

Corollary Discharge Mechanisms During Vocal Production in Marmoset Monkeys

Steven J. Eliades and Xiaoqin Wang

ABSTRACT

Interactions between motor systems and sensory processing are ubiquitous throughout the animal kingdom and play an important role in many sensorimotor behaviors, including both human speech and animal vocalization. During vocal production, the auditory system plays important roles in both encoding feedback of produced sounds, allowing one to self-monitor for vocal errors, and simultaneously maintaining sensitivity to the outside acoustic environment. Supporting these roles is an efferent motor-to-sensory signal known as a corollary discharge. This review summarizes recent work on the role of such signaling during vocalization in the marmoset monkey, a nonhuman primate model of social vocal communication.

Keywords: Auditory cortex, Auditory-vocal, Corollary discharge, Marmoset, Sensorimotor, Vocalization

<https://doi.org/10.1016/j.bpsc.2019.06.008>

Vocal communication is a fundamental behavior that plays an important role in the lives of humans and is shared by many nonhuman animal species. This interactive process requires participants to act in a coordinated fashion both as the sender and as the receiver of information. One important aspect of this process, however, is that vocalized communication sounds are heard not only by the intended recipient, but also by the individual producing them. Such vocal feedback poses a number of unique challenges to the auditory system that must be overcome for successful communication. Foremost, one must be able to correctly determine that the perceived sounds are indeed self-produced, rather than coming from an external source (1,2). Second, it is important to maintain perception of external sounds during vocal production despite masking by loud self-produced sounds (3). Finally, it is critically important to monitor one's own vocal output to ensure accurate production (4). Dysfunction in such self-monitoring, as in hearing loss, leads humans to have difficulty in acquiring and maintaining normal speech (5–7). In the past 2 decades, interest has increased in understanding these auditory-vocal (sensorimotor) mechanisms and, in particular, the role of internal motor signals, such as corollary discharges, in modulating sensory processing in the auditory system during vocal production. In this review, we summarize recent work in a nonhuman primate model of vocal production, the marmoset monkey (*Callithrix jacchus*), which has become an important mammalian model for studying behavioral and neural mechanisms underlying vocal communication, and link these results with parallel findings in human speech.

ROLE OF VOCAL SELF-MONITORING BEHAVIORS IN NONHUMAN PRIMATES

While it is clear that vocal self-monitoring is important in human speech, the role of this process in nonhuman primates is

less well understood. In contrast to humans, songbirds, and some vocal species, deafening studies in nonhuman primates have not found any large changes in the overall acoustic structure of their vocalizations (8,9). However, marmosets and other nonhuman primates do show evidence of feedback-dependent vocal control when faced with changes in the acoustic environment or auditory feedback of their own vocalizations. This includes changing the timing or structure of their vocalizations when faced with interfering noise (10–12) and increasing their vocal amplitude in more modest background noise (13–15). Similar effects have been well documented in humans (16), and these behaviors may be intended to optimize communication in the setting of variable and interfering acoustic contexts. Marmosets also exhibit precise vocal control, changing their vocalization frequency contents to compensate for frequency-shifted feedback (17,18), similar to humans (7,19) and echolocating bats (20). However, the necessity or advantage of such precise vocal control is unclear. Marmosets do exhibit acoustic variations both between and within individual animals (21–23) and tend to imitate others with which they are paired (24). Whether there is information contained within such vocal acoustics remains a field of active study.

METHODOLOGICAL CHALLENGES IN DEVELOPING ANIMAL MODELS OF VOCAL PRODUCTION

Despite the importance of auditory-vocal mechanisms to human speech, we know relatively little about them. Much progress has been made in investigating these questions using human subjects, but there are experimental limitations in the degree of specificity and invasiveness that can be performed in humans. In parallel to these human studies, there has also been increasing use of nonhuman animals, which has the

potential to yield more detailed mechanistic information, take advantage of recent advances in genetic and other experimental techniques, and provide insight into the evolutionary origins of human speech.

There are, however, technical challenges in developing an appropriate animal model to investigate questions on vocal communication and better understand human speech. First, animal vocalizations do not fully reproduce the full complexity of speech, although they share many of the same underlying principles and mechanisms as human voice. Second, there are limitations in traditional animal models of vocal production. The songbird model, for example, shares some behavioral features of speech, including vocal sequence learning, and has yielded many important insights into vocal mechanisms (25). However, birdsong is different from speech in that most songbird species have only a fixed number of songs and is sex specific. Finally, there are considerable challenges in getting animals, in particular nonhuman primates, to readily vocalize under traditional experimental conditions, such as in a laboratory testing chamber. Potential solutions to these challenges include electrical stimulation of brainstem vocal areas (26–28) or operant-conditioning training (29–31), both of which can yield useful results but may bypass normal neural mechanisms and result in vocal production outside its natural behavioral context. A preferable approach is to perform experiments in behaviorally relevant conditions more conducive to vocal production, such as interactive antiphonal calling paradigms (32–34) or experiments in a social setting (e.g., housing colony) (35,36). Such approaches have proven successful, but required the development and adaptation of new experimental techniques, including specialized implanted microelectrode arrays (37) and neural recordings from unrestrained animals (38,39).

