

The neurobiology of vocal communication in marmosets

Dori M. Grijseels¹  | Brendan J. Prendergast¹ | Julia C. Gorman^{1,2} | Cory T. Miller^{1,2}

¹Cortical Systems and Behavior Laboratory, University of California, San Diego, La Jolla, California, USA

²Neurosciences Graduate Program, University of California, San Diego, La Jolla, California, USA

Correspondence

Dori M. Grijseels, Cortical Systems and Behavior Laboratory, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.
Email: dogrijseels@ucsd.edu

Funding information

Common Fund, Grant/Award Number: R01 DC012087

Abstract

An increasingly popular animal model for studying the neural basis of social behavior, cognition, and communication is the common marmoset (*Callithrix jacchus*). Interest in this New World primate across neuroscience is now being driven by their proclivity for prosociality across their repertoire, high volubility, and rapid development, as well as their amenability to naturalistic testing paradigms and freely moving neural recording and imaging technologies. The complement of these characteristics set marmosets up to be a powerful model of the primate social brain in the years to come. Here, we focus on vocal communication because it is the area that has both made the most progress and illustrates the prodigious potential of this species. We review the current state of the field with a focus on the various brain areas and networks involved in vocal perception and production, comparing the findings from marmosets to other animals, including humans.

KEYWORDS

Callithrix jacchus, neurobiology, neuroethology, social behavior, vocal communication

INTRODUCTION

Communication is integral to social living. Despite considerable diversity in gregariousness and social organization, all animals interact with conspecifics. And to do so, they must communicate with each other. As such, selection for effective communication systems, and the supporting neural circuits, has been a dominant force in the evolution of all species, including human and non-human primates (NHPs). The common marmoset (*Callithrix jacchus*)—an increasingly popular primate model in neuroscience^{1–3}—is emerging as a keystone species for elucidating the neurobiology of natural primate vocal communication. Though a key initial factor in driving the recent surge in marmoset interest was their apparent amenability for developing gene editing technologies in a primate,^{4–7} unique facets of the species' social, cognitive, and communication systems are increasingly motivating research because of their many shared properties with humans.²

These New World primates (Figure 1A) exhibit a unique repertoire of social behaviors,^{8,9} including cooperative breeding^{10–14} and com-

plex vocal behaviors, similar to humans (for review, see Miller et al.²). Like all primates, the functional organization of the marmoset brain is relatively conserved compared to the human brain,^{15,16} further illustrating the species' suitability to model aspects of the human social cognitive faculty that, though shared with other primates, are absent in other mammalian models.² As such, the marmoset has been proposed as a translational model to study speech disorders as well as speech-related aspects of neurodegenerative disorders (see Okano¹⁷ for a review),^{17–19} which is further aided by an increasing availability of genetic tools.^{4,5,18,20} Explicating the neural basis of vocal communication in marmosets will be integral to maximizing the translational significance of this powerful model organism.

Here, we detail current knowledge about the neural basis of vocal communication in marmosets to illustrate the increasingly powerful role that this species is now playing in elucidating primate social brain functions. Though focused on marmosets, this review draws upon evidence from several model species^{21–23} to contextualize the findings. While most comparisons will be to the rhesus macaque (*Macaca*

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of New York Academy of Sciences.

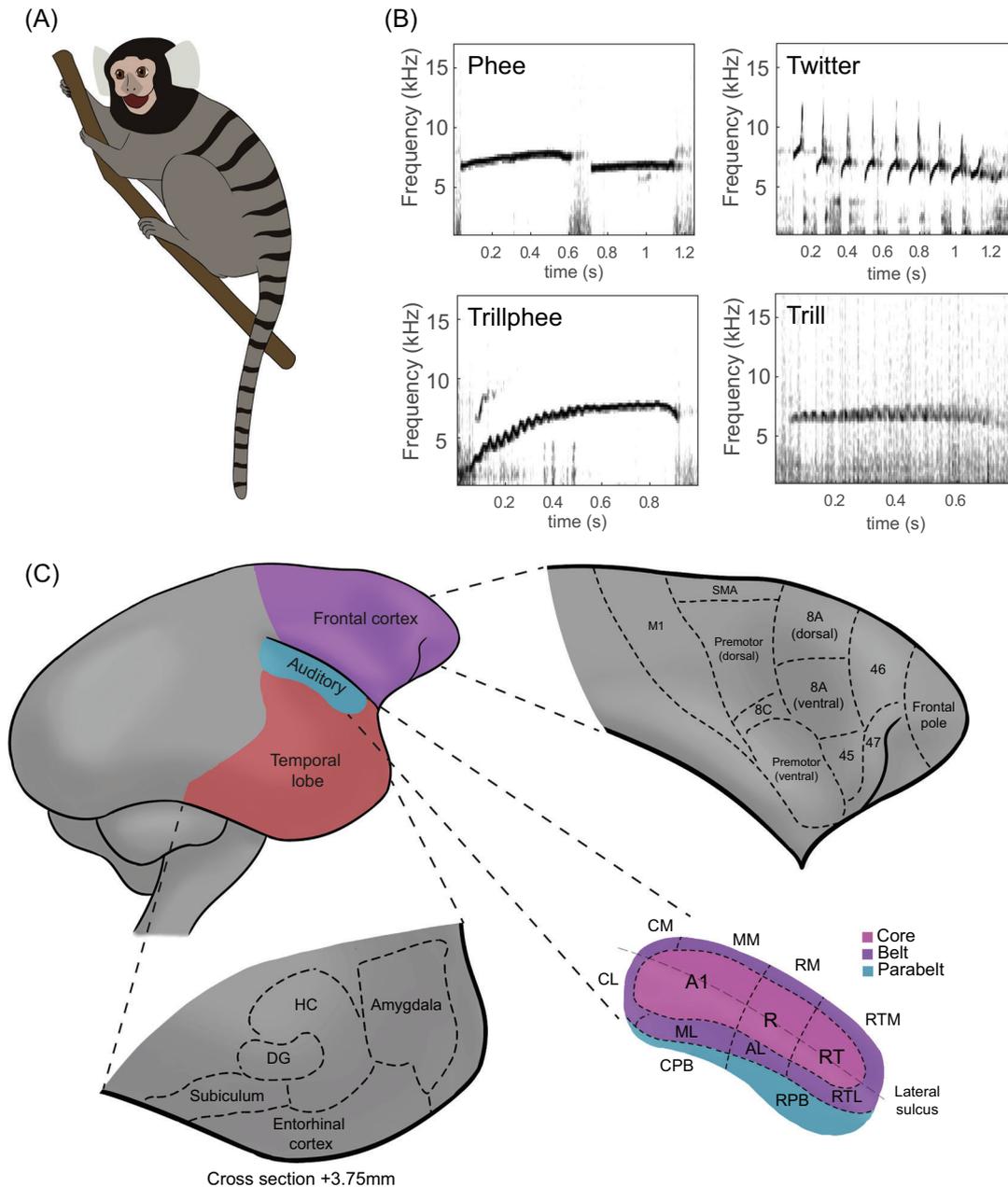


FIGURE 1 Vocal behavior and neurobiology of marmosets. (A) Graphical depiction of a marmoset. (B) Representative spectrograms of the four most common call types are shown, recorded from a freely moving marmoset performing natural behaviors using a wearable microphone. (C) Schematic of the brain areas involved in vocal communication in marmosets. Top left is sagittal view of a marmoset brain, with the frontal cortex, auditory cortex, and temporal lobe highlighted. Top right shows a detailed view of the frontal cortex (after Refs. 37,232), with areas expressed either in common name (where applicable), or Brodmann area number. Areas 45 and 47 are part of the ventrolateral prefrontal cortex and area 46 is part of the dorsolateral prefrontal cortex. Bottom left shows a detailed view of a sagittal cross-section of the medial temporal cortex at 3.75 mm from the midline (after Refs. 232,233). Bottom right shows a detailed view of the auditory cortex (after Ref. 76), with colors indicating core (pink), belt (purple), and parabelt (blue). Abbreviations: A1, primary auditory cortex; AL, anterolateral; CL, caudolateral; CM, caudomedial; CPB, caudal parabelt; DG, dentate gyrus; HC, hippocampus (includes Cornu Ammonis subfields CA1, CA2, and CA3); M1, primary motor area; ML, middle lateral; MM, middle medial; R, rostral field; RM, rostromedial; RPB, rostral parabelt; RT, rostromedial; RTL, rostromedial lateral; RTM, rostromedial medial; SMA, supplementary motor area.

mulatta), which is the most widely used NHP model animal, other species will also be highlighted when data are available. First, we will briefly review behavioral studies of marmoset vocal communication that is related to vocal perception and production. Next, we will discuss how various brain regions mediate vocal communication, describing separately literature on vocal signal processing and vocal-motor production. The final section focuses on communication in natural contexts and highlights areas of research needed to better understand the neural basis of communication more holistically.

VOCAL SIGNAL PROCESSING IN MARMOSETS

Marmosets have an extensive repertoire of calls and a dynamic corpus of communication behaviors.^{24–30} The most common types of vocalizations are phee, trill, twitter, and trillphee calls (Figure 1B),^{3,24,27} which are each perceptually distinctive to marmosets.³¹ Phee calls are loud, high-frequency calls, that function to maintain social contact when visually occluded from conspecifics, and to mediate intergroup interactions.²⁵ The acoustic structure of phees can vary based on the social context. Loud, long whistles are used for long-range communication, while shorter whistles are used for closer-ranged contact.^{3,25} Phee calls are the most widely studied call type for marmoset monkeys³ because they are easily elicited in laboratory settings by placing individuals in (visual) isolation and are used in well-established antiphonal calling paradigms.^{2,32–34} In these paradigms, two marmosets, which can hear but not see each other, engage in “conversations”—taking turns producing phee calls. One of the two monkeys may also be replaced by a computer-controlled “virtual” monkey (VM), which allows for manipulating the behavior of the partner in the conversation.^{35–37} Twitters are a second type of long-distance call, which is used both between conspecifics in the same group and during intergroup interactions.^{25,27} Trills are used for short-range contact between conspecifics,^{24,38} with bonded partners trilling at each other most frequently.³⁸ The trillphee call is a combination of a trill and a phee, with the calls usually starting as a trill and ending as a phee.²⁴ Like trills, they are used for close-range interactions. In addition to these call types, marmosets also produce a variety of other calls, such as chirp, ek, tsik, and chatter, with each related to specific social or environmental context, such as the presence of a predator or an unfamiliar human or conspecific.^{24,28}

Studies in both wild and captive marmoset populations emphasize the importance of these vocalizations for mediating a wide range of behavioral interactions.^{3,38–41} For example, simply broadcasting phee calls from an unfamiliar individual to a wild population of marmosets from within their territory elicits robust behavioral responses.⁴⁰ Interactive antiphonal playback paradigms, in which marmosets engage in vocal exchanges with a VM, is a particularly powerful tool to study vocal behavior in laboratories as it affords experimental control over different facets of a naturally occurring vocal interaction.^{33,39,41} Studies using this paradigm have shown that marmosets recognize the identity of conspecific callers in these contexts^{35,39} and modulate their response rate based on the partner's identity and sex.³²

Although trills are less often studied, as they are only produced in close range to a conspecific, recent advances in wearable technology have allowed recordings in contexts where marmosets produce these vocalizations.³⁸ This shows that marmosets are more likely to produce trills in response to calls by their pair-bonded partner, suggesting again they are able to recognize a caller's identity. Together, these behavioral results show that social information encoded in the acoustic structure of vocal signals is perceptually salient to marmosets and behaviorally relevant in a range of social contexts.

