arte venandi cum avibus*, p. 637. Stanford, CA: Stanford University Press.
15. Lewis J. 2009 As well as words: Congo Pygmy hunting, mimicry, and play. In *The cradle of language* (eds RP Botha, C Knight), pp. 236–256. Oxford, UK: Oxford University Press.
16. Armstrong EA. 1973 *A study of bird song*, p. 343. New York, NY: Dover.
17. Baylis JR. 1982 Avian vocal mimicry: its function and evolution. In *Acoustic communication in birds* (eds DE Kroodsmma, EH Miller), pp. 51–80. New York, NY: Academic Press.
18. Pollard J. 1930 Whisper song. *Emu* **30**, 62–63. (doi:10.1071/MU930062)
19. Dalziell AH, Welbergen JA, Igc B, Magrath RD. 2014 Avian vocal mimicry: a unified conceptual framework. *Biol. Rev.* **90**, 643–668. (doi:10.1111/brv.12129)
20. Lorek G, Tryjanowski P, Lorek J. 2000 Birds as prey of the great grey shrike (*Lanius excubitor*). *Ring* **22**, 37–44.
21. Pietsch TW, Grobecker DB. 1978 The compleat angler: aggressive mimicry in an antennariid anglerfish. *Science* **201**, 369–370. (doi:10.1126/science.201.4353.369)
22. Calleia F, Rohe F, Gordo M. 2009 Hunting strategy of the margay (*Leopardus wiedii*) to attract the wild pied tamarin (*Saguinus bicolor*). *Neotrop. Primates* **16**, 32–34. (doi:10.1896/044.016.0107)
23. Munn CA. 1986 Birds that 'cry wolf'. *Nature* **319**, 143–145. (doi:10.1038/319143a0)
24. Goodale E, Kotagama SW. 2006 Context-dependent vocal mimicry in a passerine bird. *Proc. R. Soc. B* **273**, 875–880. (doi:10.1098/rspb.2005.3392)
25. Goodale E, Ratnayake CP, Kotagama SW. 2014 Vocal mimicry of alarm-associated sounds by a drongo elicits flee and mobbing responses from other species that participate in mixed-species bird flocks. *Ethology* **120**, 266–274. (doi:10.1111/eth.12202)
26. Caro T. 2005 *Antipredator defenses in birds and mammals*. Chicago, IL: Chicago University Press.
27. Curio E. 1978 The adaptive significance of avian mobbing I: teleonomic hypotheses and predictions. *Z. Tierpsychol.* **48**, 175–183.
28. Owings DH, Rowe MP, Rundus AS. 2002 The rattling sound of rattlesnakes (*Crotalus viridis*) as a communicative resource for ground squirrels (*Spermophilus beecheyi*) and burrowing owls (*Athene cunicularia*). *J. Comp. Psychol.* **116**, 197–205. (doi:10.1037/0735-7036.116.2.197)
29. Wiebe KL, Moore WS. 2008 Northern flicker (*Colaptes auratus*). In *The Birds of North America Online* (ed. A Poole). Ithaca, NY: Cornell Lab of Ornithology.