AUDITORY CORTEX ACTIVITY DURING VOCAL PRODUCTION IN MARMOSETS

The mammalian auditory pathway consists of multiple processing stages and potential sites for interactions between hearing and vocalization. Most recent research has focused on the role of the auditory cortex, a critical brain area necessary for sound perception. Recordings of single-unit activities in the auditory cortex have revealed the presence of two different types of neural responses during self-initiated vocal production (40,41). The first, and most common, is a vocalization-induced suppression of neural firing (Figure 1A). During this suppression, neural firing is reduced below the level of spontaneous firing, suggestive of direct cortical inhibition. This population of suppressed neurons accounts for between 65% and 90% of neurons in the auditory cortex of marmosets (36,40), indicating its importance in vocal production and related feedback control. More important, however, is that this vocalization-induced suppression actually begins before the onset of vocal production, on average approximately 200 ms, but as long as a second before (40). This prevocal timing suggests that suppression is not a sensory response coming from the ascending auditory pathway, but rather is caused by signals from the vocal production circuit, possibly a result of a motor-to-sensory corollary discharge. Vocalization-induced suppression also appears to be a general phenomenon, with neurons

suppressed by one type of vocalization also suppressed by other types (36).

In contrast to this dominant vocalization-induced suppression, the remaining smaller population of neurons in auditory cortex exhibit a vocalization-related excitation (Figure 1B). In contrast to suppression, excitation generally does not begin until after the onset of vocalization, suggesting that it may simply be a sensory response to vocal feedback from the ascending auditory pathway (40). In contrast to suppressed neurons, excited neurons are generally responsive only to a smaller set of vocalization types (36), likely limited by their sensory receptive fields (42). Neurons that switch responses between suppression and excitation for different types of vocalizations are uncommon and of unclear significance and origin.

This two-population model for auditory cortex responses during vocalization is similar to and consistent with activities in human recordings and imaging. Speech-induced suppression, similar to marmoset results, has been observed during human experiments using a variety of methods, including electroencephalography (37,43), magnetoencephalography (44–46), positron emission tomography (47,48), functional magnetic resonance imaging (49,50), and intracranial recordings (5,9,14,34,43,44,51–54). However, in contrast to the unit recordings in animals, human studies have largely been limited to showing a reduction in brain activity during speech production compared with simply hearing speech, although brain activity is still elevated compared with prespeech baseline (44,45,48,52,53,55–57). Unit recordings during primate vocalization exhibit responses that are not only reduced compared with hearing vocalizations (42), but also are reduced compared with spontaneous rates. The difference is likely methodological, with most human brain recording techniques lacking the resolution for single-unit activity and instead showing averaged activities of many neurons. First, the types of signals typically recorded in humans (blood flow, field potentials, oscillations) may reflect different aspects of the physiologic response from spiking units. Second, the averaging of a group of neurons with diverse responses may partially mask vocal suppression. For example, a dominant population of suppressed neurons will exhibit a maximal decrease in firing to zero from their typical spontaneous firing of <10 Hz. In contrast, a small (10%–35%) group of neurons with vocal increases is capable of reaching firing rates of 100 Hz or more. Under those circumstances, it is not difficult to extrapolate how the net average vocal firing rate might still be above spontaneous. Indeed, when we created a similar model, averaging activity from a large number of neurons in marmosets, including multiunit activities, the net vocal response was an increase in firing rate (58).

DIFFERENCES BETWEEN SUPPRESSED AND EXCITED NEURONS DURING VOCAL PRODUCTION

One immediate and important question is what the differences are between suppressed and excited neurons that give rise to their dichotomous behavior during vocal production. Some possibilities include anatomic location, sensory tuning properties, and connectivity. There are some differences in the estimated laminar location of these neurons, with suppressed neurons distributed mostly in upper cortical layers and less in