BRAIN NETWORKS UNDERLYING VOCAL PERCEPTION

Vocal signal processing in primates, including marmosets, follows the ascending auditory pathway found in all mammals. Like any sound, vocalizations are first encoded by the cochlea, which sends signals through the cochlear nuclei to the superior olivary complex.⁴² The signal ascends to the inferior colliculus, eventually reaching the medial geniculate body (MGB) in the thalamus (Figure 2).^{42–45} Although these subcortical and brainstem regions involved in audition already perform varying levels of signal processing^{46–49} and have received increasing interest in the study of vocal communication,⁵⁰ here we will focus primarily on the upstream cortical regions involved in vocal signal perception. The MGB projections to the auditory cortex (ACx) relay the signal to this main cortical site for the processing of auditory stimuli. While the initial flow of spiking activity from the thalamus predominantly targets A1 before spreading to the belt and parabelt, all these regions share dense bidirectional connectivity, some of which even have thalamorecipient layers of their own.^{44,45,51} The ACx then sends output streams to multiple cortical and subcortical regions, principally to the prefrontal cortex (PFC) and other frontal lobe regions.^{52,53} In fact, auditory projections from downstream tertiary ACx fields comprise a plurality of inputs to the PFC.⁵⁴ Across these connections, a hierarchical pattern of organization exists. The incoming acoustic information is analyzed and outputted to downstream regions containing neurons that often exhibit variably nonlinear and more complex spectrotemporal receptive fields,^{55–57} then more selectively integrates salient features of this information into increasingly invariant perceptual objects in later associative areas.^{58–68}

The fronto-temporal network in the primate brain is crucial for vocal communication⁶⁹ and is organized in two pathways; the postero-dorsal and antero-ventral pathway, which parallel the visual dorsal and ventral pathways (Figure 2).⁶⁴ In primates, the postero-dorsal pathway, which involves connections from the caudal belt of the ACx to the dorsolateral PFC (dlPFC), is thought to be involved in the spatial processing of auditory signals. Experiments testing the antero-ventral pathway, which involves connections from the anterior belt of the ACx to the ventrolateral PFC (vlPFC), suggest it is integral to encoding the individual call types.^{63,64,70} Although convincing evidence for these two pathways has been found in rhesus macaques, relatively little functional work has been done in marmosets. However, given the presence of dual visual streams similar to macaques in marmosets,^{71,72} and the

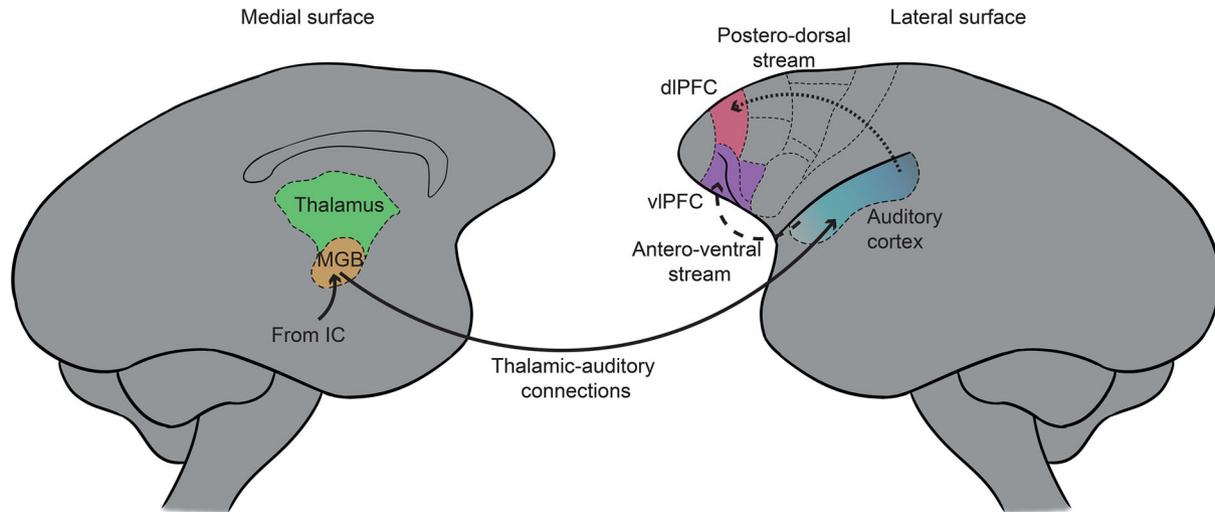


FIGURE 2 Cortical networks for vocal perception. A graphical representation of sensory networks involved in processing vocalizations. The MGB in the thalamus projects to the auditory cortex. From the anterior part of the auditory cortex, projections to the vIPFC are made, which represents the antero-ventral stream, while the posterior part projects onto the dIPFC, which represents the postero-dorsal stream. Abbreviations: dIPFC, dorsolateral prefrontal cortex; IC, inferior colliculus; MGB, medial geniculate body; vIPFC, ventrolateral prefrontal cortex.

differential connectivity of the rostral and caudal belt areas of the auditory cortex to the frontal lobe,⁴⁴ it is highly likely they would possess a similar dual pathway system for auditory processing. Those neurophysiological studies that have been performed in marmosets have tended to focus on the antero-ventral stream and investigated responses to various call types in their repertoire in different neural substrates. Studies on how marmosets perform sound localization have been less prevalent despite the facts that marmosets possess sufficient acuity to perform sound localization experiments,⁷³ localization is an important function of long-range calls in marmosets,²⁵ and initial studies show promising results.^{74,75} Increased focus on the postero-dorsal pathway may elucidate how the fronto-temporal network contributes to vocal communication in marmosets.

As evident from these various neural networks involved in auditory processing, several brain regions work in concert to process conspecific vocalizations. Although these substrates are highly interconnected, so far, most neurophysiology studies have focused on single brain regions. Below, we unpack in more detail how the more well-studied substrates in the marmoset forebrain represent calls to better inform their respective, complementary roles in the broader vocal signal processing network.

Auditory cortex

The ACx is the neocortical substrate located in the temporal cortex that first processes sound in the forebrain, and it plays a crucial role in extracting perceptually salient features of a vocal signal. In all primates, the ACx is organized along the dorsal, medial, and lateral axes in the temporal lobe, most notably along the rostral-caudal axis (Figure 1C).⁷⁶ Its fundamental pattern of organization consists of an elongated core region, where the primary sensory cortex (A1)

is located, and receives the majority of its ascending inputs from the thalamus. The core is encircled by two concentric belt-like areas—the belt and parabelt—each divided into several subareas (Figure 1C).^{52,77} Each region and subregion exhibits functional specialization, with A1 encoding more basic sound properties, while belt and parabelt areas represent increasingly abstract, integrated patterns of auditory activity.^{63,78,79}

Earlier neurophysiological studies of the marmoset auditory cortex reported that both individual neurons and populations were responsive to, or selective for, conspecific vocalizations, including in both primary (A1) and secondary (caudomedial belt) ACx regions.^{80,81} Intriguingly, Zeng et al.⁸² recorded calcium activity of cells in marmoset A1, finding that a subpopulation of A1 neurons exhibited selectivity for certain conspecific call types but not others, and became unresponsive when the call structure was acoustically manipulated. This degree of representational selectivity is somewhat greater than what is typically posited,⁸³ though consistent with early data on marmoset A1 showing response extinction when the vocal stimulus was played in reverse.⁸¹ At present, the dearth of studies examining how vocalizations are represented in the auditory cortex of behaving marmosets is a significant bottleneck in determining whether A1 is capable of more stimulus selectivity and dimensionality than was previously considered.

Recent studies with functional magnetic resonance imaging (fMRI) in marmosets have complemented neurophysiological experiments to investigate how multiple regions of the auditory cortex support vocal signal processing. Using fMRI in anesthetized marmosets, Sadagopan et al.⁸⁴ found a gradient of vocalization-specific responses across the rostral-caudal axis of the ACx, with particularly strong responses near the temporal pole (rostral superior temporal gyrus [STG]), comparable to findings in homologous regions in humans and macaques.^{85–87} Rather than finding a similar gradient, Jafari et al.⁸⁸ showed “voice patches” in awake marmosets, similar to the human voice patch

system,⁸⁹ consisting of three distinct voice areas in the temporal cortex. In addition, they reported a wider functional network of areas—with particular sensitivity to voices or vocalizations—distributed across the marmoset anterior temporal lobe, PFC, anterior cingulate cortex (ACC), and subcortical structures. This network included rostral areas of the ACC, which had not been reported in macaque studies and is generally not included in the human vocal processing network.⁸⁹ Notably, such a network parallels the similar “face patch” system, which has been found across NHPs.^{90–92} Given the integral social salience of both voices and faces to all primates, these parallel specializations of marmoset cortical pathways for processing both of these salient social signals provide a compelling framework for the future study of multimodal social communication in the primate temporal cortex.

Frontal cortex

The auditory cortex sends projections to a variety of brain regions, including the frontal cortex and other temporal regions.⁹³ Although it is clear that auditory regions already represent higher-level acoustic representations, such as species-specific vocalizations, the upstream regions are thought to further process incoming stimuli, for example, determining the identity or location of a caller, and integrating various modalities of information.

Several regions in the frontal cortex share reciprocal connections with the auditory cortex, most notably the dlPFC through the posterodorsal stream and vlPFC through the antero-ventral stream.⁶⁴ It is, therefore, no surprise that the frontal cortex exhibits responses to vocalizations in all primate species that have been tested.^{37,94,95} Neurons in the vlPFC of head-fixed macaques, for example, respond to different types of vocalizations, with most cells responding to a subset of possible call types and only a few responding to a single or all calls.⁹⁵ Although less is known about the representation of vocalizations in the dlPFC, studies using natural sounds do show that this area responds to acoustic stimuli and specifically encodes the spatial location from which they have been emitted.^{96,97} It is, therefore, likely that this area would respond to vocalizations, but the expected degree of vocalization selectivity is not clear. In addition to these two areas, the frontal pole, which is the most anterior portion of the frontal cortex, is also largely innervated by auditory inputs in both macaques and marmosets, suggesting this area may also perform a role in vocalization processing.^{98–100}

In marmosets, several studies have shown that frontal cortex neurons are responsive to vocalizations, though in contrast to macaque experiments, these studies have predominantly been performed in freely moving animals. A cFos study showed significantly increased expression in the ventrolateral area of the PFC during perception of vocalizations compared to production or antiphonal conditions as well as compared to cFos expression in the dorsal premotor cortex during vocal production.^{101,102} Further studies showed stronger evidence of frontal activity in response to vocalizations, both in individual cells and across populations, with vocalization-responsive neurons found in multiple areas of the prefrontal and premotor cortices.^{37,103} Importantly,

neural activity was highly dependent on context: individual PFC neurons and populations responded differently depending on whether the vocalization was heard in a passive-listening context, freely moving and/or in the context of a natural conversation, suggesting that vocalization representations in this substrate are highly sensitive to the behavioral context in which they are heard.^{37,103} Overall, these studies point toward a representation of conspecific vocalizations across frontal regions, but more detailed work in particular subregions is still needed. In addition, studies on more complex representations of vocal stimuli, such as the identity or location of the speaker, are critical to determining the complementary roles of the ventral and dorsal pathways for vocal signal processing in the marmoset frontal cortex.

Subcortical substrates

In addition to the frontal cortex, the ACx also projects to various subcortical structures, suggesting these areas may also be involved in the processing of vocalizations. Indeed, Jafari et al.⁸⁸ report the activation of several subcortical areas in response to hearing vocalizations and suggest these are part of a larger vocal processing network. Other subcortical areas have been implicated in the integration of auditory information across modalities, such as the hippocampus integrating individual identity from faces and voices¹⁰⁴ and the amygdala extracting affective information from the auditory input.^{105,106} However, only a limited number of studies have specifically targeted subcortical substrates in the context of vocalizations in marmosets. Although responses to vocalizations can thus be found in these areas, crucial experiments in the broader context of natural communication behaviors are needed, as we will discuss in more detail.