30. Payne RB. 1973 Vocal mimicry of the paradise whydahs (*Vidua*) and response of female whydahs to the songs of their hosts (*Pytilia*) and their mimics. *Anim. Behav.* **21**, 762–771. (doi:10.1016/S0003-3472(73)80102-2)
31. Sorenson MD, Balakrishnan CN, Payne RB. 2004 Clade-limited colonization in brood parasitic finches (*Vidua* spp.). *Syst. Biol.* **53**, 140–153. (doi:10.1080/10635150490265021)
32. Payne RB, Payne LL, Woods JL. 1998 Song learning in brood-parasitic indigobirds *Vidua chalybeata*: song mimicry of the host species. *Anim. Behav.* **55**, 1537–1553. (doi:10.1006/anbe.1997.0701)
33. Payne RB, Payne LL, Woods JL, Sorenson MD. 2000 Imprinting and the origin of parasite-host species associations in brood-parasitic indigo-birds, *Vidua chalybeata*. *Anim. Behav.* **59**, 69–81. (doi:10.1006/anbe.1999.1283)
34. Boughman JW. 1998 Vocal learning by greater spear-nosed bats. *Proc. R. Soc. Lond. B* **265**, 227–233. (doi:10.1098/rspb.1998.0286)
35. Esser K-H. 1994 Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *NeuroReport* **5**, 1718–1720. (doi:10.1097/00001756-199409080-00007)
36. Esser K-H, Schmidt U. 1989 Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae)—evidence for acoustic learning. *Ethology* **82**, 156–168. (doi:10.1111/j.1439-0310.1989.tb00496.x)
37. Esser K-H, Schubert J. 1998 Vocal dialects in the lesser spear-nosed bat *Phyllostomus discolor*. *Naturwissenschaften* **85**, 347–349. (doi:10.1007/s001140050513)
38. Lattenkamp EZ, Vernes SC, Wiegrefe L. 2018 Volitional control of social vocalisations and vocal usage learning in bats. *J. Exp. Biol.* **221**, 180729. (doi:10.1242/jeb.180729)
39. Ford JKB. 1991 Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.* **69**, 1454–1483. (doi:10.1139/z91-206)
40. Ford JKB, Fisher HD. 1983 Group-specific dialects of killer whales (*Orcinus orca*) in British Columbia. In *Communication and behavior of whales* (eds R Payne), pp. 129–161. Boulder, CO: Westview Press.
41. Yurk H, Barrett-Lennard L, Ford JKB, Matkin CO. 2002 Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Anim. Behav.* **63**, 1103–1119. (doi:10.1006/anbe.2002.3012)
42. Abramson JZ, Hernández-Lloreda MV, García L, Colmenares F, Aboitiz F, Call J. 2018 Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*). *Proc. R. Soc. B* **285**, 20172171. (doi:10.1098/rspb.2017.2171)
43. Delarue J, Todd SK, Van Parijs SM, Di Iorio L. 2009 Geographic variation in northwest Atlantic fin whale (*Balaenoptera physalus*) song: implications for stock structure assessment. *J. Acoust. Soc. Am.* **125**, 1774–1782. (doi:10.1121/1.3068454)
44. Stafford KM, Nieuwkirk SL, Fox CG. 2001 Geographic and seasonal variation of blue whale calls in the North Pacific. *J. Cetacean Res. Manage.* **3**, 65–76.
45. Stafford KM, Lydersen C, Wiig Ø, Kovacs KM. 2018 Extreme diversity in the songs of Spitsbergen's bowhead whales. *Biol. Lett.* **14**, 20180056. (doi:10.1098/rsbl.2018.0056)
46. Noad MJ, Cato D, H Bryden MM, Jenner MN, Jenner KCS. 2000 Cultural revolution in whale songs. *Nature* **408**, 537. (doi:10.1038/35046199)
47. Payne RS, McVay S. 1971 Songs of humpback whales. *Science* **173**, 583–597. (doi:10.1126/science.173.3997.585)
48. Smith JN, Goldizen AW, Dunlop RA, Noad MJ. 2008 Songs of male humpback whales, *Megaptera novaengliae*, are involved in intersexual interactions. *Anim. Behav.* **76**, 467–477. (doi:10.1016/j.anbehav.2008.02.013)
49. Herman LM. 2017 The multiple functions of male song within the humpback whale (*Megaptera novaengliae*) mating system: review, evaluation, and synthesis. *Biol. Rev.* **92**, 1795–1818. (doi:10.1111/brv.12309)
50. Dunlop RA, Noad MJ. 2016 The 'risky' business of singing: tactical use of song during joining by male humpback whales. *Behav. Ecol. Sociobiol.* **70**, 2149–2160. (doi:10.1007/s00265-016-2218-8)
