



Published in final edited form as:

Anim Behav. 2016 October ; 120: 163–172. doi:10.1016/j.anbehav.2016.07.031.

Social calls provide novel insights into the evolution of vocal learning

Kendra B. Sewall^a, Anna M. Young^b, and Timothy F. Wright^{c,*}

^aDepartment of Biological Sciences, Virginia Tech, Blacksburg, VA, U.S.A

^bDepartment of Biology and Earth Science, Otterbein University, Westerville, OH, U.S.A

^cDepartment of Biology, New Mexico State University, Las Cruces, NM, U.S.A

Abstract

Learned song is among the best-studied models of animal communication. In oscine songbirds, where learned song is most prevalent, it is used primarily for intrasexual selection and mate attraction. Learning of a different class of vocal signals, known as contact calls, is found in a diverse array of species, where they are used to mediate social interactions among individuals. We argue that call learning provides a taxonomically rich system for studying testable hypotheses for the evolutionary origins of vocal learning. We describe and critically evaluate four nonmutually exclusive hypotheses for the origin and current function of vocal learning of calls, which propose that call learning (1) improves auditory detection and recognition, (2) signals local knowledge, (3) signals group membership, or (4) allows for the encoding of more complex social information. We propose approaches to testing these four hypotheses but emphasize that all of them share the idea that social living, not sexual selection, is a central driver of vocal learning. Finally, we identify future areas for research on call learning that could provide new perspectives on the origins and mechanisms of vocal learning in both animals and humans.

Keywords

communication; contact calls; evolution; social dynamics; vocal learning

At present in the field of animal behaviour, learned vocal communication is most commonly studied in the context of the songs of male oscine birds, which are used in intrasexual competition and mate attraction and thus are shaped by sexual selection (Searcy & Andersson, 1986; Searcy & Nowicki, 2005). Birdsong has earned this research focus in part because of its elaborate, varied and conspicuous production (Catchpole & Slater, 1995; Marler & Slabbekoorn, 2004) and in part because it has notable parallels with human speech in both developmental timelines and neural underpinnings (Bolhuis et al., 2010; Doupe &

*Correspondence: T. F. Wright, Department of Biology, MSC 3AF, New Mexico State University, Las Cruces, NM 88003, U.S.A. wright@nmsu.edu (T. F. Wright).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Kuhl, 1999; Jarvis, 2004; Miyagawa et al., 2014; Petkov & Jarvis, 2012; Wilbrecht & Nottebohm, 2003). This focus on male oscine song has had two unintended consequences for the study of vocal production learning (hereafter termed ‘vocal learning’, see Table 1 for definitions): it has limited the study of other classes of communication signals and it has led to the general inference that sexual selection has been the primary force driving the evolution of vocal learning (Nowicki & Searcy, 2014; Nottebohm, 1972; Puts et al., 2007; Miller, 2000; Burling, 2007). However, several lines of evidence are inconsistent with the idea that classical sexual selection drove the origin of vocal learning across other species, or even within the songbirds. First, the recent finding that song production by females is widespread in oscine songbirds and that singing by both sexes likely represents the ancestral state in this group calls into question the common assumption that song has always been central to mate choice and, in turn, undermines the hypothesis that sexual selection is primarily responsible for the evolution of song learning (Odom et al., 2014). Second, vocal learning of less elaborate vocal signals, often termed ‘calls’, occurs in diverse taxa including parrots, whales, seals, elephants, bats and primates, and many of these taxa lack elaborate songs altogether but share a propensity to form highly social groups (Bradbury, 2003; Janik, 2014; Janik & Slater, 1997; Knörnschild, 2014; Petkov & Jarvis, 2012; Reichmuth & Casey, 2014; Stoeger & Manger, 2014; Tyack, 2008; Watson et al., 2015; Toft & Wright, 2015). These observations have led some to propose the alternative hypothesis that learned communication in animals, including humans, has evolved as a means of better mediating complex and dynamic social interactions, rather than via sexual selection driven by mate choice (Fitch et al., 2010; Freeberg et al., 2012; Janik, 2014; Pinker, 2010; Sewall, 2015; Seyfarth & Cheney, 2014; Tyack, 2008; but see Burling, 2007; Fitch, 2005; Miller, 2000; Puts et al., 2007).

In contrast to song, individuals of all age classes and both sexes in a variety of species produce and respond to calls in a range of social contexts. Some examples of vocalizations termed ‘calls’ include alarm calls, mobbing calls, food begging calls and isolation calls (Marler, 2004). Importantly, a diverse array of taxa are capable of vocal learning of a particular category of signals, known as ‘contact calls’, which are used by juveniles and adults of both sexes when contacting or coordinating behaviours with conspecifics (Kondo & Watanabe, 2009; Marler, 2004; Table 2). In this essay we argue that contact calls used to mediate social interactions represent a valuable and understudied model of vocal learning in animals that can provide important new perspectives on the evolutionary origins, developmental processes and neural mechanisms underlying learned communication.

PATTERNS OF CONTACT CALL LEARNING

Although several specific social functions have been proposed for learned calls, the unifying theme is that vocal learning can permit flexibility in social associations not possible with nonlearned signals and can encode complex information by increasing signal diversity. Thus, just as sociality is proposed to drive the evolution of intelligence and cognitive specialization (Byrne & Whiten, 1989; Jolly, 1966), it may also contribute to the origin and maintenance of vocal learning abilities (Dunbar, 2003; Fitch et al., 2010; McComb & Semple, 2005; Pinker, 2010). Here, we describe patterns of contact call learning and then consider four nonmutually exclusive hypotheses previously proposed to explain the origin and

maintenance of this ability, all of which are rooted in the broader argument that social dynamics drove the repeated evolution of vocal learning.

Vocal learning of calls has been most frequently described within contact calls, so we focus on this category of calls for review (Table 2). As with song, vocal learning of contact calls has sometimes been inferred by the existence of shared calls that are unique to a social unit, such as group-specific calls or regional dialects (Janik & Slater, 2000), although alternative processes such as reduced dispersal, biased settlement or assortative mating can also give rise to these patterns (Groth, 1993; Price, 1998; Rendell & Whitehead, 2003; Yurk, 2002). Learning of contact calls is perhaps best demonstrated through longitudinal recordings documenting *convergence* (when conspecifics collectively change their call structure to generate a novel, shared variant) or *imitation* (when one individual replicates the call of one or more companions) over time (Janik & Slater, 2000). Learning could also occur through the novel recombination of existing signals (Templeton et al., 2005), but this route has been described less frequently and for the purpose of this paper we will focus on changes in the acoustic structure of signals.

Animals may share acoustically similar contact call variants at different levels of social organization including mated pairs, family lineages, social groups and populations (Table 2). Family-specific shared calls can emerge when young imitate the calls of parents or matriline (Sewall, 2011; Yurk, 2002). Pair-specific and group-specific calls can develop when individuals imitate the existing calls of companions (Boughman, 1998; Hile et al., 2000) or when all social partners modify their calls (i.e. convergence, Farabaugh et al., 1994; Hile & Striedter, 2000). Population-specific calls, or dialects, can occur when animals living in a geographical area learn a similar call structure (Salinas-Melgoza & Wright, 2012; Sewall, 2009, 2011; Wright, 1996). When call learning is restricted to a critical period early in life, shared signals have the potential to constrain movement among families, social groups and populations (Sewall, 2009; Wright, 1996). In many taxa, however, call production learning can continue into adulthood and such flexibility may be particularly important for encoding changing social relationships and facilitating flexibility in social bonds (Dahlin et al., 2014; Salinas-Melgoza & Wright, 2012; Sewall, 2009).