MECHANISMS FOR VOCAL PRODUCTION IN MARMOSETS

Vocal communication systems are not only optimized for vocal signal processing and perception, but also the production and control of the vocalizations themselves. Species across different vertebrate taxonomic groups, including songbirds, cetaceans, and humans, exert vocal control allowing for flexible and varied vocal production.^{107–111} Humans exhibit perhaps the greatest degree of flexibility and control over vocal emissions and rapidly change speech production in real time in response to auditory feedback.^{112,113} Determining the evolutionary origins of this degree of vocal plasticity in humans, however, has been challenging because of the seemingly dramatic evolutionary gap between humans and NHPs.^{114,115} Historically, vocal production in monkeys has been assumed to be largely innate and stereotyped,¹¹⁶ a view largely based on studies in squirrel monkeys, which exert very little learning and vocal control over the timing and structure of their vocalizations.^{117,118} This view, however, has begun to change as numerous studies now show evidence of vocal control across a variety of primate species.^{21,119–124} In particular, marmosets show plasticity and flexibility to control when,¹²⁵ where,¹²⁶ and what to vocalize.¹²⁷

This suggests that the purported gap between humans and NHPs in capacities for vocal control may not be as dramatic as previously thought, and it opens the door for neurobiological investigations of this system in marmosets to serve as a model for at least some aspects of the parallel system in humans.

Vocal learning and plasticity in marmosets are evident in several naturally occurring social contexts. This is especially true during ontogenetic vocal development, which shows parallels with humans. Marmosets begin producing baby cries, which gradually develop into “babbling” and finally fully mature calls.^{128,129} Parental feedback plays a crucial role in this development.^{129–132} When subadults are separated from their parents, they will continue to produce baby-like cries and babbling vocal behavior into adulthood, which is suppressed in normally raised monkeys.¹³² Even after the initial infant stage, the call structure of individuals continues to mature, with each individual developing unique call sequences.¹²⁸ Not just the content of the calls changes over ontogeny; young marmosets also learn turn-taking during conversations with their parents.^{130,131} Interestingly, vocal plasticity does not appear to be limited to the infant and juvenile periods of development as marmoset vocal behaviors remain highly plastic into adulthood.^{36,128,133} Adult marmosets, for example, change specific features of the calls depending on the social context,^{14,35,38,133–139} such as taking on a new “dialect” of a social group after translocation.¹³⁹ The prevalence of vocal plasticity in marmosets is notable as it contrasts with other primates,^{140,141} such as squirrel monkeys¹⁴² and macaques,^{143,144} which do not show the same effects of ontogeny. This suggests a suite of potentially complementary mechanisms which may underly vocal control in this species at different stages of life and make the marmoset a compelling candidate to study aspects of speech development in humans.¹⁴⁵

Several studies in marmosets have more precisely quantified motor control in experimental contexts.^{122–124,146} Phee calls consist of a variable number of pulses, most often two, and thus can be used to elucidate whether marmosets produce a motor plan for the full multipulsed call prior to the vocal onset or in concurrence with the generation of each pulse. Miller et al.¹²² found that acoustic features early in the call were highly correlated with the call structure—both how many phee pulses the marmoset made and the spectral features of the call. Because the features of the first pulse were predictive of the number of pulses a call would consist of, they suggested that a single motor plan for the whole call is created before or at the call onset. However, later studies using acoustic perturbation at various times after phee call onset have shown that marmosets are able to change the call length¹²³ and the number of pulses¹⁴⁷ after initiation of a call. This suggests each call consists of a number of smaller motor units, rather than a single motor plan, which they are able to flexibly control throughout the call and not just at the onset.

Zhao et al.¹²⁴ further leveraged this acoustic-perturbation paradigm to show that rapid changes to the spectra of phee calls occurred during perturbation, such as shifts in the fundamental frequency and longer call phrases, while other experiments show that marmosets shift the fundamental frequency of their phee calls in the presence of noise that interferes with the particular frequency.³⁶

Similarly, Eliades and Tsunada¹⁴⁶ used a paradigm where marmosets heard a pitch-shifted version of their own call during vocalizations. They found that the marmosets would change the pitch of their own vocalizations, suggesting that marmosets could sense the error between their motor goal the received sensory input and were able to rapidly adjust their calls. Pairs of marmosets also changed the relative timing of their vocalizations to enable vocal exchanges when tested with temporally periodic and aperiodic white noise broadcasts.¹⁴⁸ Together, these studies show that marmosets are able to rapidly adjust their calls in response to sensory feedback and environmental noise, suggesting they possess a high level of vocal control throughout adulthood.

BRAIN NETWORKS UNDERLYING VOCAL PRODUCTION

The neural basis of primate vocal production has been hotly debated for many years. One recent model¹¹⁴ states a division into two separate networks (Figure 3), distinguishing between the primary vocal motor network (PVMN), which exists across species,¹⁴⁹ and the volitional articulatory motor network (VAMN). The hypothesis suggests that the VAMN only exists in primate species.¹¹⁴ Another model¹⁵⁰ described a four-level system, where the PVMN is further split into a central pattern generator (CPG) and a slower innate driving network, while the laryngeal structure forms the fourth facet.¹⁵⁰ Specifically, this model poses that the various layers contribute to vocal production at different timescales, thus addressing the coordination of vocal production across single syllables, words, or conversations. Although both hypotheses agree about the general structure and hierarchical control involved in vocal production, Hage and Nieder¹¹⁴ argue that the VAMN is a primate-specific network, whereas Zhang and Ghazanfar¹⁵⁰ describe a top layer of the hierarchy that is shared between species that engage in coordinated vocal exchanges, including monkeys, but also meerkats. Additionally, Zhang and Ghazanfar¹⁵⁰ do not assign specific brain regions to these layers of the hierarchy, but rather a functional system that would roughly map onto Hage and Nieder’s brain regions. Despite these differences, both agree that the innate production of vocalizations relies on a CPG in the brainstem, specifically in the reticular formation,¹⁵¹ which is highly conserved across vertebrate species.^{152,153} The CPG connects directly to phonatory motoneurons and receives input from the periaqueductal gray (PAG), which is located in the midbrain.¹⁵⁴ The robustness of this system is demonstrated through studies in which the PAG is stimulated, resulting in the production of innate calls across a large number of species, including various primates.^{155–157} The PAG in turn receives input from a wide variety of areas, including the amygdala and ACC.^{114,150,158} The role of these connections is thought to be the stimulation of PAG to elicit calls in the appropriate context, such as a certain social or emotional context,^{159–161} though the details on how various inputs control PAG and vocalizations in general are still debated.

In addition to this conserved mammalian vocal system, NHPs are thought to possess an additional network of structures, mainly located

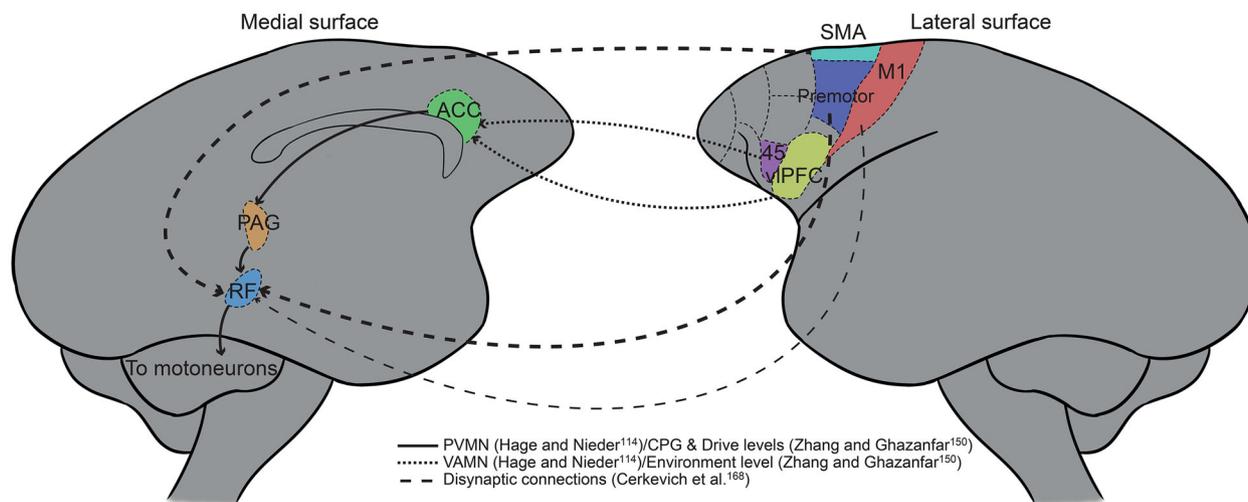


FIGURE 3 Cortical networks for vocal production. A graphical representation of the three vocal production networks discussed in this review. The PVMN¹¹⁴, or CPG and drive levels¹⁵⁰ (solid arrows), involve connections from the ACC to PAG to RF, which projects to the motoneurons. The VAMN,¹¹⁴ or environment level¹⁵⁰ (dotted arrows), shows the connection from area 45 (Broca's area) and the vIPFC on the lateral surface to the ACC on the medial surface. The connectivity described by Cerkevich et al.¹⁶⁸ (dashed arrows) involves direct connections from M1, SMA, and the ventral premotor cortex to the RF, with the two premotor connections being stronger (represented by thicker lines). Abbreviations: ACC, anterior cingulate cortex; CPG, central pattern generator; M1, primary motor cortex; PAG, periaqueductal gray; PVMN, primary vocal motor network; RF, reticular formation; SMA, supplementary motor area; VAMN, volitional articulatory motor network; vIPFC, ventrolateral prefrontal cortex.

in the frontal cortex, that control more complex vocal sequences similar to speech in humans.^{114,150} This network includes areas of the premotor and motor cortices as well as the inferior frontal gyrus. Central to this network is Broca's area, which plays a large role in the production of speech in humans,^{162,163} and the NHP homologs in the vIPFC.^{164,165} Both Hage and Nieder¹¹⁴ and Zhang and Ghazanfar¹⁵⁰ suggest these areas connect to the ACC and control vocalizations through the PAG. Although general connectivity studies seem to support such a network,¹⁶⁶ several areas in the PFC also connect directly to the PAG or reticular formation,^{167,168} and as such might be able to control vocalizations directly. In addition, connections between the ACC and vIPFC are reciprocal,¹⁶⁹ suggesting that information flows in both directions. Lastly, neither of these frameworks account for the apparent differences in vocal control between marmosets and other primates. Although other primates, such as macaques, are able to exert some level of control over their vocalizations,^{170,171} marmosets are able to flexibly suppress vocalizations¹²³ and change a vocalization's frequency³⁶ and length.¹⁴⁷ These vocalizations involve fine motor skills that have so far not been shown in other NHPs despite the similarity in neural circuits of vocal control.