51. Smith WJ. 1994 Animal duets: forcing a mate to be attentive. *J. Theor. Biol.* **166**, 221–223. (doi:10.1006/jtbi.1994.1019)
52. Baldassarre DT, Grieg EI, Webster MS. 2016 The couple that sings together stays together: duetting, aggression and extra-pair paternity in a promiscuous bird species. *Biol. Lett.* **12**, 20151025. (doi:10.1098/rsbl.2015.1025)
53. Logue DM, Gammon DE. 2004 Duet song and sex roles during territory defense in a tropical bird, the black-bellied wren, *Thryothorus fasciatoventris*. *Anim. Behav.* **68**, 721–731. (doi:10.1016/j.anbehav.2003.10.026)
54. Templeton CN, Rios-Chelen AA, Quiros-Guerrero E, Mann NI, Slater PJB. 2013 Female happy wrens select songs to cooperate with their mates rather than confront intruders. *Biol. Lett.* **9**, 20120863. (doi:10.1098/rsbl.2012.0863)
55. Sonnenschein E, Reyer HU. 1983 Mate-guarding and other functions of antiphonal duets in the slate-colored boubou (*Laniarius funebris*). *Z. Tierpsychol.* **63**, 112–140. (doi:10.1111/j.1439-0310.1983.tb00083.x)
56. Adret P *et al.* 2018 Duetting patterns of titi monkeys (primates, Pitheciidae: Callitrichinae) and relationships with phylogeny. *Animals* **8**, 178. (doi:10.3390/ani8100178)
57. Muller AE, Anzenberger G. 2002 Duetting in the titi monkey *Callitrichus cupreus*: structure, pair specificity and development of duets. *Folia Primatol.* **73**, 104–115. (doi:10.1159/000064788)
58. Okobi Jr DE, Banerjee A, Matheson AMM, Phelps SM, Long MA. 2019 Motor cortical control of vocal interaction in neotropical singing mice. *Science* **363**, 983–988. (doi:10.1126/science.aau9480)

59. Doupe AJ, Kuhl PK. 1999 Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631. (doi:10.1146/annurev.neuro.22.1.567)
60. Goldstein MH, Schwade JA. 2009 From birds to words: perception of structure in social interactions guides vocal development and language learning. In *The Oxford handbook of developmental behavioural neuroscience* (eds MS Blumberg, JH Freeman, SR Robinson), pp. 708–729. Oxford, UK: Oxford University Press.
61. Robinson CM, Snyder KT, Creanza N. 2019 Correlated evolution between repertoire size and song plasticity predicts that selection on song promotes open-ended learning. *eLife* **8**, e44454. (doi:10.7554/eLife.44454)
62. Beecher M, Brenowitz E. 2005 Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* **20**, 143–149. (doi:10.1016/j.tree.2005.01.004)
63. Odom KJ, Hall ML, Riebel K, Omland KE, Langmore NE. 2014 Female song is widespread and ancestral in songbirds. *Nat. Commun.* **5**, 3379. (doi:10.1038/ncomms4379)
64. Lobato M, Vellema M, Gahr C, Leitão A, de Lima SM, Geberzahn N, Gahr M. 2015 Mismatch in sexual dimorphism of developing song and song control system in blue-capped cordon-bleus, a songbird species with singing females and males. *Front. Ecol. Evol.* **3**, 117. (doi:10.3389/fevo.2015.00117)
65. Boughey MJ, Thompson NS. 1981 Song variety in the brown thrasher (*Toxostoma rufum*). *Z. Tierpsychol.* **56**, 47–58. (doi:10.1111/j.1439-0310.1981.tb01283.x)
66. Kershenbaum A, Freeberg TM, Gammon DE. 2015 Estimating vocal repertoire size is like collecting coupons: a theoretical framework with heterogeneity in signal abundance. *J. Theor. Biol.* **373**, 1–11. (doi:10.1016/j.jtbi.2015.03.009)
67. Colbert-White EN, Corballis MC, Frigaszy DM. 2014 Where apes and songbirds are left behind: a comparative assessment of the requisites for speech. *Comp. Cogn. Behav. Rev.* **9**, 1–28. (doi:10.3819/ccbr.2014.90001)
68. Bradbury JW, Balsby TJS. 2016 The functions of vocal learning in parrots. *Behav. Ecol. Sociobiol.* **70**, 293–312. (doi:10.1007/s00265-016-2068-4)