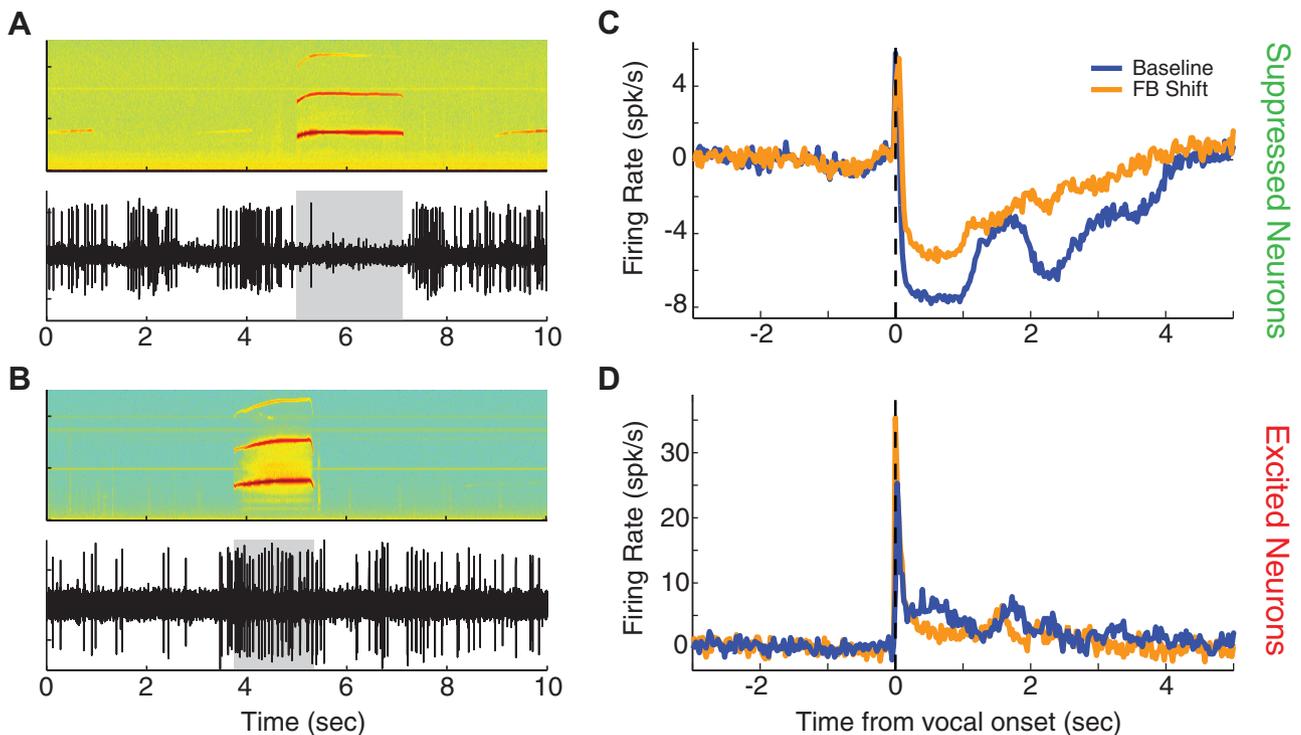


Figure 1. Suppressed and excited responses during vocalization. **(A)** Sample neuron recorded from marmoset auditory cortex (top) showing suppression during vocalization (bottom). **(B)** Sample neuron showing excitation during vocalization. **(C, D)** Vocal firing-rate responses from suppressed and excited auditory cortex neurons during normal (blue) and frequency-shifted vocal feedback (FB) (orange). Suppressed neurons **(C)** exhibit reduced suppression during shifted feedback, whereas excited neurons **(D)** are not strongly affected. spk/s, spikes per second. [Adapted with permission from Eliades and Wang (35).]

cortical input layer IV and below (58). However, no obvious differences based on hemisphere or cortical subfield (i.e., primary auditory cortex vs. more lateral areas) have been found (17,35,36). In fact, neurons recorded simultaneously from the same electrode often show no correlations in their vocal responses (58), suggestive of microarchitectural or connectivity variation to explain differences between vocal suppression and excitation. The bias toward suppression in the supragranular layers is of particular interest, given that these layers are the primary recipients of cortical-cortical connections, as opposed to ascending auditory inputs to layer IV (59,60). The microcircuits involved in vocal suppression, however, remain uncertain.

Sensory tuning properties may play a partial role in determining vocal responses. Although early analyses failed to reveal any relationship between basic acoustic response (frequency tuning, loudness sensitivity) and vocal suppression (40), more comprehensive analysis suggests that vocalization-related excitation may be explainable as a sensory response (42). First, although suppressed neurons exhibit variable responses to playback of vocal sounds, excited neurons are generally responsive to vocal playback (42). These excited neurons typically exhibit monotonic rate-level functions, increasing their firing rate with increasing sound level, which would make them more likely to respond to loud self-produced vocalizations. In contrast, suppressed neurons show a mix of monotonic and nonmonotonic (tuned/peaked) rate-level functions. Finally, computational models using frequency-tuning information to predict sensory responses to vocal sounds

show good prediction of both vocal playback and excitation during vocal production, but fail to predict vocalization-induced suppression (42). Together, these results suggest that the primary differences between excited and suppressed neurons is the presence of nonsensory (motor) inputs to suppressed neurons that are absent in the excited neurons. Excited neurons appear to respond similarly to vocal sounds whether or not they are self-produced or received externally.