A recent study sought to address this apparent disparity in vocal control between primates by performing a comparative study of the vocal motor pathways in marmosets and macaques.¹⁶⁸ Cerkevich et al.¹⁶⁸ noted that enhanced vocal control in marmosets are unlikely to be attributed to laryngeal biomechanics and other peripheral mechanisms alone,¹¹⁵ and thus they hypothesized that the brain areas driving the laryngeal muscles play a more significant role in this process. To test this conjecture, they used retrograde viral tracing from the cricothyroid muscle, the laryngeal muscle that is most specifically related to vocal motor control, to identify the cortical areas

involved in vocal motor production in both macaques and marmosets. Analyses revealed that the neurons in both premotor and primary motor areas formed disynaptic connections, via the reticular formation and nucleus ambiguus, to motoneurons. These connections are unlike those in humans, which have direct corticomotoneuronal connections. In marmosets, compared to macaques, a much larger number of these disynaptic output neurons was located in the supplementary motor area and ventral premotor cortex compared to the primary motor cortex (Figure 3), suggesting that the increased vocal plasticity and control in marmosets relative to macaques is not likely due to alterations in primary motor cortex, but from expansions in the disynaptic connections to laryngeal motor neurons in these two premotor areas. Given these neuroanatomical results, future research into the contribution of the output neurons in these areas to flexible vocal production might give an insight into the neurobiology underlying complex vocal skills. Considered together with the results discussed above and the species amenability to naturalistic, freely moving neurophysiology paradigms, these findings highlight that marmosets are a uniquely powerful primate model of vocal production and control.

Data in support of each of the aforementioned models are somewhat mixed, at least in part due to gaps in our knowledge about vocal production, control, and learning discussed above, as well as the neural mechanisms that underly each of these facets of the vocal-motor system in primates. As with the research on the neural basis of vocal signal processing discussed above, studies on vocal production have largely focused on individual brain regions that are thought to be involved. Here, we focus on the role of forebrain regions for primate vocal production (for a review on mid- and hindbrain regions, see Jürgens).¹⁷²

Frontal cortex

Several studies in marmosets have explored the possibility of a role for the frontal cortex in the production of vocalizations. Early studies using immediate early gene (IEG) expression as a proxy for neural activity showed seemingly contradictory results regarding the effect of producing vocalizations on IEG in the frontal cortex.^{102,173} Where one found increased expression in the vIPFC in animals that produced calls,¹⁷³ another study found only a significant increase in the premotor cortex.¹⁰² However, the task was slightly different between the two studies. Simões et al.¹⁷³ compared a condition where the animals listened and responded to conspecific vocalizations (antiphonal conversation) to a condition where the animals only listened but did not produce any calls themselves. Miller et al.¹⁰² included these same two conditions, in addition to a condition where an animal did not listen to any calls, but vocalized spontaneously. When the marmosets produced calls, regardless of whether they also listened to calls, the premotor cortex was activated. However, when the marmosets also listened to calls, which was the case in the study by Simões et al., the vIPFC was activated as well.

Recent neurophysiology experiments in freely moving marmosets examined the responses of single neurons in the frontal cortex during vocal production.^{101,125} Overall, these studies report evidence that neurons exhibited robust vocal-motor-related activity during call production in both dorsal and ventral premotor regions, as well as from a subset of neurons in areas 45 and 8av in the vIPFC. The authors distinguished between an antiphonal and spontaneous state during vocal production, but found no difference between these two states.¹⁰¹ Importantly, Roy et al.¹²⁵ confirmed these neurophysiology results and controlled for orofacial movement, indicating that the premotor activity is specifically related to the production of vocalizations. While each of these studies tested only a single call type, phee calls, a subsequent experiment recorded neural activity in the premotor cortex, while marmosets produced a broader corpus of vocalizations.¹⁷⁴ Results indicated that while overall neurons exhibited vocal-motor-related changes in activity just prior to or during call production, different subpopulations of premotor neurons are active depending on the call type produced. Their results show that the marmoset premotor cortex and vIPFC are involved in the spontaneous/volitional production of vocalizations during communication, though more detailed experiments testing their respective contributions to the control and production of calls are needed.

The majority of data on the neural basis of vocal production in marmosets are in the premotor cortex, despite convincing evidence from macaques showing the involvement of other regions, in particular the vIPFC and ACC.^{161,165} The vocal production networks described earlier assign the ACC the role of relay center, taking inputs from the vIPFC and premotor areas and sending signals to the PAG. This is based on results showing cue-related activity in the vIPFC and ramping activity in the ACC prior to a cued call but not for spontaneously produced calls.¹⁶¹ By contrast, neurons in the marmoset vIPFC and premotor cortex exhibited robust vocal-motor changes in activity for calls pro-

duced spontaneously and in conversations.¹⁰¹ However, each of these results are based on a limited number of neurons and animals, thus further evidence is needed to determine the role of the vIPFC and ACC in the production of vocalizations. Given the pattern of activity in the premotor cortex during vocal production, and the robust connectivity patterns, future neurophysiology experiments are critically needed in these other substrates to understand the neural basis of vocal production more fully in primates.

Auditory cortex

In addition to its role in vocal signal processing and perception, the ACx also plays an important role in vocal production through monitoring self-generated calls in a continuously updating feedback loop.^{146,175-177} Studies of auditory cortical processing in species such as rodents and songbirds have found that the ACx exhibits widespread inhibition of neural activity corresponding to background noise and self-generated locomotor activity.^{62,178,179} Indeed, Eliades and Wang¹⁷⁷ found that approximately 75% of recorded A1 neurons were significantly inhibited during the production of a marmoset's own vocalizations in a frequency-dependent manner. Eliades and Tsunada¹⁴⁶ used an online vocal pitch-shifting paradigm to further study the role of the ACx in the control of marmoset vocal behavior. They found that the population of A1 neurons, which would ordinarily be suppressed by one's own call, exhibited a significant reduction in suppression upon the manipulation of the vocal pitch away from the fundamental frequency of the initial vocalization. They also demonstrate a causal role for the ACx by directly stimulating this same population of neurons, which resulted in correlated shifts in both vocal pitch and firing rate of A1 neurons. These findings provide critical evidence that the auditory cortex is integral to vocal communication, not only for the encoding of conspecific vocalizations but for accurate and dynamic vocal control as well.

VOCAL COMMUNICATION IN THE NATURAL WORLD

Our understanding of the neural mechanisms underlying vocal communication in primates is largely based on studies in which vocal perception and production are studied independently. In fact, much of the literature on the neural representations of vocalizations have used traditional head-restrained monkey paradigms and presented isolated exemplars of vocalizations that are largely divorced from their natural contexts. However, in natural contexts, communication is an inherently interactive social behavior involving the active exchange of signals between individuals. In this sense, vocal perception and production are not separable, but reflect complementary processes that seamlessly integrate as the foundational architecture of the communication system. Evidence suggests that facets of natural communication cannot be captured using more reductionist paradigms and the consideration of each component of the system in isolation.^{103,180} As such,

it is crucial that future work takes a more holistic approach to the neurobiological study of communication and examines its processes in naturalistic contexts comprising the very challenges the system evolved to overcome.¹⁸¹

Marmosets are highly voluble and amenable to freely moving, naturalistic paradigms and neural recording systems, which makes the species particularly ideal to elucidate the neural basis of natural communication. Indeed, as highlighted throughout this manuscript, work to this end is accelerating, further illustrating the significance of this model organism in the next chapter of the field. Here, we highlight some of the recent technological advances that allow the monitoring of vocal communication in the natural world and briefly discuss two lines of inquiry that have been understudied, but are becoming increasingly feasible to study because of current methodological innovations.

Technical advances

As outlined above, most of what is known about natural marmoset vocal communication has come from studies of the species phee call during antiphonal call interactions. While clearly powerful, this type of long-distance dyadic vocal behavior reflects only a subset of the species' communicative repertoire. The study of dynamic, multicaller environments that approximate more natural social landscapes, such as the home colony, comes with technical challenges. The vocalizations of the experimental monkey, as well as the other monkeys in the colony, need to be recorded and isolated. Animals may move around and produce calls intended for different animals in various social contexts that change fluidly, making it difficult to determine where and under what context a call is made. Despite these challenges, such experiments are becoming increasingly feasible.

In recent studies, marmosets were trained to wear small microphones, which were used to record their calls within the colony.^{26,38} Likewise, several methods of automatically detecting calls from these recordings have been developed,^{26,38,182-184} which allow for rapid analysis of vast amounts of data. In addition to detecting calls, deep convolutional networks can also be used to identify which individual animal made a certain call.²⁶ Lastly, marmosets can be filmed and tracked using automated pose estimation techniques¹⁸⁵⁻¹⁸⁷ to determine the behaviors that covary with call emissions. These technological developments make it possible to quantify the detailed nuances of natural marmoset behaviors in increasingly complex natural contexts.

Despite considerable challenges, performing neuronal recordings in natural conditions is becoming increasingly feasible. Many systems for neuronal recording, such as two-photon imaging¹⁸⁸ and various methods of neurophysiology,^{103,189} that traditionally required head-fixed animals are being adapted for freely moving paradigms. Tethered miniaturized microscopes, for example, can be used to record calcium transients during natural behavior,¹⁹⁰ while wireless neural recording technologies can be used to record spiking activity with silicon probes and other electrodes.^{191,192} These methods have meant a significant improvement in the breadth of behaviors whose underlying neurobiology can be studied in naturalistic contexts,^{193,194} but their

use in marmosets comes with important considerations. Because of the marmosets relatively small body size, wireless equipment must be lightweight to not impede the animals' movements,¹⁹³ while at the same time, it must be designed to withstand the rigors of a 3D arboreal lifestyle. Recent studies using wireless recordings in marmosets have shown promising results,^{193,195} recordings from up to 96 channels simultaneously while the marmoset is freely moving in the home cage. Employing such wireless recording technologies and combining them with wearable microphones and behavior tracking opens the door to a range of new questions related to marmoset vocal communication that were not possible even just a few years ago.

Cocktail party problem

Noise is ever present in natural environments. Given the significance of conveying accurate information for natural communication, it is not surprising that the auditory system evolved mechanisms to overcome these challenges. In fact, animals are remarkably adept at recognizing conspecific vocalization and extracting meaningful information from these signals in the presence of competing background sounds, including other vocalizations,¹⁹⁶⁻¹⁹⁹ classically illustrated by the cocktail party problem (CPP) and auditory scene analysis (ASA). While studies of the CPP in NHPs are largely absent, a handful of experiments have tested the role of the ACx in ASA.^{62,79,200-205} For example, in macaques, the auditory cortex shows increased separation for tones that are mistuned, which would allow for these mistuned tones to be more easily separated from the background.²⁰⁴ In addition, the activation in macaque auditory neurons follows human behavioral results, showing increased separation of tones that were out of phase.²⁰² In humans, a series of studies recording intracortical STG activity during passive listening to multiple simultaneous voices has begun to discover differential patterns of activity in primary compared to tertiary ACx regions corresponding to the ability to attend to one sound versus another.^{86,87,206-208}

While studies in NHPs and humans employ differing behavioral paradigms and stimuli, with vocalizations and speech notably absent as stimuli in primate studies, findings in both areas support the idea that primary auditory regions respond to both the competing sounds⁸⁷ and to the attended-to target sound,⁷⁹ while downstream regions in the parabelt and rostral STG more selectively encode only the target, be it an artificial tone or a vocalization.⁶² However, how the auditory system represents the attended-to vocalization, while effectively ignoring the other sounds, remains poorly understood. To our best knowledge, the field lacks any study which has recorded neural activity in the ACx, or any other area of the brain, while monkeys listened to multiple co-occurring vocalizations. Given the converging lines of evidence pointing to the role of both the ACx and its various downstream targets in stream segregation and vocalization processing, it is crucial that these experiments be conducted. Partially in service of this goal, recent behavioral studies by Jovanovic and Miller²⁰⁹ engaged marmoset subjects in an interactive, species-specific, simulated "cocktail

party" scenario in order to better understand marmosets' behavioral strategies for successfully navigating this core challenge. Coupled with wireless electrophysiology systems, such naturalistic behavioral paradigms offer powerful opportunities to investigate neural mechanisms that underlie the CPP and ASA in the primate brain that have been difficult to study with more conventional approaches.