69. Forshaw JM. 1989 *Parrots of the world*. London, UK: Blandford.
70. Juniper T, Parr M. 1998 *Parrots: a guide to parrots of the world*. New Haven, CT: Yale University Press.
71. Kroodsmas DE, Konishi M. 1991 A subsong bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Anim. Behav.* **42**, 477–487. (doi:10.1016/S0003-3472(05)80047-8)
72. Dearing MD, Foley WJ, McLean S. 2005 The influence of plant secondary metabolites on the nutritional ecology of herbivorous terrestrial vertebrates. *Annu. Rev. Ecol. Syst.* **36**, 169–189. (doi:10.1146/annurev.ecolsys.36.102003.152617)
73. Bradbury JW. 2003 Vocal communication in wild parrots. In *Animal social complexity: intelligence, culture and individualized societies* (eds FBM de Waal, PL Tyack), pp. 293–316. Cambridge, MA: Harvard University Press.
74. Nottebohm F, Nottebohm M. 1969 The parrots of Bush Bush. *Anim. Kingdom* **72**, 18–23.
75. Farabaugh SM, Linzenbold A, Dooling RJ. 1994 Vocal plasticity in budgerigars (*Melopsittacus undulatus*): evidence for social factors in the learning of contact calls. *J. Comp. Psychol.* **108**, 81–92. (doi:10.1037/0735-7036.108.1.81)
76. Strahl SD. 1988 The social organization and behaviour of the hoatzin *Opisthocomus hoazin* in central Venezuela. *Ibis* **130**, 483–502. (doi:10.1111/j.1474-919X.1988.tb02714.x)
77. Pepperberg IM. 1990 Referential mapping: attaching functional significance to the innovative utterances of an African grey parrot. *Appl. Psycholing.* **11**, 23–44. (doi:10.1017/S0142716400008274)
78. Venuto V, Massa R, Bottoni L. 2001 African parrot vocalizations and their functional significance. *Ostrich Suppl.* **15**, 224–228.
79. Hile AG, Plummer TK, Striedter GF. 2000 Male vocal imitation produces call convergence during pair bonding in budgerigars. *Anim. Behav.* **59**, 1209–1218. (doi:10.1006/anbe.1999.1438)
80. Hile AG, Striedter GF. 2000 Call convergence within groups of female budgerigars (*Melopsittacus undulatus*). *Ethology* **106**, 1105–1114. (doi:10.1046/j.1439-0310.2000.00637.x)
81. Cortopassi KA, Bradbury JW. 2000 The comparison of harmonically rich sounds using spectrographic cross-correlation and principal coordinates analysis. *Bioacoustics* **11**, 89–127. (doi:10.1080/09524622.2000.9753454)
82. Balsby TJS, Scarl JC. 2008 Sex-specific responses to vocal convergence and divergence of contact calls in orange-fronted conures (*Aratinga canicularis*). *Proc. R. Soc. B* **275**, 2147–2154. (doi:10.1098/rspb.2008.0517)
83. Vehrencamp SL, Ritter AF, Keever M, Bradbury JW. 2003 Responses to playback of local vs. distant contact calls in the orange-fronted conure, *Aratinga canicularis*. *Ethology* **109**, 37–54. (doi:10.1046/j.1439-0310.2003.00850.x)
84. Balsby TJS, Momberg JV, Dabelsteen T. 2012 Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network. *PLoS ONE* **7**, e49747. (doi:10.1371/journal.pone.0049747)
85. Berg KS, Delgado S, Cortopassi KA, Beissinger SR, Bradbury JW. 2012 Vertical transmission of learned signatures in a wild parrot. *Proc. R. Soc. B* **279**, 585–591. (doi:10.1098/rspb.2011.0932)
86. Berg KS, Beissinger SR, Bradbury JW. 2013 Factors shaping the ontogeny of vocal signatures in a wild parrot. *J. Exp. Biol.* **216**, 338–345. (doi:10.1242/jeb.073502)
87. Boysson-Bardies B. 1993 Ontogeny of language-specific syllabic productions. In *Developmental neurocognition: speech and face processing in the first year of life* (eds B Boysson-Bardies, S Schonen, P Juszczyk, P MacNeilage, J Morton), pp. 353–363. Dordrecht, The Netherlands: Kluwer.