HYPOTHESES FOR THE EVOLUTIONARY ORIGINS OF CALL LEARNING

Ultimate explanations for the origins of call learning must specify the benefits accrued by the first individuals to display this complex trait in order to explain adequately its establishment (Nowicki & Searcy, 2014). These benefits may be distinct from those that explain the maintenance of vocal learning, which address current function and fitness benefits of learning once it has spread within a population. In the context of contact calls used to coordinate social interactions, the potential current benefits of vocal learning to individuals are those of group membership, such as cooperative defence of resources (Wilkinson & Boughman, 1998), improved foraging efficiency (Smith et al., 1999) and shared predator vigilance (Elgar, 1989). In contrast, there are four hypotheses in the literature that attempt to address the benefits that could have been reaped by the first individuals that evolved the capacity for vocal learning (Table 3): (1) improving signal recognition by intended receivers (Tyack, 2008); (2) signalling familiarity with a local

environment (Nottebohm, 1972) while maintaining the ability to move across geographical and social boundaries; (3) signalling social alliances both to unfamiliar group members and to nongroup members (Feekes, 1982; Wilkinson & Boughman, 1998); and (4) increasing the amount of information that could be encoded in signals through greater signal complexity (Freeberg et al., 2012; Nowicki & Searcy, 2014). Below, we first consider the potential costs and benefits of contact call learning to signallers and receivers in the context of signal evolution and honesty. In the following section we then discuss each of these hypotheses, their limitations, and their predictions (see also Table 4) and conclude with suggestions for future research directions.

Given that many species that do not learn to modify their call production are still capable of learning to use call variants in particular contexts (i.e. contextual learning; Table 1) and to display appropriate behaviours in response to hearing call variants (i.e. comprehension learning; Janik & Slater, 2000), it is reasonable to hypothesize that such learning in receivers preceded vocal production learning (Tyack, 2008). Indeed, many species that live in stable social groups learn to recognize unique individual variation in unlearned calls (e.g. 'signature calls') to distinguish among familiar conspecifics (Aubin & Jouventin, 2002; Cheney et al., 1995; Insley, 2001; Kober et al., 2008; Oda, 2002; Rendall et al., 1996; Sousa-Lima et al., 2002). Since comprehension learning provides an alternative, and potentially cognitively simpler, strategy for individual recognition of group membership (Tyack, 2008; Vehrencamp et al., 2003), call production learning must be less costly than comprehension learning in order to evolve and persist. This could occur either if circumstances make comprehension learning very costly, or if vocal learning of calls has very low costs. The first scenario, of comprehension learning becoming costly or inefficient, is predicted to occur when group membership changes rapidly or frequently, or when groups are very large (Tyack, 2008; Vehrencamp et al., 2003). Importantly, in the case of contact call learning, all individuals play the roles of both signallers and receivers and thus the costs of vocal production learning and comprehension learning are incurred by the same individuals, albeit at different moments. This is different from song learning, when signallers and receivers incur different costs because signallers must learn to produce and respond to vocalizations, while receivers may not bear the cost of learning to produce signals. Because every individual that learns to produce novel calls has a reduced burden of learning to recognize multiple signals, evolutionary conflict between sender and receiver may be eliminated and only the net cost of learning should influence trait evolution.

The alternative, that vocal learning is not costly, is also possible in some cases. Although several mechanisms might link learned signals to the experience, phenotype or intent of a signaller, the costs of vocal learning itself are unclear. One possible cost could be imposed by the specialized neural mechanisms underpinning vocal learning (Bolhuis et al., 2010; Feenders et al., 2008; Jarvis, 2004), which may be physiologically costly to develop and maintain (Isler & Schaik, 2006; Mink et al., 1981; Nowicki et al., 2002). If the neural machinery of vocal learning imposes a physiological cost, then the accuracy or speed of vocal learning could reflect an individual's quality or condition, permitting companions to assess the value of a new individual seeking access to their social group. Second, there are temporal and resource trade-offs inherent to all learning processes, such as the time required to learn relative to the time that could be spent engaged in other activities. Such temporal

costs could ensure that call learning reliably reflects a signaller's prior social experience because the time and social interaction required by the learning process effectively encode an individual's social exposure. Third, there is the potential cost of mistakes of learning to produce inaccurate calls, which could lead to misidentification by other group members. However, the possibility of costly mistakes also exists with contextual or comprehension learning. Fourth, social retaliation for producing dishonest signals can negatively affect a signaller and may promote learning and production of signals that reliably encode intent (Akçay et al., 2009; Smith et al., 2000). However, because social calls also mediate group interactions, there are many cases when signallers and receivers are not in conflict and communication can be mediated by *low-cost conventional signals*, defined as signals with low intrinsic costs and an arbitrary relationship between a signal's form and its message that is mutually agreed upon by sender and receiver (Maynard Smith & Harper, 2003). Specifically, when shared calls function to ensure coordinated group behaviours from which both signallers and receivers benefit, such as cooperative foraging, then there is no benefit to cheating and thus signal honesty need not be enforced by signalling costs.

Improving Signal Recognition

Shared calls have the potential to permit rapid identification of companions because of the mechanisms of auditory perception underlying signal reception (Endler & Basolo, 1998; Guilford & Dawkins, 1991). Specifically, to facilitate the motor learning essential to vocal production, animals' auditory processing systems are sensitized not only to their own vocalizations but also to signals that are similar to their own signals (Margoliash, 1983; Theunissen et al., 2004). This selective sensitivity could ensure that listeners will recognize and pay attention to imitations of their own vocalizations, even in noisy environments (the cocktail party effect; Busnel & Mebes, 1975) and suggests that signallers who produce these imitations will receive enhanced attention from the intended listener (Miller et al., 2004; Sugiura, 1998). Such a process could be controlled by auditory neurons that are sensitive to both the calls of an individual and to acoustically similar calls produced by others (e.g. auditory mirror neurons, Prather et al., 2008).

Evidence that imitation of conspecifics' calls improves signal recognition and benefits both signallers and receivers comes from behavioural studies of several taxa, including dolphins (reviewed in Janik & Sayigh, 2013) and parrots. In spectacled parrotlets, *Forpus conspicillatus*, each member of a parrotlet family produces a unique, 'signature' call but also produces a mimicked version of the signature contact call of each companion when interacting with that specific bird (Wanker & Fischer, 2001; Wanker et al., 1998, 2005). Parrotlets respond more strongly to the imitations of their own signature call than to other calls, consistent with imitated calls drawing the attention of a targeted receiver (Wanker et al., 2005). In orange-fronted conures, *Eupsittula* (formerly *Aratinga*) *canicularis*, playbacks of contact calls to pairs of wild-caught birds elicited faster and stronger vocal responses from the bird whose calls were more closely matched by the playback exemplar (Balsby et al., 2012). Similarly, playbacks of signature whistle calls in wild bottlenose dolphins, *Tursiops truncatus*, only elicited responses from individuals when their own whistles were broadcast (King & Janik, 2013). Budgerigars, *Melopsittacus undulatus*, produce contact call variants that are shared with mates and flock members, as well as unshared variants; shared

calls contain acoustic signatures of both the sender and receiver (Dahlin et al., 2014), and hearing playback of shared and nonshared calls results in different patterns of brain activity in receivers (Brauth et al., 2002). Such differential neural response is consistent with individuals being sensitized to calls like their own and offers a mechanism that could ensure that receivers recognize and attend to the calls of companions, even in noisy environments (Tyack, 2008). However, improved signal recognition via shared calls may not be a unitary selection pressure driving the origin of call learning, as hypotheses about the benefits accrued by individuals using call imitation to facilitate cooperative relationships (below) also depend upon receivers showing enhanced responsiveness to imitated calls. Thus, improved recognition of calls was likely linked to another benefit of vocal learning when it first evolved. Furthermore, our understanding of this phenomenon is based on just a few taxa. Testing for an enhanced response to shared calls using subjects with known prior experience with signallers and naturally varying playback exemplars in a wider range of taxa will be important to test the generality of this phenomenon.