EVIDENCE FOR A COROLLARY DISCHARGE MECHANISM FOR VOCALIZATION-INDUCED SUPPRESSION

Is a motor-to-sensory corollary discharge the cause of vocal suppression? Numerous pieces of evidence point toward this mechanism, as opposed to an alternative hypothesis of vocal suppression as a result of passive sensory responses. First, vocally suppressed neurons are generally responsive to vocal playback. Second, prevocal suppression is present, with no auditory inputs. Finally, suppressed vocal responses, in contrast to excited ones, are not well predicted from sensory tuning properties, although it might be possible for a corollary discharge to take advantage of underlying sensory tuning. Unfortunately, direct experimental evidence is thus far lacking.

One possible source of a corollary discharge is the frontal cortex, which is known to be involved in human speech production, in its reciprocal communication with the auditory cortex. To what extent the frontal cortex may be involved in

nonhuman primate vocalizations remains controversial (61). However, neurons in both prefrontal and premotor cortex in marmosets have been found to exhibit activity before and during vocal production (62–65), similar to what has been found in other nonhuman primates (30,31,66). There is also evidence that bidirectional anatomic connections exist between the frontal lobe and auditory cortex in both marmosets (60,67) and macaques (68,69). Stimulation of frontal cortex in primates can evoke inhibition in auditory cortex (70,71), but whether such connections are active during vocalization is unknown. Additionally, many of these frontal cortex regions, particularly the premotor cortex, exhibit both vocal premotor activity and auditory responses (30,62,66,72,73), making the frontal cortex a region of interest as a possible source for corollary discharge signals as well as a possible recipient area of auditory cortex vocal outputs. Similar frontotemporal connectivity has been seen in humans during speech, although its relationship to auditory processing is unknown (74–76). Some evidence may also be extrapolated from rodents, where optogenetic methods have shown an inhibitory motor to auditory connection during locomotion (77,78), which may represent an evolutionarily conserved sensorimotor pathway.

Another question is whether such a corollary discharge is acting directly on the auditory cortex or is targeting subcortical centers from which auditory cortex might inherit its vocal suppression. During vocal production, there is evidence for contraction of the middle ear muscles in multiple species (79–81) as well as reduced average auditory brainstem responses in bats (82,83). Although these subcortical mechanisms could account for a 20- to 40-dB reduction in effective sound level, modeling marmoset auditory cortical responses to sound level with these reductions does not account for the observed degree of vocalization-induced suppression (40). Moreover, such subcortical attenuation would not explain the close match in production and playback responses that is preserved in excited cortical neurons. As the middle ear muscles also similarly contract when hearing loud sounds (in the vocalization range), it is possible that a similar attenuation is present during both production and perception of speech or vocalization. Along the ascending auditory pathway, there is a notable absence of vocalization-induced suppression in the central nucleus of inferior colliculus, a midbrain auditory nucleus above the brainstem that is an obligate portion of the ascending auditory pathway, as demonstrated in the squirrel monkey (84–86), and would also argue against a lower pathway mechanism. Whether a thalamic or thalamocortical mechanism could be involved in producing vocalization-induced suppression, however, is uncertain, as only a single example of auditory thalamus recording during vocalization exists (87). Observations from recent rodent locomotion experiments argue for a direct cortical-cortical mechanism, as the optogenetic methods used have more specifically labeled and inactivated such direct connections (77,78), and similar cortical inhibition is seen even when ascending cortical inputs from the thalamus are artificially created (88).

ENVIRONMENTAL MONITORING DURING VOCAL PRODUCTION

What roles might these two populations of neurons play in solving the challenges faced by the auditory system during

vocal production? One such role is environmental monitoring, that is, maintaining sensitivity to the outside world during vocalization. When faced with an external sound that overlaps with loud self-produced vocalizations, the external sound could be masked. Surprisingly, relatively little is known about how large an issue this really is. One might predict that vocalization-induced suppression could be used to subtract out responses to self-produced vocal sounds, thereby improving the perception of concurrently received external sounds. This role has been suggested by studies in insects (89) and by tone detection studies in mice using locomotion-evoked sound presentation (90). Similar suppression of auditory responses has been found during human speech (91).