Multimodality

Primate communication is often not restricted solely to vocalizations, and acoustic social signals are closely integrated with visual ones, such as faces and facial expressions.^{30,210–212} Indeed, both macaques^{213,214} and marmosets^{30,215} exhibit a range of expressions, and marmosets have been shown to vary their expression depending on the affective value of a stimulus.³⁰ In addition, macaques are sensitive to the congruence of faces articulating vocalizations (i.e., McGurk Effect),²¹⁰ strongly suggesting these primates integrate audio–visual information.

Both macaque⁹¹ and marmoset⁹⁰ frontal lobes contain patches of face-selective cells, which are integrated with more extensive face processing and voice processing networks.^{85,88,216–220} Primates match the identity of an individual across their face and voice,^{104,221} though data on how these social signals are integrated for social recognition at the neural level are only beginning to emerge.²²² A recent study in marmosets suggests that the hippocampus may be integral to the cross-modal integration of faces and voices for cohesive representations of individual identity.¹⁰⁴ Tyree et al.¹⁰⁴ showed that, like humans,²²³ individual neurons in the marmoset were highly selective to specific individuals when seeing their faces or hearing their voices. In contrast to human experiments, however, a parallel population was discovered that represented the cross-modal identity of multiple individuals within individual neurons. Analyses of the population activity revealed not only that the cross-modal identity of all animals tested was encoded, but that other social categories (i.e., family vs. nonfamily) were also represented in the same population. This suggests the hippocampus may be involved in encoding critical information about social context (e.g., partner identity and social relationship) during communicative interactions. In addition to the hippocampus, the frontal cortex,²²⁴ amygdala,²²⁵ other areas of the temporal lobe,²²⁶ and auditory cortex²²⁷ all show multimodal responses to auditory and visual stimuli, suggesting that these representations recruit a broad network of substrates. Especially in marmosets, which use eye contact and facial expressions during their communication, we would expect to see an interplay of visual and auditory signals to drive communication behaviors.^{30,212} This multimodal aspect of communication has so far received little attention, with studies mainly focusing on phee calls, which are made in visual isolation. The recent interest in other types of calls, such as trills,³⁸ which are made while in visual contact with a partner, will enable future cross-modal experiments exploring how social signals across modalities are integrated and functionally complement each other during communication in natural contexts.

CONCLUSION

In this review, we discussed the current stage of research on the neural basis of vocal communication in marmosets and complemented this summary with suggestions for both exciting new avenues and topics with important remaining questions that need more work. A recurring theme throughout has been an emphasis on vocal communication in natural behavioral contexts, as it is in these settings that the neural processes that evolution selected to overcome the idiosyncratic challenges of communication will be most apparent.¹⁸¹ Answers to these challenges are becoming increasingly attainable with the recent developments in gene editing, allowing for genetically expressed calcium indicators,⁶ and high-density silicon probe recordings^{192,193} to record large populations of neurons from the brain. While these cutting-edge neural recording approaches are critical, elucidating the neural basis of natural behavior hinges on techniques being developed to be used in freely moving experiments^{192–194} combined with novel behavioral paradigms,^{209,228} which exploit the wide range of natural behaviors in marmosets.^{37,101,194,229} This is particularly important as various behaviors can cause widespread activation across the brain and strongly predict activity across various regions,^{230,231} suggesting that vocal communication should be studied in concert with other natural behaviors. After all, communication is a social behavior, not merely the process of hearing and producing signals. The continuing popularity of marmosets as a model animal across neuroscience disciplines^{1,17} illustrates their many advantages to explore facets of the natural brain and behaviors, not only because of their practical benefits for gene-editing and freely moving neural paradigms but also as a result of their dynamic social, cognitive, and communicative behaviors that are only beginning to be explored. The marmoset repertoire can serve as a powerful engine of discovery in the coming years to elucidate the complex nuances of the neural mechanisms underlying primate social brain functions.

AUTHOR CONTRIBUTIONS

D.M.G., B.J.P., J.C.G., and C.T.M. wrote the paper jointly.

ACKNOWLEDGMENTS

The authors were partially supported by NIH grant R01 DC012087 while writing this manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ORCID

Dori M. Grijseels  <https://orcid.org/0000-0002-3822-5323>

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/nyas.15057>.

REFERENCES

- Mitchell, J. F., Reynolds, J. H., & Miller, C. T. (2014). Active vision in marmosets: A model system for visual neuroscience. *Journal of Neuroscience*, 34, 1183–1194.
- Miller, C. T., Freiwald, W. A., Leopold, D. A., Mitchell, J. F., Silva, A. C., & Wang, X. (2016). Marmosets: A neuroscientific model of human social behavior. *Neuron*, 90, 219–233.
- Eliades, S. J., & Miller, C. T. (2017). Marmoset vocal communication: Behavior and neurobiology. *Developmental Neurobiology*, 77, 286–299.
- Kumita, W., Sato, K., Suzuki, Y., Kurotaki, Y., Harada, T., Zhou, Y., Kishi, N., Sato, K., Aiba, A., Sakakibara, Y., Feng, G., Okano, H., & Sasaki, E. (2019). Efficient generation of knock-in/knock-out marmoset embryo via CRISPR/Cas9 gene editing. *Scientific Reports*, 9, 1–13.
- Kishi, N., Sato, K., Sasaki, E., & Okano, H. (2014). Common marmoset as a new model animal for neuroscience research and genome editing technology. *Development Growth and Differentiation*, 56, 53–62.
- Park, J. E., Zhang, X. F., Choi, S.-H. O., Okahara, J., Sasaki, E., & Silva, A. C. (2016). Generation of transgenic marmosets expressing genetically encoded calcium indicators. *Scientific Reports*, 6, 34931.
- Park, J. E., & Sasaki, E. (2020). Assisted reproductive techniques and genetic manipulation in the common marmoset. *Ilar Journal*, 61, 286–303.
- Koski, S. E., & Burkart, J. M. (2015). Common marmosets show social plasticity and group-level similarity in personality. *Scientific Reports*, 5, 1–7.
- Finkenwirth, C., & Burkart, J. M. (2017). Long-term-stability of relationship structure in family groups of common marmosets, and its link to proactive prosociality. *Physiology & Behavior*, 173, 79–86.
- Koenig, A., & Rothe, H. (1991). Social relationships and individual contribution to cooperative behaviour in captive common marmosets (*Callithrix jacchus*). *Primates: Journal of Primatology*, 32, 183–195.
- Yamamoto, M. E. (2005). Infant care in callitrichids: Cooperation and competition. *Annual Review of Biomedical Sciences*, 7, 149–160.
- Digby, L. J., Ferrari, S. F., & Saltzman, W. (2006). The role of competition in cooperatively breeding species. In C. Campbell, A. Fuentes, K. MacKinnon, S. Bearder, & R. Stumpf (Eds.), *Primates in perspective* (pp. 85–106) Oxford University Press.
- Burkart, J. M., & Van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Animal Cognition*, 13, 1–19.
- Snowdon, C. T. (2001). Social processes in communication and cognition in callitrichid monkeys: A review. *Animal Cognition*, 4, 247–257.
- Chaplin, T. A., Yu, H.-H., Soares, J. G. M., Gattass, R., & Rosa, M. G. P. (2013). A conserved pattern of differential expansion of cortical areas in simian primates. *Journal of Neuroscience*, 33, 15120–15125.
- Hori, Y., Cléry, J. C., Selvanayagam, J., Schaeffer, D. J., Johnston, K. D., Menon, R. S., & Everling, S. (2021). Interspecies activation correlations reveal functional correspondences between marmoset and human brain areas. *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2110980118.
- Okano, H. (2021). Current status of and perspectives on the application of marmosets in neurobiology. *Annual Review of Neuroscience*, 44, 27–48.
- Kato, M., Okanoya, K., Koike, T., Sasaki, E., Okano, H., Watanabe, S., & Iriki, A. (2014). Human speech- and reading-related genes display partially overlapping expression patterns in the marmoset brain. *Brain and Language*, 133, 26–38.
- Jones, C. A., Duffy, M. K., Hoffman, S. A., Schultz-Darken, N. J., Braun, K. M., Ciucci, M. R., & Emborg, M. E. (2018). Vocalization development in common marmosets for neurodegenerative translational modeling. *Neurological Research*, 40, 303–311.
- Feng, G., Jensen, F. E., Greely, H. T., Okano, H., Treue, S., Roberts, A. C., Fox, J. G., Caddick, S., Poo, M.-M., Newsome, W. T., & Morrison, J. H. (2020). Opportunities and limitations of genetically modified nonhuman primate models for neuroscience research. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 24022–24031.
- Miller, C. T., Flusberg, S., & Hauser, M. D. (2003). Interruptibility of long call production in tamarins: Implications for vocal control. *Journal of Experimental Biology*, 206, 2629–2639.
- Newman, J. D. (1985). Squirrel monkey communication. In L. A. Rosenblum, & C. L. Coe (Eds.), *Handbook of squirrel monkey research* (pp. 99–126).
- Manser, M. B. (2013). Semantic communication in vervet monkeys and other animals. *Animal Behaviour*, 86, 491–496.
- Agamaite, J. A., Chang, C.-J., Osmanski, M. S., & Wang, X. (2015). A quantitative acoustic analysis of the vocal repertoire of the common marmoset (*Callithrix jacchus*). *Journal of the Acoustical Society of America*, 138, 2906–2928.
- Bezerra, B. M., & Souto, A. (2008). Structure and usage of the vocal repertoire of *Callithrix jacchus*. *International Journal of Primatology*, 29, 671–701.
- Oikarinen, T., Srinivasan, K., Meisner, O., Hyman, J. B., Parmar, S., Fanucci-Kiss, A., Desimone, R., Landman, R., & Feng, G. (2019). Deep convolutional network for animal sound classification and source attribution using dual audio recordings. *Journal of the Acoustical Society of America*, 145, 654–662.
- Pistorio, A. L., Vintch, B., & Wang, X. (2006). Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*). *Journal of the Acoustical Society of America*, 120, 1655–1670.
- Epple, G. (1968). Comparative studies on vocalization in marmoset monkeys (Hapalidae). *Folia Primatologica*, 8, 1–40.
- De Boer, R. A., Overduin-De Vries, A. M., Louwerse, A. L., & Sterck, E. H. M. (2013). The behavioral context of visual displays in common marmosets (*Callithrix jacchus*). *American Journal of Primatology*, 75, 1084–1095.
- Kemp, C., & Kaplan, G. (2013). Facial expressions in common marmosets (*Callithrix jacchus*) and their use by conspecifics. *Animal Cognition*, 16, 773–788.
- Osmanski, M. S., & Wang, X. (2023). Perceptual specializations for processing species-specific vocalizations in the common marmoset (*Callithrix jacchus*). *Proceedings of the National Academy of Sciences of the United States of America*, 120, e2221756120.
- Miller, C. T., & Wang, X. (2006). Sensory-motor interactions modulate a primate vocal behavior: Antiphonal calling in common marmosets. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 192, 27–38.
- Miller, C. T., Mandel, K., & Wang, X. (2010). The communicative content of the common marmoset phee call during antiphonal calling. *American Journal of Primatology*, 72, 974–980.
- Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Current Biology*, 23, 2162–2168.
- Toarmino, C. R., Wong, L., & Miller, C. T. (2017). Audience affects decision-making in a marmoset communication network. *Biology Letters*, 13, 20160934.
- Zhao, L., Rad, B. B., & Wang, X. (2019). Long-lasting vocal plasticity in adult marmoset monkeys. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 286, 20190817.
- Nummela, S. U., Jovanovic, V., De La Mothe, L., & Miller, C. T. (2017). Social context-dependent activity in marmoset frontal cortex populations during natural conversations. *Journal of Neuroscience*, 37, 7036–7047.
- Landman, R., Sharma, J., Hyman, J. B., Fanucci-Kiss, A., Meisner, O., Parmar, S., Feng, G., & Desimone, R. (2020). Close-range vocal interaction in the common marmoset (*Callithrix jacchus*). *PLoS ONE*, 15, e0227392.