Signalling Local Knowledge

Vocal learning encodes prior experience by virtue of the learning process, which can take weeks or even longer in some species (Mundinger, 1979). The time required for learning can therefore honestly reflect the degree to which an individual is familiar with the local ecological environment, which can make calls shared at the level of populations indicators of local knowledge (Nottebohm, 1972; Nowicki & Searcy, 2014). Receivers that associate with a companion that has learned calls similar to its own may be assured that this associate is familiar with local food resources and predators (Deecke et al., 2010; Mammen & Nowicki, 1981). Thus, signallers that are capable of learning will benefit by being accepted into a group and receivers that prefer vocal learners will benefit from a knowledgeable group member. Importantly, unlike genetically encoded signals that might be associated with geographical areas or stable groups of individuals, learned signals permit individuals to move among populations and groups during their lifetime. This hypothesis can explain the maintenance of vocal learning but its utility in explaining the origin of vocal learning requires the preexistence of genetically based dialects that are used for social discrimination (Nowicki & Searcy, 2014). As there are nonvocal learning species that produce group-specific calls (Townsend et al., 2010), and environmental gradients can lead to distinct vocalizations in nonlearning species even in a similar habitat (Kirschel et al., 2009; Tobias et al., 2010), it is plausible that genetically based group signals represent an ancestral state. This appears to be the case in capybaras, *Hydrochoerus hydrochaeris*, where groups defend local territories against nongroup members, mediate interactions with vocalizations, and have differences in acoustic properties of calls (Barros et al., 2011). Thus, vocal learning could have conferred a selective advantage on the first learners by permitting them to associate with groups in more than one geographical area or to move to a new population. Determining how signalling familiarity with the local environment might contribute to the evolution of call learning requires comparative studies examining the frequency with which geographical dialects correspond with call learning and the extent to which dialects are associated with group benefits. At present, no such phylogenetic comparison has been conducted either for learned or unlearned call dialects, making the ancestral state unclear, and thus drawing into question the potential evolutionary benefit of vocal learning.

Additionally, future studies within dialect systems must demonstrate that receivers discriminate among dialects and only allow individuals with local dialects to benefit from social interactions such as cooperative foraging or predator vigilance. Overall, although the hypothesis that vocal learning provides a benefit by encoding local knowledge is a long-standing one (Nottebohm, 1972), neither phylogenetic nor experimental evidence is sufficient to support this idea at this time.

Signalling Social Affiliations or Group Membership

While the hypothesis that shared vocalizations indicate regional familiarity originated in studies of learned song and song dialects, studies of shared calls at smaller geographical scales have generated a related hypothesis: that shared vocalizations can encode group membership and permit the recognition of social alliances (the ‘badge’ or ‘password’ hypothesis; Feekes, 1982; Wilkinson & Boughman, 1998). Possessing a group badge or password could be beneficial to individuals in a group if membership is associated with cooperative interactions that are both costly to participants and vulnerable to cheating. The potential for learned shared vocalizations to facilitate recognition of other group members in especially large or fluid social groups has been proposed for diverse taxa because these species represent cases when learning to produce a shared calls has the potential to be more efficient than learning and remembering many individual calls (Bradbury & Balsby, 2016). For example, greater spear-nosed bats, *Phyllostomus hastatus*, learn roost-specific calls while foraging and these shared calls permit individuals to quickly identify other group members, possibly to facilitate the cooperative defence of rich food resources (Wilkinson & Boughman, 1998). Shared group-specific calls also have the potential to reduce cognitive burden or signal processing time if individuals must only learn a single call, rather than the distinctive call of each group member (Tyack, 2008; Vehrencamp et al., 2003). Additionally, shared calls can signal the size and strength of an alliance either within a cohesive group or to other social groups (Wilkinson & Boughman, 1998). Evidence that shared calls permit alliance recognition by nongroup members comes from male bottlenose dolphins, who use their alliance-specific whistles during competitive interactions with unfamiliar individuals and their unique signature whistles during interactions with familiar group members (Janik & Slater, 1998).

As with the proposed association between vocal learning and local knowledge, it is easier to understand how shared calls facilitate group cohesion once vocal learning is widespread in a population than it is to posit a role in the origin of vocal learning (Nowicki & Searcy, 2014). Again, the first learners would have gained a new capacity to move among social groups, either during early development if they were close-ended learners or throughout life if open-ended learners, presuming group membership was previously based on ‘innate’ vocal similarity, and this capacity could potentially provide a fitness advantage. In the case of fission–fusion social groups, selection could favour life-long learning rather than learning only during a critical period at the time of dispersal and group recruitment. Additionally, if receivers co-evolved enhanced sensitivity to imitated calls (see Improving Signal Recognition, above), the first vocal-learning lineage could have benefited from improved coordination of cooperative behaviours such as group foraging. Again, testing the hypothesis that signalling social affiliation drove the evolution of vocal learning entails comparative

studies of the frequency with which complex social dynamics are associated with call learning. This requires carefully defining social complexity (Bergman & Beehner, 2015; Blumstein & Armitage, 1997) and then testing for associations between complexity and call learning. While operationally defining social complexity is beyond the scope of this review, several alternatives have been proposed beyond simply measuring group size, including the number of demographic roles held by a single group member (Blumstein & Armitage, 1997) and the number of differentiated relationships maintained by individuals within groups (Bergman & Beehner, 2015). Finally, playback studies demonstrating that receivers discriminate among call types and that shared signals are associated with cooperative interactions are needed to support the functional hypothesis that learned calls signal group membership.

Increasing Information Complexity

Vocal learning increases signal diversity and complexity, thereby permitting more information to be encoded by communication systems (Freeberg et al., 2012; Jackendoff, 1999; Nowicki & Searcy, 2014). In the case of contact calls, vocal learning can generate variation within this signal category that is specific to social associations and, importantly, can permit new signals to emerge and indicate new social bonds (Table 2). In fact, learned calls can be shared at multiple levels, permitting a single call to simultaneously reflect information about social associations within a nested hierarchy of social groups such as social pairs, groups and populations, a phenomenon termed *hierarchical mapping* (Bradbury & Vehrencamp, 2011). These complex patterns of call sharing may be important both in denoting past associations and negotiating future associations by individuals in groups (Bradbury & Balsby, 2016). For example, the calls of budgerigars reflect individual identity as well as pair or group membership (Dahlin et al., 2014). Similarly, the calls of red crossbills, *Loxia curvirostra*, reflect family and pair affiliation within the bounds of broader dialects (Sewall, 2009, 2011).