In contrast, on one hand, responses in marmoset auditory cortex during simultaneous vocalization and sound presentation have shown that the same vocalization-suppressed neurons also exhibit a parallel reduction in external sound-evoked responses (40). Vocalization-excited neurons, on the other hand, exhibited similar responses, on average, to sound presentation whether or not the animal was vocalizing. These findings suggest that the nonsuppressed neurons continue to encode outside sounds in a faithful fashion. It should be noted, however, that the presented stimuli were not systematically tested to determine if there was a degree of acoustic specificity to the reduction in external sound responses. It is also as yet unclear what mechanisms or features the auditory system uses to determine whether a sound is from an external source or is, in fact, self-produced.

VOCALIZATION-INDUCED SUPPRESSION IN SELF-MONITORING AND VOCAL CONTROL

A second, and likely more important, potential function of vocalization-induced suppression is to self-monitor one's own vocalization and to use this information to detect and correct any errors in vocal production. Such self-monitoring is best tested by perturbing the sounds that a human or an animal hears of itself (auditory feedback), which can be accomplished through a variety of acoustic manipulations. When presented with frequency-shifted auditory feedback during vocalization, there is a reduction in the degree of vocal suppression in the auditory cortex (17,35). This feedback sensitivity is most prominent for suppressed neurons and less evident for excited neurons, consistent with a possible role of vocal suppression in self-monitoring (Figure 1C, D). Moreover, feedback sensitivity appears not to be dependent on the underlying sensory tuning of these auditory cortex neurons, with no correlation between feedback responses and frequency tuning or responses to playback of vocal sounds. Direct comparisons of responses to frequency shifts during vocal production and playback of shifted vocal sounds reveal that suppressed neurons are actually more sensitive to vocal frequency shifts than would be predicted based on passive sensory responses. A similar feedback sensitivity is also seen in the human auditory cortex using a variety of experimental methods (43,55,92,93). In addition to sensitivity to frequency-shifted feedback, experiments using masking noise have shown a similar reduction in vocal suppression (15). Such masking is known to result in compensatory increases in vocal amplitude, a common behavior known as the Lombard effect (16,94) that is also

Vocal Corollary Discharge Mechanisms in Marmosets

present in marmosets and may be predicted by changes in auditory cortex activity (15).

These observations from feedback perturbation experiments suggest that vocalization-induced suppression somehow changes or biases sensory processing in the auditory system to allow for more accurate self-monitoring. The origin and mechanisms of this altered sensory processing are unclear, but it has been suggested that it is a result of a corollary discharge containing a specific sensory prediction of the motor action plan, the latter termed an efferent copy (95) or forward model (96,97,98). There has been considerable variability and confusion in the use of these terms, but here we focus on the sensory prediction of the expected auditory feedback during vocalization. In theory, the sensory prediction could be compared with auditory inputs resulting in an auditory cortical response reflecting the feedback error—a so-called error signal. Whereas vocalization-induced suppression is reduced during altered feedback, as predicated by an error hypothesis, direct evidence for such an error signal is more limited. For example, an error signal would be expected to reflect the magnitude of the difference between vocal production and feedback, and some human data suggest such a role (43,98–100). Marmoset recordings have thus far been limited to a single feedback manipulation, and so a parallel process is unclear. It is possible that both marmoset and human vocal responses reflect a nonspecific suppression combined with underlying sensory tuning, rather than a specific prediction. Only studies systematically testing individual neurons with a wide range of feedback shifts, as opposed to passive sensory responses to similar frequency shifts, will better distinguish these two hypotheses.

A similar phenomenon is observed under conditions in which nonvocal actions result in a predictable sensory input. For example, tone presentation to locomoting rodents suggest a suppression of sensory responses that is less evident when tone frequencies do not match those that were expected or predicted (90). This finding closely parallels observed sensitivity to frequency-shifted vocal feedback seen in both humans (4,7,11,55,93,101) and marmosets (17,35). An intriguing question is whether these represent a more general sensory-prediction mechanism, of which speech and vocal production are just special cases.

Given this evidence of neural responses reflecting self-monitoring in auditory cortex of both humans and marmosets, the question arises as to whether marmosets are able to control their vocal production and what role auditory cortical activity might play in such vocal control. Humans are capable of rapidly adjusting the acoustics of their voice and speech to compensate for perceived changes in vocal feedback (19,102,103). Using frequency-shifted feedback, triggered to begin in the middle of vocalization, recent studies have demonstrated that marmosets also exhibit compensatory vocal control, raising their vocal frequency when presented with downward feedback shifts or decreasing frequency when faced with upward feedback shifts (17). The magnitude (approximately 0.3–0.4 semitones) and latency (approximately 200 ms) of this vocal compensation is comparable to that of humans (19,102,104), suggesting that marmosets are a good model for such feedback-dependent vocal control. Auditory cortical recordings performed during this compensatory vocal

control also show that there is a greater decrease in vocal suppression in trials in which marmosets exhibit greater compensation (17), a pattern also seen using human intracranial recordings (55,92). This difference in suppression precedes vocal changes and, in fact, predicts them, suggesting a possible role in a vocal control circuit involving the auditory cortex. Additionally, different groups of neurons were found that predicted both positive and negative directions of vocal compensation, perhaps a push-pull system for vocal frequency control.