88. Snow D. 2002 Intonation in the monosyllabic utterances of 1-year-olds. *Infant Behav. Dev.* **24**, 393–407. (doi:10.1016/S0163-6383(02)00084-X)
89. Vihman MM. 1993 Variable paths to early word production. *J. Phonetics* **21**, 61–82. (doi:10.1016/S0095-4470(19)31321-X)
90. Iyer SN, Oller DK. 2009 Fundamental frequency development in typically developing infants and infants with severe-to-profound hearing loss. *Clin. Linguist. Phon.* **22**, 917–936. (doi:10.1080/02699200802316776)
91. Locke JL. 1993 *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
92. Locke JL, Pearson DM. 1992 Vocal learning and the emergence of phonological capacity: a neurobiological approach. In *Phonological development: models, research, implications* (eds CA Ferguson, L Menn, C Stoel-Gammon), pp. 91–129. Timonium, MD: York Press.
93. Oller DK, Eilers RE. 1988 The role of audition in infant babbling. *Child Dev.* **59**, 441–449. (doi:10.2307/1130323)
94. Oller DK, Eilers RE, Bull DH, Carney AE. 1985 Prespeech vocalizations of a deaf infant: a comparison with normal metaphonological development. *J. Speech Hear. Res.* **28**, 47–63. (doi:10.1044/jshr.2801.47)
95. Stoel-Gammon C, Otomo K. 1986 Babbling development of hearing-impaired and normally hearing subjects. *J. Speech Hear. Dis.* **51**, 33–41. (doi:10.1044/jshd.5101.33)
96. Goldstein MH, King AP, West MJ. 2003 Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl Acad. Sci. USA* **100**, 8030–8035. (doi:10.1073/pnas.1332441100)
97. Goldstein MH, Schwade JA. 2008 Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* **19**, 515–523. (doi:10.1111/j.1467-9280.2008.02117.x)
98. Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, Bakeman R. 2013 Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl Acad. Sci.* **110**, 6318–6323. (doi:10.1073/pnas.1300337110)
99. Griebel U, Oller DK. 2008 Evolutionary forces favoring contextual flexibility. In *Evolution of communicative flexibility: complexity, creativity and adaptability in human and animal communication* (eds DK Oller, U Griebel), pp. 9–40. Cambridge, MA: MIT Press.
100. Oller DK et al. 2019 Preterm and full term infant vocalization and the origin of language. *Sci. Rep.* **9**, 1–10. (doi:10.1038/s41598-019-51352-0)
101. Janik V, Slater P. 2000 The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11. (doi:10.1006/anbe.2000.1410)
102. Goldstein MH, Schwade JA, Bornstein MH. 2009 The value of vocalizing: five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Dev.* **80**, 636–644. (doi:10.1111/j.1467-8624.2009.01287.x)

103. Franklin B, Warlaumont AS, Messinger D, Bene E, Nathani-Iyer S, Lee CC. 2013 Effects of parental interaction on infant vocalization rate, variability and vocal type. *Lang. Learn. Dev.* **10**, 1–18. (doi:10.1080/15475441.2013.849176)
104. Amsel A. 1958 The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.* **55**, 102–119. (doi:10.1037/h0043125)
105. Warlaumont AS, Richards JA, Gilkerson J, Oller DK. 2014 A social feedback loop for speech development and its reduction in autism. *Psychol. Sci.* **25**, 1314–1324. (doi:10.1177/0956797614531023)
106. Locke JL. 2006 Parental selection of vocal behavior: crying, cooing, babbling, and the evolution of language. *Hum. Nat.* **17**, 155–168. (doi:10.1007/s12110-006-1015-x)
107. Goldstein MH, West MJ. 1999 Consistent responses of human mothers to prelinguistic infants: the effect of prelinguistic repertoire size. *J. Comp. Psychol.* **113**, 52–58. (doi:10.1037/0735-7036.113.1.52)
108. Gros-Louis J, West MJ, Goldstein MH, King AP. 2006 Mothers provide differential feedback to infants' prelinguistic sounds. *Int. J. Behav. Dev.* **30**, 509–516. (doi:10.1177/0165025406071914)
109. Albert RR, Schwade JA, Goldstein MH. 2017 The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Dev. Sci.* **18**, e12641. (doi:10.1111/desc.12641)
110. Elmlinger SL, Schwade JA, Goldstein MH. 2019 The ecology of prelinguistic vocal learning: parents simplify the structure of their speech in response to babbling. *J. Child Lang.* **46**, 998–1011. (doi:10.1017/S0305000919000291)
111. Goldstein MH, Schwade J, Briesch J, Syal S. 2010 Learning while babbling: prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy* **15**, 362–391. (doi:10.1111/j.1532-7078.2009.00020.x)
112. Catchpole CK, Slater PJB. 2008 *Bird song: themes and variations*, 2nd edn. New York, NY: Cambridge University Press.