The first individuals capable of vocal learning would have thus been able to signal their affiliations among multiple social levels and with multiple individuals, thereby reaping benefits from more than one social affiliation. Like the previous hypotheses, affiliations would have had to be encoded by nonlearned but shared calls prior to vocal learning first evolving. In fact, the role of learned calls in signalling social affiliation or group membership may have preceded the potential for learning to encode multiple levels of social complexity, making this hypothesis a subcategory of the previous hypothesis, in the context of contact calls. The capacity to learn, and particularly for companions to converge on novel signals, could have increased the signalling repertoire and permitted such ‘new’ signals to encode newly established social bonds. The potential for learned vocalizations to encode more information than nonlearned calls is particularly compelling in the context of complex social dynamics, both because learned signals can be modified sufficiently to encode new bonds within fluid social environments, and because their potential to encode more complex or specific information is open-ended (Freeberg et al., 2012; Nowicki & Searcy, 2014). Examining the relationship between signal complexity, social complexity and vocal learning across taxa is an important area of future study in the field of vocal learning and can be best addressed with comparative approaches (Blumstein & Armitage, 1997; Freeberg et al., 2012;

Sewall, 2015). If call learning encodes complex social information, then the number of groups at an equivalent level of the social hierarchy to which individuals of a given species belong should be positively associated with the number of contact call variants those individuals produce. Additionally, playback studies in species with learned calls should determine whether receivers discriminate among variants and behave in a way consistent with recognizing the social bond with different signallers (for example see Cheney et al., 1995). Similarly, quantifying the degree of call similarity among members of different social cohorts or demographics within a larger group will determine whether different degrees of call similarity reflect different types of social bonds, as is predicted by this hypothesis. Changing social group composition and mapping changes in calls with the formation of new social bonds will provide definitive evidence that new, multilevel social connections are encoded by learning. Finally, it is possible that vocal learning encodes information through combinatorial complexity of multiple call syllables as well (Freeberg et al., 2012). Note, however, that the information complexity hypothesis may be difficult to distinguish from the social affiliation hypothesis in species that encode varied social affiliations with contact calls (Table 4); studies of call learning in predator or food calls may provide better tests of the information complexity hypothesis (Freeberg et al., 2012; Templeton et al., 2005).

A Role for Sexual Selection

While the production of learned calls by females and juveniles in many species diminishes the explanatory power of sexual selection as a selective force driving the origin of call learning, call imitation can play a role in social bonding and perhaps even mate choice. The association between call learning and social affiliation found in many species suggests that learning itself may serve as an honest signal of commitment or affiliation, which could facilitate social bonding at multiple levels of social organization including within mated pairs (Hile et al., 2000; Mammen & Nowicki, 1981; Mundinger, 1979). Furthermore, the ability to learn new calls quickly or with greater fidelity could provide receivers with information about a signaller's quality or other learning abilities (Boogert et al., 2011; Nowicki & Searcy, 2004). For example, female budgerigars prefer males that are tutored to produce imitations of the female's calls before pairing, and females paired with brain-lesioned males incapable of learning engaged in more extrapair copulations (Hile et al., 2005). Whether call learning reflects affiliation, a male's cognitive ability, or local knowledge is unclear, but female preference for vocal learning in budgerigars suggests that mate choice can provide additional selection for call learning, even if it was not the primary force driving its origin.

FUTURE DIRECTIONS

We see four key areas of future research for understanding the function of learned calls and the evolution of vocal learning: (1) determining the functional relationship between vocal learning and social dynamics; (2) better understanding costs and constraints on the evolution of vocal learning; (3) determining how call learning in animals relates to human speech learning; and (4) resolving the neural mechanisms underlying call learning across diverse taxa. Below we suggest future avenues for research in each of these key areas.

The Relationship between Vocal Learning and Social Dynamics

Complex social dynamics are likely driving the ongoing evolution of learning, pushing vocal learning to be faster and more flexible. The use of individual-level versus group-level signatures for recognition may be associated with group size or other aspects of social complexity. The social brain hypothesis maintains that neocortex size places a constraint on the size of social groups; supporting evidence for this has been found in primates (Dunbar, 1992). Such a group size limit may be imposed by cognitive limits on how many individuals one group member can recognize and maintain social relationships with. Above that limit of recognizable associates, group-level signatures to differentiate nonassociates would be favoured. In budgerigars, individuals can discriminate between groups based on shared contact calls, but only among individuals within their flock (Ali, Farabaugh & Dooling, 1993; Dooling, 1986). An informal survey across taxa with contact call learning suggests that in addition to large social aggregations of unrelated individuals, transient social bonds, noisy social environments and cooperative defence of resources are all associated with call learning and therefore may also be the factors that originally drove its evolution (Feekes, 1982; Janik, 2000; Rendell & Whitehead, 2003; Tyack, 2008; Vehrencamp et al., 2003), but this hypothesis awaits more rigorous phylogenetic comparative tests. Cases of very rapid vocal learning are found in common bottlenose dolphins, orange-fronted conures and galahs, *Eolophus roseicapilla*, species that live in fission–fusion social groups and imitate the vocalizations of new group members (Table 2; Cortopassi & Bradbury, 2006; Janik & Slater, 1998; Scarl & Bradbury, 2009; Walløe et al., 2015). Orange-fronted conures and galahs have also been shown to rapidly converge on more similar contact call variants during the course of a single vocal exchange (Scarl & Bradbury, 2009; Vehrencamp et al., 2003). The matching of call variants allowed by vocal learning can permit individuals to move among social units throughout their lifetime and to flexibly encode both present and future social relationships (Bradbury & Balsby, 2016). Whether this flexibility is a major driver of vocal learning is best tested by comparative studies across taxa that vary in the temporal dynamics of associations. This hypothesis also raises interesting, and currently unanswered, questions about the degree to which these signals remain reliable indicators of social associations if they can be changed so rapidly. One fundamental question is the extent to which species do differ in the rapidity of their vocal learning: captive operant studies in which the reward for vocal modification and the challenge of the vocal target are controlled (e.g. Manabe & Dooling, 1997) could provide important data. Beyond that basic work, determining whether learned calls in species with rapid vocal change are conventional signals will require two very challenging experimental tasks: assessing the costs of learning and the degree of conflict between signallers and receivers.

Constraints on the Evolution of Call Learning

Another outstanding question in the area of vocal learning is, given the potential benefits of vocal learning and the diversity of taxa that show some version of call imitation, why hasn't call learning evolved in every group-living lineage? First, the specialized neural mechanisms that underpin vocal learning (Bolhuis et al., 2010; Chakraborty et al., 2015; Feenders et al., 2008; Jarvis, 2004) may be difficult to evolve (Chakraborty & Jarvis, 2015; Isler & Schaik, 2006; Mink et al., 1981). Second, the functional costs of learning processes, including time and social retaliation for making errors (Akçay et al., 2009; Smith et al., 2000), can be

avoided if unlearned vocalizations are sufficient for mediating social dynamics. When species live in social groups that are small, stable and/or genetically homogenous, then nonlearned calls, or calls learned during a single critical period, can function to mediate social interactions within and among groups without these associated costs (Seyfarth & Cheney, 2014). It may be that only large, dynamic social groups make it sufficiently challenging for individuals to remember and recognize individually distinctive vocalizations for vocal learning of shared calls to be a more beneficial alternative. Future work examining the relative costs of comprehension and production learning will help determine which factors serve as evolutionary constraints on vocal learning. Both modelling and comparative approaches would be particularly useful in addressing this key area of research. Another fruitful approach could be molecular manipulations of the neural plasticity that underlies vocal learning coupled with examination of the social consequences for individuals with either enhanced or diminished learning capacities.