39. Miller, C. T., & Wren Thomas, A. (2012). Individual recognition during bouts of antiphonal calling in common marmosets. *Journal of Comparative Physiology A*, 198, 337–346.
40. Caselli, C. B., Ayres, P. H. B., Castro, S. C. N., Souto, A., Schiel, N., & Miller, C. T. (2018). The role of extragroup encounters in a Neotropical, cooperative breeding primate, the common marmoset: A field playback experiment. *Animal Behaviour*, 136, 137–146.
41. Miller, C. T., Beck, K., Meade, B., & Wang, X. (2009). Antiphonal call timing in marmosets is behaviorally significant: Interactive playback experiments. *Journal of Comparative Physiology A*, 195, 783–789.
42. Aitkin, L., & Park, V. (1993). Audition and the auditory pathway of a vocal New World primate, the common marmoset. *Progress in Neurobiology*, 41, 345–367.
43. De La Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2006). Thalamic connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *Journal of Comparative Neurology*, 496, 72–96.
44. De La Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2006). Cortical connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *Journal of Comparative Neurology*, 496, 27–71.
45. De La Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2012). Thalamic connections of auditory cortex in marmoset monkeys: Lateral belt and parabelt regions. *Anatomical Record-Advances in Integrative Anatomy and Evolutionary Biology*, 295, 822–836.
46. Tammer, R., Ehrenreich, L., & Jürgens, U. (2004). Telemetrically recorded neuronal activity in the inferior colliculus and bordering tegmentum during vocal communication in squirrel monkeys (*Saimiri sciureus*). *Behavioural Brain Research*, 151, 331–336.
47. Hage, S. R., Jürgens, U., & Ehret, G. (2006). Audio-vocal interaction in the pontine brainstem during self-initiated vocalization in the squirrel monkey. *European Journal of Neuroscience*, 23, 3297–3308.
48. Wang, X., Zhang, Y., Bai, S., Qi, R., Sun, H., Li, R., Zhu, L., Cao, X., Jia, G., Li, X., & Gao, L. (2022). Corticofugal modulation of temporal and rate representations in the inferior colliculus of the awake marmoset. *Cerebral Cortex*, 32, 4080–4097.
49. Nelson, P. C., Smith, Z. M., & Young, E. D. (2009). Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. *Journal of Neuroscience*, 29, 2553–2562.
50. Pomberger, T., & Hage, S. R. (2019). Semi-chronic laminar recordings in the brainstem of behaving marmoset monkeys. *Journal of Neuroscience Methods*, 311, 186–192.
51. Scott, B. H., Leccese, P. A., Saleem, K. S., Kikuchi, Y., Mullarkey, M. P., Fukushima, M., Mishkin, M., & Saunders, R. C. (2015). Intrinsic connections of the core auditory cortical regions and rostral supratemporal plane in the macaque monkey. *Cerebral Cortex*, 27(1), bhv277.
52. Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 11793–11799.
53. Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*, 2, 1131–1136.
54. Barbas, H., Bunce, J. G., & Medalla, M. (2013). Prefrontal pathways that control attention. In D. T. Stuss, & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 31–48).
55. Thorson, I. L., Liénard, J., & David, S. V. (2015). The essential complexity of auditory receptive fields. *PLoS Computational Biology*, 11, 1–33.
56. Machens, C. K., Wehr, M. S., & Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. *Journal of Neuroscience*, 24, 1089–1100.
57. Atencio, C. A., Sharpee, T. O., & Schreiner, C. E. (2009). Hierarchical computation in the canonical auditory cortical circuit. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 21894–21899.
58. Sharpee, T. O., Atencio, C. A., & Schreiner, C. E. (2011). Hierarchical representations in the auditory cortex. *Current Opinion in Neurobiology*, 21, 761–767.
59. Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature*, 424, 669–674.
60. Rabinowitz, N. C., Willmore, B. D. B., King, A. J., & Schnupp, J. W. H. (2013). Constructing noise-invariant representations of sound in the auditory pathway. *PLoS Biology*, 11, e1001710.
61. Knudsen, D. P., & Gentner, T. Q. (2013). Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. *Journal of Neurophysiology*, 109, 1690–1703.
62. Schneider, D. M., & Woolley, S. M. N. (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron*, 79, 141–152.
63. Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science*, 292, 290–293.
64. Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12, 718–724.
65. Winer, J. A., & Schreiner, C. E. (2011). *The auditory cortex*.
66. Kell, A. J. E., & Mcdermott, J. H. (2019). Invariance to background noise as a signature of non-primary auditory cortex. *Nature Communications*, 10, 3958.
67. Sen, K., Theunissen, F. E., & Doupe, A. J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. *Journal of Neurophysiology*, 86, 1445–1458.
68. Carruthers, I. M., Laplagne, D. A., Jaegle, A., Briguglio, J. J., Mwilambwe-Tshilobo, L., Natan, R. G., & Geffen, M. N. (2015). Emergence of invariant representation of vocalizations in the auditory cortex. *Journal of Neurophysiology*, 114, 2726–2740.
69. Rocchi, F., Oya, H., Balezeau, F., Billig, A. J., Kocsis, Z., Jenison, R. L., Nourski, K. V., Kovach, C. K., Steinschneider, M., Kikuchi, Y., Rhone, A. E., Dlouhy, B. J., Kawasaki, H., Adolphs, R., Greenlee, J. D. W., Griffiths, T. D., Howard, M. A., & Petkov, C. I. (2021). Common fronto-temporal effective connectivity in humans and monkeys. *Neuron*, 109, 852–868.e8.
70. Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 11800–11806.
71. Rosa, M. G. P., Palmer, S. M., Gamberini, M., Burman, K. J., Yu, H.-H., Reser, D. H., Bourne, J. A., Tweedale, R., & Galletti, C. (2009). Connections of the dorsomedial visual area: Pathways for early integration of dorsal and ventral streams in extrastriate cortex. *Journal of Neuroscience*, 29, 4548–4563.
72. Solomon, S. G., & Rosa, M. G. P. (2014). A simpler primate brain: The visual system of the marmoset monkey. *Frontiers in Neural Circuits*, 8, 96.
73. Chen, C., Remington, E. D., & Wang, X. (2023). Sound localization acuity of the common marmoset (*Callithrix jacchus*). *Hearing Research*, 430, 108722.
74. Zhou, Y., & Wang, X. (2012). Level dependence of spatial processing in the primate auditory cortex. *Journal of Neurophysiology*, 108, 810–826.
75. Remington, E. D., & Wang, X. (2019). Neural representations of the full spatial field in auditory cortex of awake marmoset (*Callithrix jacchus*). *Cerebral Cortex*, 29, 1199–1216.

76. Eliades, S. J., & Tsunada, J. (2019). Marmosets in auditory research. In R. P. Marini, L. M. Wachtman, S. D. Tardif, K. Mansfield, & J. G. Fox (Eds.), *The common marmoset in captivity and biomedical research* (pp. 451–475). Elsevier.
77. Hackett, T. A., De La Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., & Schroeder, C. E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: Refining the hierarchical model. *Frontiers in Neuroscience*, 8, 1–21.
78. Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience*, 13, 1428–1432.
79. Schneider, F., Dheerendra, P., Balezeau, F., Ortiz-Rios, M., Kikuchi, Y., Petkov, C. I., Thiele, A., & Griffiths, T. D. (2018). Auditory figure-ground analysis in rostral belt and parabelt of the macaque monkey. *Scientific Reports*, 8, 1–8.
80. Kadia, S. C., & Wang, X. (2003). Spectral integration in A1 of awake primates: Neurons with single- and multiplexed tuning characteristics. *Journal of Neurophysiology*, 89, 1603–1622.
81. Wang, X., Merzenich, M. M., Beitel, R., & Schreiner, C. E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. *Journal of Neurophysiology*, 74, 2685–2706.
82. Zeng, H.-H., Huang, J.-F., Li, J.-R., Shen, Z., Gong, N., Wen, Y.-Q., Wang, L., & Poo, M.-M. (2021). Distinct neuron populations for simple and compound calls in the primary auditory cortex of awake marmosets. *National Science Review*, 8, nwab126.
83. Rauschecker, J. P. (2019). Dual stream models of auditory vocal communication. In S. Frühholz, & P. Belin (Eds.), *Oxford handbook of voice perception* (pp. 413–428). Oxford University Press.
84. Sadagopan, S., Temiz-Karayol, N. Z., & Voss, H. U. (2015). High-field functional magnetic resonance imaging of vocalization processing in marmosets. *Scientific Reports*, 5, 1–15.
85. Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Current Biology*, 21, 1408–1415.
86. Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343, 1006–1010.
87. O'sullivan, J., Herrero, J., Smith, E., Schevon, C., Mckhann, G. M., Sheth, S. A., Mehta, A. D., & Mesgarani, N. (2019). Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron*, 104, 1195–1209.e3.
88. Jafari, A., Dureux, A., Zanini, A., Menon, R. S., Gilbert, K. M., & Everling, S. (2023). A vocalization-processing network in marmosets. *Cell Reports*, 42, 112526.
89. Pernet, C. R., Mcaleer, P., Latinus, M., Gorgolewski, K. J., Charest, I., Bestelmeyer, P. E. G., Watson, R. H., Fleming, D., Crabbe, F., Valdes-Sosa, M., & Belin, P. (2015). The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *Neuroimage*, 119, 164–174.
90. Schaeffer, D. J., Selvanayagam, J., Johnston, K. D., Menon, R. S., Freiwald, W. A., & Everling, S. (2020). Face selective patches in marmoset frontal cortex. *Nature Communications*, 11, 4856.
91. Tsao, D. Y., Schweers, N., Moeller, S., & Freiwald, W. A. (2008). Patches of face-selective cortex in the macaque frontal lobe. *Nature Neuroscience*, 11, 877–879.
92. Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 19514–19519.
93. Munoz-Lopez, M. M., Mohedano-Moriano, A., & Insausti, R. (2010). Anatomical pathways for auditory memory in primates. *Frontiers in Neuroanatomy*, 4, 129.
94. Newman, J. D., & Lindsley, D. F. (1976). Single unit analysis of auditory processing in squirrel monkey frontal cortex. *Experimental Brain Research*, 25, 169–181.
95. Romanski, L. M., Averbeck, B. B., & Diltz, M. (2004). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *Journal of Neurophysiology*, 93, 734–747.
96. Azuma, M., & Suzuki, H. (1984). Properties and distribution of auditory neurons in the dorsolateral prefrontal cortex of the alert monkey. *Brain Research*, 298, 343–346.
97. Vaadia, E., Benson, D. A., Hienz, R. D., & Goldstein, M. H. (1986). Unit study of monkey frontal cortex: Active localization of auditory and of visual stimuli. *Journal of Neurophysiology*, 56, 934–952.
98. Barbas, H., & Mesulam, M.-M. (1985). Cortical afferent input to the principals region of the rhesus monkey. *Neuroscience*, 15, 619–637.
99. Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydin, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., ... & Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551, 232–236.
100. Burman, K. J., Reser, D. H., Yu, H.-H., & Rosa, M. G. P. (2011). Cortical input to the frontal pole of the marmoset monkey. *Cerebral Cortex*, 21, 1712–1737.
101. Miller, C. T., Thomas, A. W., Nummela, S. U., & De La Mothe, L. A. (2015). Responses of primate frontal cortex neurons during natural vocal communication. *Journal of Neurophysiology*, 114, 1158–1171.
102. Miller, C. T., Dimauro, A., Pistorio, A., Hendry, S., & Wang, X. (2010). Vocalization induced cFos expression in marmoset cortex. *Frontiers in Integrative Neuroscience*, 4, 128.
103. Jovanovic, V., Fishbein, A. R., De La Mothe, L., Lee, K.-F., & Miller, C. T. (2022). Behavioral context affects social signal representations within single primate prefrontal cortex neurons. *Neuron*, 110, 1318–1326.e4.
104. Tyree, T., Metke, M., & Miller, C. (2022). Cross-modal representation of identity in primate hippocampus. doi: <https://doi.org/10.1101/2022.09.12.507611>
105. Pannese, A., Grandjean, D., & Frühholz, S. (2016). Amygdala and auditory cortex exhibit distinct sensitivity to relevant acoustic features of auditory emotions. *Cortex*, 85, 116–125.
106. Kato, M., Yokoyama, C., Kawasaki, A., Takeda, C., Koike, T., Onoe, H., & Iriki, A. (2018). Individual identity and affective valence in marmoset calls: In vivo brain imaging with vocal sound playback. *Animal Cognition*, 21, 331–343.
107. West, M. J., Stroud, A. N., & King, A. P. (1983). Mimicry of the human voice by European starlings: The role of social interaction. *Wilson Bulletin*, 95(4), 635–640.
108. Taylor, D., Clay, Z., Dahl, C. D., Zuberbühler, K., Davila-Ross, M., & Dezechache, G. (2022). Vocal functional flexibility: What it is and why it matters. *Animal Behaviour*, 186, 93–100.
109. Leonardo, A., & Konishi, M. (1999). Decrystallization of adult bird-song by perturbation of auditory feedback. *Nature*, 399, 466–470.
110. Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, 289, 1355–1357.
111. Hammerschmidt, K., & Fischer, J. (2013). Constraints in primate vocal production. In D. K. Oller, & U. Griebel (Eds.), *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication*. The MIT Press.
112. Fairbanks, G. (1955). Selective vocal effects of delayed auditory feedback. *Journal of Speech and Hearing Disorders*, 20, 333–346.
113. Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, 22, 824–826.
114. Hage, S. R., & Nieder, A. (2016). Dual neural network model for the evolution of speech and language. *Trends in Neuroscience*, 39, 813–829.
115. Fitch, W. T., De Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2, e1600723.
116. Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26, 235–258.