However, establishing a causal role for auditory cortex and vocalization-induced suppression in self-monitoring and vocal control requires direct manipulation of cortical activity. Experiments using electrical microstimulation of auditory cortex during vocalization demonstrate rapid changes in vocal production and provide early evidence for such causal inference (17). This stimulation evokes vocal changes with a latency of approximately 40 ms, slightly longer than the latency from laryngeal motor cortex in humans and other species (105–107), suggesting a short reflex pathway between the auditory cortex and larynx. Moreover, the effects of this stimulation on vocal acoustics significantly correlate with neural responses recorded from the same electrode locations, suggesting that stimulation may have activated the same cortical circuits and resulting in a similar behavioral outcome. Interestingly, stimulation effects were strongly biased toward the right hemisphere, suggesting a potential hemispherical lateralization that is consistent with some previous human auditory cortex studies (47,92,108) and specializations of the right asymmetry of the motor cortex in vocal pitch control for both humans (106) and monkeys (109). Despite this evidence, it still remains unclear whether the auditory cortex is necessary for vocal control or whether a subcortical mechanism might also be involved.

BEYOND THE AUDITORY SYSTEM

What other brain areas might be involved in vocal corollary discharge mechanisms and auditory-vocal comparisons? Two possibilities are the basal ganglia and the cerebellum, but neither has been well studied in mammalian vocal production. The basal ganglia have long been studied in the songbird literature for their role in song sequence learning and control (25) but have largely been ignored in nonhuman primates. However, the basal ganglia are reciprocally connected with the auditory cortex and exhibit responses to sound stimuli (110,111). The cerebellum has a long history of study in sensorimotor behavior and coordination, and anatomic tracing and lesion studies have implicated it in the vocal motor pathway (29,112–115), but its role in auditory-vocal processing is unknown. These areas remain intriguing avenues for future investigation.

CONCLUSIONS

Suppression of the neural activity in auditory cortex during both speech and vocalization is a ubiquitous phenomenon for both humans and nonhuman primates. Evidence suggests that this suppression is mediated by a corollary discharge mechanism, likely originating in brain areas that initiate or control vocal production, that is likely conserved across species. This mechanism appears to play a critical role in self-monitoring

during vocalization and resulting feedback-dependent control of speech and vocal production. Further comparative approaches studying vocal production in nonhuman primates at circuit and cellular levels may yield valuable insights into the evolutionary origins of human speech and its control circuit in the brain as well as the underlying mechanisms of many vocal disorders.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institutes of Health (Grant No. DC014299 [to SJE] and Grant No. DC005808 [to XW]).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Auditory and Communication Systems Laboratory (SJE), Department of Otorhinolaryngology: Head and Neck Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; and Laboratory of Auditory Neurophysiology (XW), Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Address correspondence to Steven J. Eliades, M.D., Ph.D., Department of Otorhinolaryngology: Head and Neck Surgery, 3400 Spruce Street, 5 Ravdin, Philadelphia, PA 19104; E-mail: seliades@pennmedicine.upenn.edu.

Received Apr 22, 2019; revised and accepted Jun 24, 2019.