Relationship between Call Learning and Human Speech Learning

Call learning is a useful model for studying the evolutionary origins, social contexts and proximate mechanisms of speech learning in humans because calls share some key features with language. For example, calls have been proposed to serve as referential signals (Herman, 2006; Janik & Slater, 2000; Templeton et al., 2005; Wanker et al., 2005; Watson et al., 2015), call learning is associated with social complexity and cooperation (Tyack, 2008), and learning can be socially motivated (Farabaugh et al., 1994; Freeberg et al., 2012; Manabe & Dooling, 1997; Sewall, 2009). Additionally, call imitation and convergence in animals are analogous to the process of human vocal accommodation – the imitation of speech prosody, intonation and cadence (Giles et al., 1991; Snowdon & Elowson, 1999; Tyack, 2008; Vehrencamp et al., 2003). This similarity is especially striking in the case of rapid vocal convergence seen within the course of a single communicative interaction in some parrots (Scarl & Bradbury 2009; Balsby & Bradbury, 2009). Finally, because call learning often occurs in some capacity throughout life, it has parallels with human adult vocal learning in the context of second-language learning in immigrants and language relearning in stroke victims; studying strictly closed-ended song learners provides limited insight into these processes. Additional support for call learning as a model for human speech learning would include evidence of contextual learning of imitated calls (though see Herman, 2006; King & Janik, 2013; Wanker et al., 2005) and reports of the vocal learning of calls other than contact calls (such as alarm and food calls; but see Goodale & Kotagama, 2006; Watson et al., 2015). Further work on the early ontogeny of call learning could also identify parallels with early speech learning in human children.

Neural Mechanisms of Call Learning

Conserved motor circuits underlie the vocal production of song in birds and have parallels with the circuits controlling speech production in humans (Bolhuis et al., 2010; Doupe & Kuhl, 1999; Jarvis, 2004). Similarly, specialized brain regions are involved in the auditory processing and recognition of songs and these are analogous to mammalian brain regions (Bolhuis et al., 2010; Chew et al., 1996; Terpstra et al., 2006). These similarities in brain mechanisms have helped propel research on song learning in birds. While fewer studies have examined mechanisms of call learning, there is evidence that some key genes, particularly

FoxP2, play similar roles in promoting song and call learning in birds (Hara et al., 2015; Whitney et al., 2015). Moreover, there is evidence that the song control pathway in the brain, which controls song production, also controls the vocal production of unlearned calls (Ter Maat et al., 2014), although, to our knowledge, no studies have yet examined the role of these regions in the production of learned calls in species that learn both calls and song. Similarly, there is some evidence that the brain regions involved in song recognition are also involved in call recognition (Brauth et al., 2002; Eda-Fujiwara et al., 2011), but comparisons of responses to shared and unshared calls have not yet been made. Most importantly, the neural mechanisms of call perception, production and learning in other vocal learning species like bats and cetaceans are even less resolved than are those of birds (Knörnschild, 2014; Petkov & Jarvis, 2012; Stoeger & Manger, 2014). Proximate studies of the brain mechanisms involved in call production learning and call recognition will inform our thinking about the evolution of call learning, its current function, and the relevance to human speech.

Conclusions

Consideration of the past and current selection pressures driving vocal learning informs our understanding of the social factors important for the development of species-typical communication and the qualities of learned vocalizations that are important for effective communication. Call learning is particularly widespread in animals, is associated with social complexity, and occurs in individuals of all ages and sexes. For these reasons we suggest that a focus on the function and mechanisms of call learning across the broad range of taxa may provide important insights into the evolution of vocal learning, and even of human language. Future studies investigating the origins of vocal learning should focus on testing alternatives to sexual selection through comparative approaches, modelling and focused experiments that assess the costs and benefits of learned calls in dynamic social systems.

Acknowledgments

We thank J. Bradbury, T. P. Hahn, P. R. Marler, C. Toft, W. A. Searcy and members of the Wright, Sewall and Young Labs for discussion of these ideas. The manuscript was greatly improved by the thoughtful comments of the anonymous referees and editor G. Patricelli. T. Wright and A. Young's work on call learning in budgerigars is supported by National Institutes of Health grant 9SC1GM112582. Work in the Sewall lab is supported by an award from the Jeffress Memorial Trust.

References

- Akçay Ç, Wood WE, Searcy WA, Templeton CN, Campbell SE, Beecher MD. Good neighbour, bad neighbour: song sparrows retaliate against aggressive rivals. *Anim. Behav.* 2009; 78:97–102.
- Ali NJ, Farabaugh S, Dooling R. Recognition of contact calls by the budgerigar (*Melopsittacus undulatus*). *Bull. Psychonomic Soc.* 1993; 31(5):468–470.
- Aubin T, Jouventin P. How to vocally identify kin in a crowd: The penguin model. *Adv. Study Behav.* 2002; 31:243–277.
- Balsby T, Bradbury J. Vocal matching by orange-fronted conures (*Aratinga canicularis*). *Behav. Proc.* 2009; 82:133–139.
- Balsby TJS, Momberg JV, Dabelsteen T. Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network. *Plos One.* 2012; 7:8.
- Barros KS, Tokumaru RS, Pedroza JP, Nogueira SSC. Vocal repertoire of captive capybara (*Hydrochoerus hydrochaeris*): structure, context and function. *Ethology.* 2011; 117:83–93.

- Bergman TJ, Beehner JC. Measuring social complexity. *Anim. Behav.* 2015; 103:203–209.
- Blumstein DT, Armitage KB. Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *Am. Nat.* 1997; 150:179–200. [PubMed: 18811281]
- Bolhuis JJ, Okanoya K, Scharff C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci.* 2010; 11:747–759. [PubMed: 20959859]
- Boogert NJ, Anderson RC, Peters S, Searcy WA, Nowicki S. Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Anim. Behav.* 2011; 81:1209–1216.
- Boughman JW. Vocal learning by greater spear-nosed bats. *Proc. R. Soc. B Biol. Sci.* 1998; 265:227–233.
- Bradbury JW, Balsby TJS. The functions of vocal learning in parrots. *Behav. Ecol. Sociobiol.* 2016; 70(3):293–312.
- Bradbury, JW., Vehrencamp, SL. Principles of animal communication. Sunderland, MA: Sinauer; 1998.
- Brauth S, Liang W, Roberts TF, Scott LL, Quinlan EM. Contact call-driven zenk protein induction and habituation in telencephalic auditory pathways in the budgerigar (*Melopsittacus undulatus*): implications for understanding vocal learning processes. *Learn. Mem.* 2002; 9:76–88. [PubMed: 11992018]
- Burling, R. The talking ape: How language evolved. Oxford, U.K: Oxford University Press; 2007.
- Busnel RG, Mebes HD. Hearing and communication in birds: cocktail-party-effect in intraspecific communication of *Agapornis roseicollis* (Aves, Psittacidae). *Life Sci.* 1975; 17:1567–1570. [PubMed: 1207375]
- Byrne, R., Whiten, A. Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans. Oxford, U.K: Oxford University Press; 1989.
- Catchpole, CK., Slater, PJB. Bird song: Biological themes and variations. Cambridge, U.K: Cambridge University Press; 1995.
- Chakraborty M, Jarvis ED. Brain evolution by brain pathway duplication. *Phil. Trans. Roy. Soc. B-Biol. Sci.* 2015; 370:20150056.
- Chakraborty M, Walloe S, Nedergaard S, Fridel EE, Dabelsteen T, Pakkenberg B, et al. Core and shell song systems unique to the parrot brain. *Plos One.* 2015; 10:e0118496. [PubMed: 26107173]
- Cheney DL, Seyfarth RM, Silk JB. The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning? *J. Comp. Psychol.* 1995; 109:134–141. [PubMed: 7758289]
- Chew SJ, Vicario DS, Nottebohm F. A large-capacity memory system that recognizes the calls and songs of individual birds. *Proc. Natl. Acad. Sci. U. S. A.* 1996; 93:1950–1955. [PubMed: 8700865]
- Cortopassi KA, Bradbury JW. Contact call diversity in wild orange-fronted parakeet pairs, *Aratinga canicularis*. *Anim. Behav.* 2006; 71:1141–1154.
- Dahlin CR, Young AM, Cordier B, Mundry R, Wright TF. A test of multiple hypotheses for the function of call sharing in female budgerigars, *Melopsittacus undulatus*. *Behav. Ecol. Sociobiol.* 2014; 68:145–161. [PubMed: 24860236]
- Deecke VB, Barrett-Lennard LG, Spong P, Ford JKB. The structure of stereotyped calls reflects kinship and social affiliation in resident killer whales (*Orcinus orca*). *Naturwissenschaften.* 2010; 97:513–518. [PubMed: 20213302]
- Dooling RJ. Perception of vocal signals by budgerigars (*Melopsittacus undulatus*). *Exp Biol.* 1986; 45:195–218. [PubMed: 3732497]
- Doupe AJ, Kuhl PK. Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* 1999; 22:567–631. [PubMed: 10202549]
- Dunbar RIM. Neocortex size as a constraint on group-size in primates. *J. Human Evol.* 1992; 22:469–493.
- Dunbar RIM. The social brain: Mind, language, and society in evolutionary perspective. *Annu. Rev. Anthropol.* 2003; 32:163–181.