113. Krebs JR. 1977 The significance of song repertoires: the Beau Geste hypothesis. *Anim. Behav.* **25**, 475–478. (doi:10.1016/0003-3472(77)90022-7)
114. Freeberg T, Dunbar RIM, Ord TJ. 2012 Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B* **367**, 1785–1801. (doi:10.1098/rstb.2011.0213)
115. Freeberg T. 2006 Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychol. Sci.* **17**, 557–561. (doi:10.1111/j.1467-9280.2006.01743.x)
116. Marler P, Nelson DA. 1992 Action-based learning: a new form of developmental plasticity in bird song. *Neth. J. Zool.* **43**, 91–103. (doi:10.1163/156854293X00232)
117. West MJ, King AP. 1988 Female visual displays affect the development of male song in the cowbird. *Nature* **334**, 244–246. (doi:10.1038/334244a0)
118. Langmore NE, Maurer G, Adcock GJ, Kilner RM. 2008 Socially acquired host-specific mimicry and the evolution of host races in Horsfield's bronze-cuckoo *Chalcites basalis*. *Evolution* **62**, 1689–1699. (doi:10.1111/j.1558-5646.2008.00405.x)
119. Garland EC, Goldizen AW, Rekdahl ML, Constantine R, Garrigue C, Hauser ND, Poole MM, Robbins J, Noad MJ. 2011 Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Curr. Biol.* **21**, 687–691. (doi:10.1016/j.cub.2011.03.019)
120. Nottebohm F, Nottebohm ME, Crane L. 1986 Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei. *Behav. Neural Biol.* **46**, 445–471. (doi:10.1016/S0163-1047(86)90485-1)
121. Allen JA, Garland EC, Dunlop RA, Noad MJ. 2018 Cultural revolutions reduce complexity in the songs of humpback whales. *Proc. R. Soc. B* **285**, 20182088. (doi:10.1098/rspb.2018.2088)
122. Fitch WT. 2010 *The evolution of language*. Cambridge, UK: Cambridge University Press.
123. Lowry H, Lill A, Wong BBM. 2013 Behavioural responses of wildlife to urban environments. *Biol. Rev.* **88**, 537–549. (doi:10.1111/brv.12012)
124. Hoppitt W, Laland KN. 2008 Social processes influencing learning in animals: a review of the evidence. *Adv. Study Behav.* **38**, 105–165. (doi:10.1016/S0065-3454(08)00003-X)
125. Rendell L *et al.* 2010 Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213. (doi:10.1126/science.1184719)
126. Payne RB. 1996 Song traditions in indigo buntings: origin, improvisation, dispersal and extinction in cultural evolution. In *Ecology and evolution of acoustic communication in birds* (eds DE Kroodsma, EH Miller), pp. 198–220. Ithaca, NY: Cornell University Press.