REFERENCES

- Johns LC, McGuire PK (1999): Verbal self-monitoring and auditory hallucinations in schizophrenia. *Lancet* 353:469–470.
- Frith CD (1992): *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Earlbaum Associates.
- Bekesy Gv (1949): The structure of the middle ear and the hearing of one's own voice by bone conduction. *J Acoust Soc Am* 21:217–232.
- Levelt WJ (1983): Monitoring and self-repair in speech. *Cognition* 14:41–104.
- Smith CR (1975): Residual hearing and speech production in deaf children. *J Speech Hear Res* 18:795–811.
- Gold T (1980): Speech production in hearing-impaired children. *J Commun Disord* 13:397–418.
- Lane H, Webster JW (1991): Speech deterioration in postlingually deafened adults. *J Acoust Soc Am* 89:859–866.
- Talmage-Riggs G, Winter P, Ploog D, Mayer W (1972): Effect of deafening on the vocal behavior of the squirrel monkey (*Saimiri sciureus*). *Folia Primatol (Basel)* 17:404–420.
- Egnor SE, Hauser MD (2004): A paradox in the evolution of primate vocal learning. *Trends Neurosci* 27:649–654.
- Roy S, Miller CT, Gottsch D, Wang X (2011): Vocal control by the common marmoset in the presence of interfering noise. *J Exp Biol* 214:3619–3629.
- Egnor SE, Wickelgren JG, Hauser MD (2007): Tracking silence: Adjusting vocal production to avoid acoustic interference. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 193:477–483.
- Pomberger T, Risueno-Segovia C, Loschner J, Hage SR (2018): Precise motor control enables rapid flexibility in vocal behavior of marmoset monkeys. *Curr Biol* 28:788–794.e783.
- Egnor SE, Hauser MD (2006): Noise-induced vocal modulation in cotton-top tamarins (*Saguinus oedipus*). *Am J Primatol* 68:1183–1190.
- Brumm H, Voss K, Kollmer I, Todt D (2004): Acoustic communication in noise: Regulation of call characteristics in a New World monkey. *J Exp Biol* 207:443–448.
- Eliades SJ, Wang X (2012): Neural correlates of the lombard effect in primate auditory cortex. *J Neurosci* 32:10737–10748.
- Lane H, Tranel B (1971): The Lombard sign and the role of hearing in speech. *J Speech Hear Res* 14:677–709.
- Eliades SJ, Tsunada J (2018): Auditory cortical activity drives feedback-dependent vocal control in marmosets. *Nat Commun* 9:2540.
- Zhao L, Rad BB, Wang X (2019): Long-lasting vocal plasticity in adult marmoset monkeys. *Proc R Soc B* 296:20198017.
- Bauer JJ, Larson CR (2003): Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *J Acoust Soc Am* 114:1048–1054.
- Schuller G, Beuter K, Schnitzler H-U (1974): Response to frequency shifted artificial echoes in the bat *Rhinolophus ferrumequinum*. *J Comp Physiol* 89:275–286.
- Agamaite JA, Chang CJ, Osmanski MS, Wang X (2015): A quantitative acoustic analysis of the vocal repertoire of the common marmoset (*Callithrix jacchus*). *J Acoust Soc Am* 138:2906.
- Pistorio AL, Vintch B, Wang X (2006): Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*). *J Acoust Soc Am* 120:1655–1670.
- DiMattina C, Wang X (2006): Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 95:1244–1262.
- Snowdon CT, Elowson AM (1999): Pygmy marmosets modify call structure when paired. *Ethology* 105:893–908.
- Doupe AJ, Kuhl PK (1999): Birdsong and human speech: Common themes and mechanisms. *Annu Rev Neurosci* 22:567–631.
- Jurgens U (1994): The role of the periaqueductal grey in vocal behaviour. *Behav Brain Res* 62:107–117.
- Jurgens U (2000): Localization of a pontine vocalization-controlling area. *J Acoust Soc Am* 108:1393–1396.
- Jurgens U, Pratt R (1979): Role of the periaqueductal grey in vocal expression of emotion. *Brain Res* 167:367–378.
- Hage SR, Gavrilov N, Nieder A (2013): Cognitive control of distinct vocalizations in rhesus monkeys. *J Cogn Neurosci* 25:1692–1701.
- Hage SR, Nieder A (2013): Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat Commun* 4:2409.
- Coudé G, Ferrari PF, Roda F, Maranesi M, Borelli E, Veroni V, et al. (2011): Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One* 6:e26822.
- Miller CT, Wang X (2006): Sensory-motor interactions modulate a primate vocal behavior: Antiphonal calling in common marmosets. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 192:27–38.
- Norcross JL, Newman JD (1997): Social context affects phee call production by nonreproductive common marmosets (*Callithrix jacchus*). *Am J Primatol* 43:135–146.
- Ghazanfar AA, Flombaum JI, Miller CT, Hauser MD (2001): The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): Playback experiments with long calls. *J Comp Physiol A* 187:27–35.
- Eliades SJ, Wang X (2008): Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453:1102–1106.
- Eliades SJ, Wang X (2013): Comparison of auditory-vocal interactions across multiple types of vocalizations in marmoset auditory cortex. *J Neurophysiol* 109:1638–1657.
- Eliades SJ, Wang X (2008): Chronic multi-electrode neural recording in free-roaming monkeys. *J Neurosci Methods* 172:201–214.
- Mohseni P, Najafi K, Eliades SJ, Wang X (2005): Wireless multi-channel biopotential recording using an integrated FM telemetry circuit. *IEEE Trans Neural Syst Rehabil Eng* 13:263–271.
- Roy S, Wang X (2012): Wireless multi-channel single unit recording in freely moving and vocalizing primates. *J Neurosci Methods* 203:28–40.
- Eliades SJ, Wang X (2003): Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J Neurophysiol* 89:2194–2207.
- Muller-Preuss P, Ploog D (1981): Inhibition of auditory cortical neurons during phonation. *Brain Res* 215:61–76.
- Eliades SJ, Wang X (2017): Contributions of sensory tuning to auditory-vocal interactions in marmoset auditory cortex. *Hear Res* 348:98–111.