117. Kirzinger, A., & Jürgens, U. (1982). Cortical lesion effects and vocalization in the squirrel monkey. *Brain Research*, 233, 299–315.
118. Jürgens, U. (1998). Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Die Naturwissenschaften*, 85, 376–388.
119. Mitani, J., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour*, 135, 1041–1064.
120. Brumm, H., Voss, K., Köllmer, I., & Todt, D. (2004). Acoustic communication in noise: Regulation of call characteristics in a New World monkey. *Journal of Experimental Biology*, 207, 443–448.
121. Egnor, S. E. R., Iguina, C. G., & Hauser, M. D. (2006). Perturbation of auditory feedback causes systematic perturbation in vocal structure in adult cotton-top tamarins. *Journal of Experimental Biology*, 209, 3652–3663.
122. Miller, C. T., Eliades, S. J., & Wang, X. (2009). Motor planning for vocal production in common marmosets. *Animal Behaviour*, 78, 1195–1203.
123. Pomberger, T., Risueno-Segovia, C., Löschner, J., & Hage, S. R. (2018). Precise motor control enables rapid flexibility in vocal behavior of marmoset monkeys. *Current Biology*, 28, 788–794.e3.
124. Zhao, L., Roy, S., & Wang, X. (2019). Rapid modulations of the vocal structure in marmoset monkeys. *Hearing Research*, 384, 107811.
125. Roy, S., Zhao, L., & Wang, X. (2016). Distinct neural activities in pre-motor cortex during natural vocal behaviors in a New World primate, the common marmoset (*Callithrix jacchus*). *Journal of Neuroscience*, 36, 12168–12179.
126. Choi, J. Y., Takahashi, D. Y., & Ghazanfar, A. A. (2015). Cooperative vocal control in marmoset monkeys via vocal feedback. *Journal of Neurophysiology*, 114, 274–283.
127. Pomberger, T., Risueno-Segovia, C., Gultekin, Y. B., Dohmen, D., & Hage, S. R. (2019). Cognitive control of complex motor behavior in marmoset monkeys. *Nature Communications*, 10, 3796.
128. Gultekin, Y. B., Hildebrand, D. G. C., Hammerschmidt, K., & Hage, S. R. (2021). High plasticity in marmoset monkey vocal development from infancy to adulthood. *Science Advances*, 7, eabf2938.
129. Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, 349, 734–738.
130. Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371, 20150370.
131. Chow, C. P., Mitchell, J. F., & Miller, C. T. (2015). Vocal turn-taking in a non-human primate is learned during ontogeny. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 282, 20150069.
132. Gultekin, Y. B., & Hage, S. R. (2018). Limiting parental interaction during vocal development affects acoustic call structure in marmoset monkeys. *Science Advances*, 4, eaar4012.
133. Rukstalis, M., Fite, J. E., & French, J. A. (2003). Social change affects vocal structure in a callitrichid primate (*Callithrix kuhlii*). *Ethology*, 109, 327–340.
134. Norcross, J. L., & Newman, J. D. (1997). Social context affects phee call production by nonreproductive common marmosets (*Callithrix jacchus*). *American Journal of Primatology*, 43, 135–146.
135. Norcross, J. L., & Newman, J. D. (1993). Context and gender-specific differences in the acoustic structure of common marmoset (*Callithrix jacchus*) phee calls. *American Journal of Primatology*, 30, 37–54.
136. Yamaguchi, C., Izumi, A., & Nakamura, K. (2010). Time course of vocal modulation during isolation in common marmosets (*Callithrix jacchus*). *American Journal of Primatology*, 72, 681–688.
137. Norcross, J. L., Newman, J. D., & Cofrancesco, L. M. (1999). Context and sex differences exist in the acoustic structure of phee calls by newly-paired common marmosets (*Callithrix jacchus*). *American Journal of Primatology*, 49, 165–181.
138. Snowdon, C. (2018). Cognitive components of vocal communication: A case study. *Animals*, 8, 126.
139. Zürcher, Y., Willems, E. P., & Burkart, J. M. (2019). Are dialects socially learned in marmoset monkeys? Evidence from translocation experiments. *PLoS ONE*, 14, e0222486.
140. Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neuroscience*, 27, 649–654.
141. Fischer, J., & Hammerschmidt, K. (2020). Towards a new taxonomy of primate vocal production learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375, 20190045.
142. Hammerschmidt, K., Freudenstein, T., & Jürgens, U. (2001). Vocal development in squirrel monkeys. *Behaviour*, 138, 1179–1204.
143. Hammerschmidt, K., Newman, J. D., Champoux, M., & Suomi, S. J. (2000). Changes in rhesus macaque 'coo' vocalizations during early development. *Ethology*, 106, 873–886.
144. Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, 26, 389–406.
145. Burkart, J. M., Adriaense, J. E. C., Brügger, R. K., Miss, F. M., Wierucka, K., & Van Schaik, C. P. (2022). A convergent interaction engine: Vocal communication among marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377, 20210098.
146. Eliades, S. J., & Tsunada, J. (2018). Auditory cortical activity drives feedback-dependent vocal control in marmosets. *Nature Communications*, 9, 2540.
147. Löschner, J., Pomberger, T., & Hage, S. R. (2023). Marmoset monkeys use different avoidance strategies to cope with ambient noise during vocal behavior. *Iscience*, 26, 106219.
148. Roy, S., Miller, C. T., Gottsch, D., & Wang, X. (2011). Vocal control by the common marmoset in the presence of interfering noise. *Journal of Experimental Biology*, 214, 3619–3629.
149. Kelley, D. B. (2022). Convergent and divergent neural circuit architectures that support acoustic communication. *Frontiers in Neural Circuits*, 16, 112.
150. Zhang, Y. S., & Ghazanfar, A. A. (2020). A hierarchy of autonomous systems for vocal production. *Trends in Neuroscience*, 43, 115–126.
151. Hage, S. R., & Jürgens, U. (2006). Localization of a vocal pattern generator in the pontine brainstem of the squirrel monkey. *European Journal of Neuroscience*, 23, 840–844.
152. Rhodes, H. J., Yu, H. J., & Yamaguchi, A. (2007). Xenopus vocalizations are controlled by a sexually differentiated hindbrain central pattern generator. *Journal of Neuroscience*, 27, 1485–1497.
153. Hage, S. R. (2020). The role of auditory feedback on vocal pattern generation in marmoset monkeys. *Current Opinion in Neurobiology*, 60, 92–98.
154. Jürgens, U. (1994). The role of the periaqueductal grey in vocal behaviour. *Behavioural Brain Research*, 62, 107–117.
155. Brown, T. G. (1915). Note on the physiology of the basal ganglia and mid-brain of the anthropoid ape, especially in reference to the act of laughter. *Journal of Physiology*, 49, 195.
156. Jürgens, U., & Ploog, D. (1970). Cerebral representation of vocalization in the squirrel monkey. *Experimental Brain Research*, 10, 532–554.
157. Magoun, H. W., Atlas, D., Ingersoll, E. H., & Ranson, S. W. (1937). Associated facial, vocal and respiratory components of emotional expression: An experimental study. *Journal of Neurology and Psychopathology*, 17, 241–255.
158. Dujardin, E., & Jürgens, U. (2006). Call type-specific differences in vocalization-related afferents to the periaqueductal gray of squirrel monkeys (*Saimiri sciureus*). *Behavioural Brain Research*, 168, 23–36.
159. Matsumoto, Y. K., Okanoya, K., & Seki, Y. (2012). Effects of amygdala lesions on male mouse ultrasonic vocalizations and copulatory behaviour. *Neuroreport*, 23, 676–680.
160. Hall, I. C., Ballagh, I. H., & Kelley, D. B. (2013). The Xenopus amygdala mediates socially appropriate vocal communication signals. *Journal of Neuroscience*, 33, 14534–14548.