127. Mesoudi A, Whiten A, Laland KN. 2004 Is human cultural evolution Darwinian? Evidence reviewed from the perspective of 'The Origin of Species'. *Evolution* **58**, 1–11. (doi:10.1111/j.0014-3820.2004.tb01568.x)
128. Cantor M, Shoemaker LG, Cabral RB, Flores CO, Varga M, Whitehead H. 2015 Multilevel animal societies can emerge from cultural transmission. *Nat. Commun.* **6**, 8091. (doi:10.1038/ncomms9091)
129. Williams H, Levin II, Norris DR, Newman AEM, Wheelwright NT. 2013 Three decades of cultural evolution in savannah sparrow songs. *Anim. Behav.* **85**, 213–223. (doi:10.1016/j.anbehav.2012.10.028)
130. Otter KA, Mckenna A, LaZerte SE, Ramsay SM. 2020 Continent-wide shifts in song dialects of white-throated sparrows. *Curr. Biol.* **30**, 3231–3235.e3. (doi:10.1016/j.cub.2020.05.084)
131. Mikula P, Valcu M, Brumm H, Bulla M, Forstmeier J, Petrusková T, Kempnaers B, Albrecht T. 2021 A global analysis of song frequency in passerines provides no support for the acoustic adaptation hypothesis but suggests a role for sexual selection. *Ecol. Lett.* **24**, 477–486. (doi:10.1111/ele.13662)
132. Slabbekoorn H, Peet M. 2003 Birds sing at a higher pitch in urban noise. *Nature* **424**, 267. (doi:10.1038/424267a)
133. Brumm H, Zollinger SA. 2013 Avian vocal production in noise. In *Animal communication and noise* (eds H Brumm), pp. 187–227. Berlin, Germany: Springer.
134. Rendell LE, Whitehead H. 2003 Vocal clans in sperm whales (*Physeter macrocephalus*). *Proc. R. Soc. Lond. B* **270**, 225–231. (doi:10.1098/rspb.2002.2239)
135. Garland EC, McGregor PK. 2020 Cultural transmission, evolution, and revolution in vocal displays: insights from bird and whale song. *Front. Psychol.* **11**, 544929. (doi:10.3389/fpsyg.2020.544929)
136. Uomini N, Fairlie J, Gray RD, Griesser M. 2020 Extended parenting and the evolution of cognition. *Phil. Trans. R. Soc. B* **375**, 20190495. (doi:10.1098/rstb.2019.0495)
137. Faust KM, Carouso-Peck S, Elson MR, Goldstein MH. 2020 The origins of social knowledge in altricial species. *Ann. Rev. Dev. Psychol.* **2**, 225–246. (doi:10.1146/annurev-devpsych-051820-121446)
138. Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. 2016 Vocal fold control beyond the species-specific repertoire in an orangutan. *Sci. Rep.* **6**, 30315. (doi:10.1038/srep30315)
139. Lameira AR, Shumaker RW. 2019 Orangutans show active voicing through a membranophone. *Sci. Rep.* **9**, 12289. (doi:10.1038/s41598-019-48760-7)
140. Sayigh LS, Wells RS, Janik VM. 2017 What's in a voice? Dolphins do not use voice cues for individual recognition. *Anim. Cognit.* **20**, 1067–1079. (doi:10.1007/s10071-017-1123-5)
141. Madsen PT, Jensen FH, Carder D, Ridgway S. 2011 Dolphin whistles: a functional misnomer revealed by heliox breathing. *Biol. Lett.* **8**, 211–213. (doi:10.1098/rsbl.2011.0701)
142. Feeles F. 1982 Song mimesis within colonies of *Cacicus c. cela* (Icteridae, Aves): a colonial password? *Z. Tierpsychol.* **58**, 119–152. (doi:10.1111/j.1439-0310.1982.tb00312.x)
143. Ritchison G. 1983 The function of singing in female black-headed grosbeaks (*Pheucticus melanocephalus*): family-group maintenance. *Auk* **100**, 105–116. (doi:10.1093/auk/100.1.105)
144. Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S. 2005 Elephants are capable of vocal learning. *Nature* **434**, 455–456. (doi:10.1038/434455a)
145. Ralls K, Fiorelli P, Gish S. 1985 Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Can. J. Zool.* **63**, 1050–1056. (doi:10.1139/z85-157)
146. Musser WB, Bowles AE, Grebner DM, Grance JL. 2014 Differences in acoustic features of vocalizations produced by killer whales cross-socialized with bottlenose dolphins. *J. Acoust. Soc. Am.* **136**, 1990–2002. (doi:10.1121/1.4893906)