- Eda-Fujiwara H, Kanesada A, Okamoto Y, Satoh R, Watanabe A, Miyamoto T. Long-term maintenance and eventual extinction of preference for a mate's call in the female budgerigar. *Anim. Behav.* 2011; 82:971–979.
- Elgar MA. Predator vigilance and group size in mammals and birds: A critical review of the empirical evidence. *Biol. Rev.* 1989; 64:13–33. [PubMed: 2655726]
- Endler JA, Basolo AL. Sensory ecology, receiver biases and sexual selection. *Trends Ecol. Evol.* 1998; 13:415–420. [PubMed: 21238370]
- Farabaugh SM, Linzenbold A, Dooling RJ. Vocal plasticity in budgerigars (*Melopsittacus undulatus*): evidence for social factors in the learning of contact calls. *J. Comp. Psychol.* 1994; 108:81–92. [PubMed: 8174348]
- Feekes F. Song mimesis within colonies of *Cacicus c. cela* (Icteridae, Aves). A colonial password? *Z. für Tierpsychol.* 1982; 58:119–152.
- Feenders G, Liedvogel M, Rivas M, Zapka M, Horita H, Hara E, et al. Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. *PLoS One.* 2008; 3:e1768. [PubMed: 18335043]
- Fitch WT. The evolution of language: A comparative review. *Biol. Philos.* 2005; 20:193–203.
- Fitch WT, Huber L, Bugnyar T. Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron.* 2010; 65:795–814. [PubMed: 20346756]
- Ford JKB. Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.* 1991; 69:1454–1483.
- Freeberg TM, Dunbar RIM, Ord TJ. Social complexity as a proximate and ultimate factor in communicative complexity. *Philos. Trans. R. Soc. B Biol. Sci.* 2012; 367:1785–1801.
- Giles, H., Coupland, J., Coupland, N. Accommodation theory: Communication, context, and consequence. In: Giles, H., Coupland, J., Coupland, N., editors. *Contexts of accommodation*. Cambridge, U.K: Cambridge University Press; 1991. p. 1–68.
- Goodale E, Kotagama SW. Vocal mimicry by a passerine bird attracts other species involved in mixed-species flocks. *Anim. Behav.* 2006; 72:471–477.
- Groth JG. Call matching and positive assortative mating in red crossbills. *Auk.* 1993; 110:398–401.
- Guilford T, Dawkins MS. Receiver psychology and the evolution of animal signals. *Anim. Behav.* 1991; 42:1–14.
- Hara E, Perez JM, Whitney O, Chen Q, White SA, Wright TF. Neural *FoxP2* and *FoxP1* expression in the budgerigar, an avian species with adult vocal learning. *Behav. Brain Res.* 2015; 283:22–29. [PubMed: 25601574]
- Herman, LM. Intelligence and rational behaviour in the bottlenosed dolphin. In: Hurley, S., Nudds, M., editors. *Rational animals?*. New York, NY: Oxford University Press; 2006. p. 439–467.
- Hile AG, Striedter GF. Call convergence within groups of female budgerigars (*Melopsittacus undulatus*). *Ethology.* 2000; 106:1105–1114.
- Hile Plummer, Striedter. Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Anim. Behav.* 2000; 59:1209–1218. [PubMed: 10877900]
- Hile AG, Burley NT, Coopersmith CB, Foster VS, Striedter GF. Effects of male vocal learning on female behavior in the budgerigar, *Melopsittacus undulatus*. *Ethology.* 2005; 111:901–923.
- Insley SJ. Mother-offspring vocal recognition in northern fur seals is mutual but asymmetrical. *Anim. Behav.* 2001; 61:129–137. [PubMed: 11170703]
- Isler K, Schaik CP, van. Metabolic costs of brain size evolution. *Biol. Lett.* 2006; 2:557–560. [PubMed: 17148287]
- Jackendoff R. Possible stages in the evolution of the language capacity. *Trends Cogn. Sci.* 1999; 3:272–279. [PubMed: 10377542]
- Janik VM. Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science.* 2000; 289:1355–1357. [PubMed: 10958783]
- Janik VM. Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* 2014; 28:60–65. [PubMed: 25057816]
- Janik VM, Sayigh LS. Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A.* 2013; 199:479–489.

- Janik VM, Slater PJB. Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Anim. Behav.* 1998; 56:829–838. [PubMed: 9790693]
- Janik VM, Slater PJB. The different roles of social learning in vocal communication. *Anim. Behav.* 2000; 60:1–11. [PubMed: 10924198]
- Janik, V., Slater, PB. Vocal learning in mammals. In: Peter, JSR.Slater, JB.Snowdon, CT., Milinski, M., editors. *Advances in the study of behavior*. London, U.K: Academic Press; 1997. p. 59-99.
- Jarvis ED. Learned birdsong and the neurobiology of human language. *Ann. N. Y. Acad. Sci.* 2004; 1016:749–777. [PubMed: 15313804]
- Jolly A. Lemur social behavior and primate intelligence. *Science.* 1966; 153:501–506. [PubMed: 5938775]
- King SL, Janik VM. Bottlenose dolphins can use learned vocal labels to address each other. *Proc. Natl. Acad. Sci.* 2013; 110:13216–13221. [PubMed: 23878217]
- Kirschel ANG, Blumstein DT, Cohen RE, Buermann W, Smith TB, Slabbekoorn H. Birdsong tuned to the environment: green hylia song varies with elevation, tree cover, and noise. *Behav. Ecol.* 2009; 20:1089–1095.
- Knörnschild M. Vocal production learning in bats. *Curr. Opin. Neurobiol.* 2014; 28:80–85. [PubMed: 25050812]
- Kober M, Trillmich F, Naguib M. Vocal mother-offspring communication in guinea pigs: females adjust maternal responsiveness to litter size. *Front. Zool.* 2008; 5:13. [PubMed: 18783602]
- Kondo N, Watanabe S. Contact calls: information and social function. *Jap. Psychol. Res.* 2009; 51:197–208.
- Mammen DL, Nowicki S. Individual differences and within-flock convergence in chickadee calls. *Behav. Ecol. Sociobiol.* 1981; 9:179–186.
- Manabe K, Dooling RJ. Control of vocal production in budgerigars (*Melopsittacus undulatus*): selective reinforcement, call differentiation, and stimulus control. *Behav. Processes.* 1997; 41:117–132. [PubMed: 24896668]
- Margoliash D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* 1983; 3:1039–1057. [PubMed: 6842281]
- Marler, P. Bird calls: a cornucopia for communication. In: Marler, P., Slabbekoorn, H., editors. *Nature's music: The science of birdsong*. San Diego, CA: Elsevier Academic Press; 2004. p. 132-176.
- Marler, P., Slabbekoorn, HW. *Nature's music: the science of birdsong*. San Diego, CA: Elsevier Academic Press; 2004.
- Maynard Smith, J., Harper, D. *Animal signals*. Oxford, U.K: Oxford University Press; 2003.
- McComb K, Semple S. Coevolution of vocal communication and sociality in primates. *Biol. Lett.* 2005; 1:381–385. [PubMed: 17148212]
- Miller, G. *The mating mind: How sexual choice shaped the evolution of human nature*. New York, NY: Doubleday; 2000.
- Miller PJO, Shapiro AD, Tyack PL, Solow AR. Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Anim. Behav.* 2004; 67:1099–1107.
- Mink JW, Blumenshine RJ, Adams DB. Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. *Am. J. Physiol.* 1981; 241:R203–R212. [PubMed: 7282965]
- Miyagawa S, Ojima S, Berwick RC, Okanoya K. The integration hypothesis of human language evolution and the nature of contemporary languages. *Lang. Sci.* 2014; 5:564.
- Moravec ML, Striedter GF, Burley NT. Assortative pairing based on contact call similarity in budgerigars, *Melopsittacus undulatus*. *Ethology.* 2006; 112:1108–1116.
- Mundinger PC. Vocal imitation and individual recognition of finch calls. *Science.* 1970; 168:480–482. [PubMed: 17838128]
- Mundinger PC. Call learning in the Carduelinae: ethological and systematic considerations. *Syst. Zool.* 1979; 28:270–283.
- Nottebohm F. The origins of vocal learning. *Am. Nat.* 1972; 106:116–140.