161. Gavrillov, N., Hage, S. R., & Nieder, A. (2017). Functional specialization of the primate frontal lobe during cognitive control of vocalizations. *Cell Reports*, *21*, 2393–2406.
162. Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: From action to language. *Physiology*, *20*, 60–69.
163. Rilling, J. K. (2014). Comparative primate neurobiology and the evolution of brain language systems. *Current Opinion in Neurobiology*, *28*, 10–14.
164. Gavrillov, N., & Nieder, A. (2021). Distinct neural networks for the volitional control of vocal and manual actions in the monkey homologue of Broca's area. *eLife*, *10*, e62797.
165. Hage, S. R., & Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications*, *4*, 2409.
166. Yeterian, E. H., Pandya, D. N., Tomaiuolo, F., & Petrides, M. (2012). The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex*, *48*, 58–81.
167. An, X., Bandler, R., Ongür, D., & Price, J. L. (1998). Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in macaque monkeys. *Journal of Comparative Neurology*, *401*, 455–479.
168. Cerkevich, C. M., Rathelot, J.-A., & Strick, P. L. (2022). Cortical basis for skilled vocalization. *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2122345119.
169. Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*, 417–424.
170. Hage, S. R., Gavrillov, N., & Nieder, A. (2013). Cognitive control of distinct vocalizations in rhesus monkeys. *Journal of Cognitive Neuroscience*, *25*, 1692–1701.
171. Aitken, P. G. (1981). Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain and Language*, *13*, 171–184.
172. Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice*, *23*, 1–10.
173. Simões, C. S., Vianney, P. V. R., De Moura, M. M., Freire, M. A. M., Mello, L. E., Sameshima, K., Araújo, J. F., Nicolelis, M. A. L., Mello, C. V., & Ribeiro, S. (2010). Activation of frontal neocortical areas by vocal production in marmosets. *Frontiers in Integrative Neuroscience*, *4*, 123.
174. Zhao, L., & Wang, X. (2021). Frontal cortex activity underlying the production of diverse vocal signals during social communication in marmoset monkeys. doi: <https://doi.org/10.1101/2021.08.30.458300>
175. Eliades, S. J., & Wang, X. (2019). Corollary discharge mechanisms during vocal production in marmoset monkeys. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*, 805–812.
176. Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*, 1102–1106.
177. Eliades, S. J., & Wang, X. (2005). Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex*, *15*, 1510–1523.
178. Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, *513*, 189–194.
179. Schneider, D. M., Sundararajan, J., & Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. *Nature*, *561*, 391–395.
180. McMahon, D. B. T., Jones, A. P., Bondar, I. V., & Leopold, D. A. (2014). Face-selective neurons maintain consistent visual responses across months. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8251–8256.
181. Miller, C. T., Gire, D., Hoke, K., Huk, A. C., Kelley, D., Leopold, D. A., Smear, M. C., Theunissen, F., Yartsev, M., & Niell, C. M. (2022). Natural behavior is the language of the brain. *Current Biology*, *32*, R482–R493.
182. Wisler, A., Brattain, L. J., Landman, R., & Quatieri, T. (2016). A framework for automated marmoset vocalization detection and classification. In *INTERSPEECH* pp. 2592–2596.
183. Turesson, H. K., Ribeiro, S., Pereira, D. R., Papa, J. P., & De Albuquerque, V. H. C. (2016). Machine learning algorithms for automatic classification of marmoset vocalizations. *PLoS ONE*, *11*, e0163041.
184. Liu, S. T., Montes-Lourido, P., Wang, X., & Sadagopan, S. (2019). Optimal features for auditory categorization. *Nature Communications*, *10*, 1302.
185. Pereira, T. D., Tabris, N., Matsliah, A., Turner, D. M., Li, J., Ravindranath, S., Papadoyannis, E. S., Normand, E., Deutsch, D. S., Wang, Z. Y., McKenzie-Smith, G. C., Mitelut, C. C., Castro, M. D., D'Uva, J., Kislin, M., Sanes, D. H., Kocher, S. D., Wang, S. S.-H., Falkner, A. L., ... Murthy, M. (2022). SLEAP: A deep learning system for multi-animal pose tracking. *Nature Methods*, *19*, 486–495.
186. Lauer, J., Zhou, M. U., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M. M., Di Santo, V., Soberanes, D., Feng, G., Murthy, V. N., Lauder, G., Dulac, C., Mathis, M. W., & Mathis, A. (2022). Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nature Methods*, *19*, 496–504.
187. Ngo, V., Gorman, J. C., De La Fuente, M. F., Souto, A., Schiel, N., & Miller, C. T. (2022). Active vision during prey capture in wild marmoset monkeys. *Current Biology*, *32*, 3423–3428.e3.
188. Ebina, T., Masamizu, Y., Tanaka, Y. R., Watakabe, A., Hirakawa, R., Hirayama, Y., Hira, R., Terada, S.-I., Koketsu, D., Hikosaka, K., Mizukami, H., Nambu, A., Sasaki, E., Yamamori, T., & Matsuzaki, M. (2018). Two-photon imaging of neuronal activity in motor cortex of marmosets during upper-limb movement tasks. *Nature Communications*, *9*, 1879.
189. Hung, C.-C., Yen, C. C., Ciuchta, J. L., Papoti, D., Bock, N. A., Leopold, D. A., & Silva, A. C. (2015). Functional MRI of visual responses in the awake, behaving marmoset. *Neuroimage*, *120*, 1–11.
190. Kondo, T., Saito, R., Otaka, M., Yoshino-Saito, K., Yamanaka, A., Yamamori, T., Watakabe, A., Mizukami, H., Schnitzer, M. J., Tanaka, K. F., Ushiba, J., & Okano, H. (2018). Calcium transient dynamics of neural ensembles in the primary motor cortex of naturally behaving monkeys. *Cell Reports*, *24*, 2191–2195.e4.
191. Eliades, S. J., & Wang, X. (2008). Chronic multi-electrode neural recording in free-roaming monkeys. *Journal of Neuroscience Methods*, *172*, 201–214.
192. Jendritza, P., Klein, F. J., & Fries, P. (2023). Multi-area recordings and optogenetics in the awake, behaving marmoset. *Nature Communications*, *14*, 577.
193. Walker, J. D., Pirschel, F., Sundiang, M., Niekrasz, M., Maclean, J. N., & Hatsopoulos, N. G. (2021). Chronic wireless neural population recordings with common marmosets. *Cell Reports*, *36*, 109379.
194. Roy, S., & Wang, X. (2012). Wireless multi-channel single unit recording in freely moving and vocalizing primates. *Journal of Neuroscience Methods*, *203*, 28–40.
195. Wong, R. K., Selvanayagam, J., Johnston, K. D., & Everling, S. (2023). Delay-related activity in marmoset prefrontal cortex. *Cerebral Cortex*, *33*, 3523–3537.
196. Bee, M. A., & Michey, C. (2008). The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, *122*, 235.
197. Shinn-Cunningham, B., Best, V., & Lee, A. K. C. (2017). *The auditory system at the cocktail party*. Springer International Publishing.
198. Bee, M. A. (2015). Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. *International Journal of Psychophysiology*, *95*, 216–237.
199. Itatani, N., & Klump, G. M. (2017). Animal models for auditory streaming. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, 20160112.

200. Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 11854–11859.
201. Christison-Lagay, K. L., Gifford, A. M., & Cohen, Y. E. (2015). Neural correlates of auditory scene analysis and perception. *International Journal of Psychophysiology*, 95, 238–245.
202. Knyazeva, S., Selezneva, E., Gorkin, A., Aggelopoulos, N. C., & Brosch, M. (2018). Neuronal correlates of auditory streaming in monkey auditory cortex for tone sequences without spectral differences. *Frontiers in Integrative Neuroscience*, 12, 1–14.
203. Itatani, N., & Klump, G. M. (2011). Neural correlates of auditory streaming of harmonic complex sounds with different phase relations in the songbird forebrain. *Journal of Neurophysiology*, 105, 188–199.
204. Fishman, Y. I., & Steinschneider, M. (2010). Neural correlates of auditory scene analysis based on inharmonicity in monkey primary auditory cortex. *Journal of Neuroscience*, 30, 12480–12494.
205. Fishman, Y. I., Arezzo, J. C., & Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America*, 116, 1656–1670.
206. Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485, 233–236.
207. Patel, P., Long, L. K., Herrero, J. L., Mehta, A. D., & Mesgarani, N. (2018). Joint representation of spatial and phonetic features in the human core auditory cortex. *Cell Reports*, 24, 2051–2062.e2.e2.
208. Mesgarani, N., David, S. V., Fritz, J. B., & Shamma, S. A. (2014). Mechanisms of noise robust representation of speech in primary auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 6792–6797.
209. Jovanovic, V., & Miller, C. T. (2021). Mechanisms for communicating in a marmoset 'cocktail party.' doi: <https://doi.org/10.1101/2020.12.08.416693>
210. Ghazanfar, A. A., & Logothetis, N. K. (2003). Facial expressions linked to monkey calls. *Nature*, 423, 937–938.
211. Adachi, I., & Hampton, R. R. (2011). Rhesus monkeys see who they hear: Spontaneous cross-modal memory for familiar conspecifics. *PLoS ONE*, 6, e23345.
212. Kotani, M., Shimono, K., Yoneyama, T., Nakako, T., Matsumoto, K., Ogi, Y., Konoike, N., Nakamura, K., & Ikeda, K. (2017). An eye tracking system for monitoring face scanning patterns reveals the enhancing effect of oxytocin on eye contact in common marmosets. *Psychoneuroendocrinology*, 83, 42–48.
213. Partan, S. (2002). Single and multichannel signal composition: Facial expressions and vocalizations of rhesus macaques (*Macaca mulatta*). *Behaviour*, 139, 993–1027.
214. Zhang, H., Ding, X., Liu, N., Nolan, R., Ungerleider, L. G., & Japee, S. (2023). Equivalent processing of facial expression and identity by macaque visual system and task-optimized neural network. *NeuroImage*, 273, 120067.
215. Stevenson, M. F., & Poole, T. B. (1976). An ethogram of the common marmoset (*Calithrix jacchus jacchus*): General behavioural repertoire. *Animal Behaviour*, 24, 428–451.
216. Weiner, K. S., & Grill-Spector, K. (2015). The evolution of face processing networks. *Trends in Cognitive Sciences*, 19, 240–241.
217. Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science*, 320, 1355–1359.
218. Dureux, A., Zanini, A., & Everling, S. (2023). Face-selective patches in marmosets are involved in dynamic and static facial expression processing. *Journal of Neuroscience*, 43, 3477–3494.
219. Belin, P., Bodin, C., & Aglieri, V. (2018). A "voice patch" system in the primate brain for processing vocal information? *Hearing Research*, 366, 65–74.
220. Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience*, 11, 367–374.
221. Sliwa, J., Duhamel, J.-R., Pascalis, O., & Wirth, S. (2011). Spontaneous voice–face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 1735–1740.
222. Sliwa, J., Planté, A., Duhamel, J.-R., & Wirth, S. (2016). Independent neuronal representation of facial and vocal identity in the monkey hippocampus and inferotemporal cortex. *Cerebral Cortex*, 26, 950–966.
223. Trinkler, I., King, J. A., Doeller, C. F., Rugg, M. D., & Burgess, N. (2009). Neural bases of autobiographical support for episodic recollection of faces. *Hippocampus*, 19, 718–730.
224. Khandhadia, A. P., Murphy, A. P., Romanski, L. M., Bizley, J. K., & Leopold, D. A. (2021). Audiovisual integration in macaque face patch neurons. *Current Biology*, 31, 1826–1835.e3.
225. Morrow, J., Mosher, C., & Gothard, K. (2019). Multisensory neurons in the primate amygdala. *Journal of Neuroscience*, 39, 3663–3675.
226. Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2014). Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. *Journal of Neuroscience*, 34, 2524–2537.
227. Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience*, 25, 5004–5012.
228. Walker, J., Maclean, J., & Hatsopoulos, N. G. (2017). The marmoset as a model system for studying voluntary motor control. *Developmental Neurobiology*, 77, 273–285.
229. Courellis, H. S., Nummela, S. U., Metke, M., Diehl, G. W., Bussell, R., Cauwenberghs, G., & Miller, C. T. (2019). Spatial encoding in primate hippocampus during free navigation. *PLoS Biology*, 17, e3000546.
230. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364, 255.
231. Zagha, E., Erlich, J. C., Lee, S., Lur, G., O'connor, D. H., Steinmetz, N. A., Stringer, C., & Yang, H. (2022). The importance of accounting for movement when relating neuronal activity to sensory and cognitive processes. *Journal of Neuroscience*, 42, 1375–1382.
232. Paxinos, G., Watson, C., Rosa, M., & Petrides, M. (2012). *The marmoset brain in stereotaxic coordinates*. Elsevier Academic Press.
233. Senoo, A., Tokuno, H., & Watson, C. (2015). Mini-atlas of the marmoset brain. *Neuroscience Research*, 93, 128–135.

How to cite this article: Grijseels, D. M., Prendergast, B. J., Gorman, J. C., & Miller, C. T. (2023). The neurobiology of vocal communication in marmosets. *Ann NY Acad Sci*, 1–16. <https://doi.org/10.1111/nyas.15057>