- Nowicki S, Searcy WA, Peters S. Brain development, song learning and mate choice in birds: a review and experimental test of the 'nutritional stress hypothesis'. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol.* 2002; 188:1003–1014. [PubMed: 12471497]
- Nowicki S, Searcy WA. Song function and the evolution of female preferences: why birds sing, why brains matter. *Ann. N. Y. Acad. Sci.* 2004; 1016:704–723. [PubMed: 15313801]
- Nowicki S, Searcy WA. The evolution of vocal learning. *Curr. Opin. Neurobiol.* 2014; 28:48–53. [PubMed: 25033109]
- Oda R. Individual distinctiveness of the contact calls of ring-tailed lemurs. *Folia Primatol.* 2002; 73:132–136. [PubMed: 12207060]
- Odom KJ, Hall ML, Riebel K, Omland KE, Langmore NE. Female song is widespread and ancestral in songbirds. *Nat. Commun.* 2014:5.
- Petkov CI, Jarvis E. Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 2012; 4:12. [PubMed: 22912615]
- Pinker S. The cognitive niche: Coevolution of intelligence, sociality, and language. *Proc. Natl. Acad. Sci.* 2010:200914630.
- Prather JF, Peters S, Nowicki S, Mooney R. Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature.* 2008; 451:305–310. [PubMed: 18202651]
- Price JJ. Family- and sex-specific vocal traditions in a cooperatively breeding songbird. *Proc. Biol. Sci.* 1998; 265:497–502.
- Puts DA, Hodges CR, Cárdenas RA, Gaulin SJC. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* 2007; 28:340–344.
- Reichmuth C, Casey C. Vocal learning in seals, sea lions, and walruses. *Curr. Opin. Neurobiol.* 2014; 28:66–71. [PubMed: 25042930]
- Rendall D, Rodman PS, Emond RE. Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Anim. Behav.* 1996; 51:1007–1015.
- Rendell LE, Whitehead H. Vocal clans in sperm whales (*Physeter macrocephalus*). *Proc. Biol. Sci.* 2003; 270:225–231. [PubMed: 12614570]
- Salinas-Melgoza A, Wright TF. Evidence for vocal learning and limited dispersal as dual mechanisms for dialect maintenance in a parrot. *PLoS One.* 2012; 7:e48667. [PubMed: 23139809]
- Scarl JC, Bradbury JW. Rapid vocal convergence in an Australian cockatoo, the galah *Eolophus roseicapillus*. *Anim. Behav.* 2009; 77:1019–1026.
- Searcy WA, Andersson M. Sexual selection and the evolution of song. *Ann. Rev. Ecol. Syst.* 1986; 17:507–533.
- Searcy, WA., Nowicki, S. The evolution of animal communication: Reliability and deception in signaling systems. Princeton, NJ: Princeton University Press; 2005.
- Sewall KB. Limited adult vocal learning maintains call dialects but permits pair-distinctive calls in red crossbills. *Anim. Behav.* 2009; 77:1303–1311.
- Sewall KB. Early learning of discrete call variants in red crossbills: implications for reliable signaling. *Behav. Ecol. Sociobiol.* 2011; 65:157–166.
- Sewall KB. Social complexity as a driver of communication and cognition. *Integr. Comp. Biol.* 2015; 55(3):384–395. 55. [PubMed: 26078368]
- Seyfarth RM, Cheney DL. The evolution of language from social cognition. *Curr. Opin. Neurobiol.* 2014; 28:5–9. [PubMed: 24813180]
- Smith JW, Benkman CW, Coffey K. The use and misuse of public information by foraging red crossbills. *Behav. Ecol.* 1999; 10:54–62.
- Smith VA, King AP, West MJ. A role of her own: female cowbirds, *Molothrus ater*, influence the development and outcome of song learning. *Anim. Behav.* 2000; 60:599–609. [PubMed: 11082230]
- Smolker R, Pepper JW. Whistle convergence among allied male bottlenose dolphins (Delphinidae, *Tursiops* sp.). *Ethology.* 1999; 105:595–617.
- Snowdon CT, Elowson AM. Pygmy marmosets modify call structure when paired. *Ethology.* 1999; 105:893–908.

- Sousa-Lima RS, Paglia AP, Da Fonseca GAB. Signature information and individual recognition in the isolation calls of Amazonian manatees, *Trichechus inunguis* (Mammalia: Sirenia). *Anim. Behav.* 2002; 63:301–310.
- Stoeger AS, Manger P. Vocal learning in elephants: neural bases and adaptive context. *Curr. Opin. Neurobiol.* 2014; 28:101–107. [PubMed: 25062469]
- Sugiura H. Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim. Behav.* 1998; 55:673–687. [PubMed: 9514664]
- Templeton CN, Greene E, Davis K. Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science.* 2005; 308:1934–1937. [PubMed: 15976305]
- Ter Maat A, Trost L, Sagunsky H, Seltmann S, Gahr M. Zebra finch mates use their forebrain song system in unlearned call communication. *PLoS One.* 2014; 9:e109334. [PubMed: 25313846]
- Terpstra NJ, Bolhuis JJ, Riebel K, van der Burg JMM, den Boer-Visser AM. Localized brain activation specific to auditory memory in a female songbird. *J. Comp. Neurol.* 2006; 494:784–791. [PubMed: 16374807]
- Theunissen FE, Amin N, Shaevitz SS, Woolley SMN, Fremouw T, Hauber ME. Song selectivity in the song system and in the auditory forebrain. *Ann. N. Y. Acad. Sci.* 2004; 1016:222–245. [PubMed: 15313778]
- Tobias JA, Aben J, Brumfield RT, Derryberry EP, Halfwerk W, Slabbekoorn H, et al. Song divergence by sensory drive in Amazonian birds. *Evolution.* 2010; 64:2820–2839. [PubMed: 20561048]
- Toft, CA., Wright, TF. *Parrots of the wild: A natural history of the world's most captivating birds.* Berkeley, CA: University of California Press; 2015.
- Townsend SW, Hollén LI, Manser MB. Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. *Anim. Behav.* 2010; 80:133–138.
- Tyack PL. Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *J. Comp. Psychol.* 2008; 122:319–331. [PubMed: 18729661]
- Vehrencamp SL, Ritter AF, Keever M, Bradbury JW. Responses to playback of local vs. distant contact calls in the orange-fronted conure, *Aratinga canicularis*. *Ethology.* 2003; 109:37–54.
- Walløe S, Thomsen H, Balsby TJ, Dabelsteen T. Differences in short-term vocal learning in parrots, a comparative study. *Behaviour.* 2015; 152:1433–1461.
- Wanker R, Fischer J. Intra- and interindividual variation in the contact calls of spectacled parrotlets (*Forpus conspicillatus*). *Behaviour.* 2001; 138:709.
- Wanker R, Apcin J, Jennerjahn B, Waibel B. Discrimination of different social companions in spectacled parrotlets (*Forpus conspicillatus*): Evidence for individual vocal recognition. *Behav. Ecol. Sociobiol.* 1998; 43:197–202.
- Wanker R, Sugama Y, Prinage S. Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. *Anim. Behav.* 2005; 70:111–118.
- Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, et al. Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* 2015; 25:495–499. [PubMed: 25660548]
- Whitney O, Voyles T, Hara E, Chen Q, White SA, Wright TF. Differential *FoxP2* and *FoxP1* expression in a vocal learning nucleus of the developing budgerigar. *Dev. Neurobiol.* 2015; 75:778–790. [PubMed: 25407828]
- Wilbrecht L, Nottebohm F. Vocal learning in birds and humans. *Ment. Retard. Dev. Disabil. Res. Rev.* 2003; 9:135–148. [PubMed: 12953292]
- Wilkinson GS, Boughman JW. Social calls coordinate foraging in greater spear-nosed bats. *Anim. Behav.* 1998; 55:337–350. [PubMed: 9480702]
- Wright TF. Regional dialects in the contact call of a parrot. *Proc. R. Soc. Lond. B Biol. Sci.* 1996; 263:867–872.
- Yurk H. Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Anim. Behav.* 2002; 63:1103–1119.

Highlights

- Learning of social calls is more common among birds and mammals than song learning.
- Unlike song learning, call learning is not well explained by sexual selection.
- We examine hypotheses for the origin and maintenance of call learning.
- Learned calls give new insight into the evolution and mechanisms of vocal learning.

Table 1

Definitions of forms of vocal learning

Type of learning	Definition
Production learning	The ability to acquire signal variants through a process of social experience and auditory feedback either by modifying existing sounds for which individuals may have some innate template, or by copying of entirely novel sounds; often simplified to 'vocal learning'
Contextual/usage learning	The ability to change a pattern of usage of an existing signal based on experience
Comprehension learning	The ability to learn to display appropriate behaviours in response to hearing specific signal variants

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Examples of species in which vocal production learning of conspecific contact calls has been documented with evidence that extends beyond possessing a shared call, and the social context with which it is associated

Order	Species	Level of social organization with shared call	Evidence of call learning	Social context	Reference
Psittaciformes	Budgerigar, <i>Melopsittacus undulatus</i>	Pair	Imitation over time	Mate choice	Moravec et al. 2006
				Pair bond formation or mate choice	Hile et al. 2000, Hile et al. 2005
		Group	Convergence over time	Social cohesion	Hile and Striedter 2000
			Imitation or convergence over time	Mediating social associations	Dahlin et al. 2014
				Social cohesion	Farabaugh et al. 1994
	Galah, <i>Eolophus roseicapilla</i>	Interaction	Rapid convergence during social interactions	Mediating social associations	Scarl and Bradbury 2009
	Orange-fronted conure, <i>Eupsittula (formerly Aratinga) canicularis</i>	Interaction	Rapid convergence during social interactions	Mediating social associations	Vehrencamp et al. 2003, Balsby and Bradbury 2009
	Yellow-naped amazon, <i>Amazona auropalliata</i>	Populations	Imitation over time	Social integration	Salinas-Melgoza and Wright 2012
Passeriformes	European siskin, <i>Spinus spinus</i>	Pair	Imitation over time	Social integration	Mundinger 1970
	Red crossbill, <i>Loxia curvirostra</i>	Family	Imitation over time	Family cohesion	Sewall 2011
		Pair	Convergence over time	Social affiliation	Sewall 2009

Order	Species	Level of social organization with shared call	Evidence of call learning	Social context	Reference
	Black-capped chickadee, <i>Parus atricapillus</i>	Group	Convergence over time	Social cohesion during cooperative foraging and territory defence	Mammen and Nowicki 1981
Chiroptera	Greater spear-nosed bat, <i>Phyllostomus hastatus</i>	Group	Convergence over time	Group badge and cohesion during cooperative foraging and territory defence	Boughman 1998
Cetacea	Bottlenose dolphins, <i>Tursiops truncatus</i>	Group Group	Convergence over time	Group badge	Smolker and Pepper 1999
			Matching and imitation	Social cohesion	Janik and Slater 1998
	Orca, <i>Orcinus orca</i>	Family	Divergence over time	Changing social affiliation	Ford 1991
		Family	Matching	Group-calling bouts	Miller et al. 2004

Table 3

Hypotheses proposed to explain the evolutionary origins of call learning

Hypothesis	Description
Improved signal recognition	Imitating signals results in enhanced response or faster processing time by receivers
Signalling local familiarity	Learners can move among populations with local dialects
Signalling affiliation and group membership	Learners can move among different social groups and/or group members signal cohesion to nongroup members
Increased information encoding/complexity	Learned calls permit the encoding of multiple levels of social affiliation ranging from pairs and families to populations

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Potential factors contributing to the evolutionary origin and maintenance of call learning across taxa

Hypothesis	Benefit	Predictions	Conditions for origin of vocal learning	Summary
Improved signal recognition	More rapid or robust communication between senders and receivers	Receivers respond more strongly or quickly to imitations of their own signals; imitated signals occur most commonly in noisy social environments	Receivers must have specialized neural circuitry to permit detection of calls like their own	Improved signal recognition is more likely to have supported other hypotheses than to have driven vocal learning independently
Signalling local familiarity	Learners benefit from the ability to move among populations; receivers only accept group members, or prefer mates, who have been in the area long enough to know the local dialects and thus have knowledge of local resources	Vocal learning is most common in lineages that have local dialects; receivers discriminate among dialects	Unlearned dialects must have preceded vocal learning	In the context of call learning this hypothesis is difficult to distinguish from the 'signalling affiliation' hypothesis and is not well supported in the literature
Signalling affiliation and group membership	Learners benefit from the ability to move among social groups; group members benefit from better coordination of group efforts/decisions; potential to improve cooperative defence of resources against competing groups	Vocal learning is most common in lineages with very large or fission–fusion groups and/or competition among social groups; receivers discriminate among group-specific calls and regulate access to group benefits based on call	Unlearned calls shared among group members must have preceded vocal learning	At present this hypothesis is the most strongly supported based on empirical studies and informal phylogenetic surveys
Increased information encoding	Learners can signal affiliation across multiple social levels and therefore move through complex and dynamic social groups; receivers can identify and associate members of different social units even if they are not familiar with an individual	Vocal learning is most common in lineages in which individuals have multiple social demographic memberships and tiered social structures; levels of call similarity are linked to social demography; receivers recognize the demographic status of companions based on calls	Some form of social complexity must have preceded vocal learning and used unlearned calls to reflect social grouping	Increased information encoding is linked to the 'signalling affiliation' hypothesis both functionally and conceptually in the context of contact calls; this hypothesis could be tested independently in the context of food or predator/alarm calls