Vocal Corollary Discharge Mechanisms in Marmosets

43. Behroozmand R, Karvelis L, Liu H, Larson CR (2009): Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin Neurophysiol* 120:1303–1312.
44. Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R (2000): Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp* 9:183–191.
45. Houde JF, Nagarajan SS, Sekihara K, Merzenich MM (2002): Modulation of the auditory cortex during speech: An MEG study. *J Cogn Neurosci* 14:1125–1138.
46. Numminen J, Curio G (1999): Differential effects of overt, covert and replayed speech on vowel-evoked responses of the human auditory cortex. *Neurosci Lett* 272:29–32.
47. Wise RJ, Greene J, Buchel C, Scott SK (1999): Brain regions involved in articulation. *Lancet* 353:1057–1061.
48. Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC (1996): Modulation of cerebral blood flow in the human auditory cortex during speech: Role of motor-to-sensory discharges. *Eur J Neurosci* 8:2236–2246.
49. Christoffels IK, Formisano E, Schiller NO (2007): Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Hum Brain Mapp* 28:868–879.
50. Fu CH, Vythelingum GN, Brammer MJ, Williams SC, Amaro E Jr, Andrew CM, *et al.* (2006): An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cereb Cortex* 16:969–977.
51. Creutzfeldt O, Ojemann G, Lettich E (1989): Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Exp Brain Res* 77:476–489.
52. Flinker A, Chang EF, Kirsch HE, Barbaro NM, Crone NE, Knight RT (2010): Single-trial speech suppression of auditory cortex activity in humans. *J Neurosci* 30:16643–16650.
53. Greenlee JD, Jackson AW, Chen F, Larson CR, Oya H, Kawasaki H, *et al.* (2011): Human auditory cortical activation during self-vocalization. *PLoS One* 6:e14744.
54. Chen CM, Mathalon DH, Roach BJ, Cavus I, Spencer DD, Ford JM (2011): The corollary discharge in humans is related to synchronous neural oscillations. *J Cogn Neurosci* 23:2892–2904.
55. Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF (2013): Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A* 110:2653–2658.
56. Ford JM, Mathalon DH, Heinks T, Kalba S, Faustman WO, Roth WT (2001): Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *Am J Psychiatry* 158:2069–2071.
57. Crone NE, Hao L, Hart J Jr, Boatman D, Lesser RP, Irizarry R, *et al.* (2001): Electrographic gamma activity during word production in spoken and sign language. *Neurology* 57:2045–2053.
58. Eliades SJ, Wang X (2005): Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cereb Cortex* 15:1510–1523.
59. de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2006): Thalamic connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *J Comp Neurol* 496:72–96.
60. de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2006): Cortical connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *J Comp Neurol* 496:27–71.
61. Hage SR, Nieder A (2016): Dual neural network model for the evolution of speech and language. *Trends Neurosci* 39:813–829.
62. 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*). *J Neurosci* 36:12168–12179.
63. Miller CT, Dimauro A, Pistorio A, Hendry S, Wang X (2010): Vocalization induced cfos expression in marmoset cortex. *Front Integr Neurosci* 4:128.
64. Miller CT, Thomas AW, Nummela SU, de la Mothe LA (2015): Responses of primate frontal cortex neurons during natural vocal communication. *J Neurophysiol* 114:1158–1171.
65. Simões CS, Vianney PV, de Moura MM, Freire MA, Mello LE, Sameshima K, *et al.* (2010): Activation of frontal neocortical areas by vocal production in marmosets. *Front Integr Neurosci* 4.
66. Gavrilov N, Hage SR, Nieder A (2017): Functional specialization of the primate frontal lobe during cognitive control of vocalizations. *Cell Rep* 21:2393–2406.
67. de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2012): Cortical connections of auditory cortex in marmoset monkeys: Lateral belt and parabelt regions. *Anat Rec* 295:800–821.
68. Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999): Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 2:1131–1136.
69. Petrides M, Pandya DN (1988): Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J Comp Neurol* 273:52–66.
70. Alexander GE, Newman JD, Symmes D (1976): Convergence of prefrontal and acoustic inputs upon neurons in the superior temporal gyrus of the awake squirrel monkey. *Brain Res* 116:334–338.
71. Muller-Preuss P, Newman JD, Jurgens U (1980): Anatomical and physiological evidence for a relationship between the ‘cingular’ vocalization area and the auditory cortex in the squirrel monkey. *Brain Res* 202:307–315.
72. Hage SR, Nieder A (2015): Audio-vocal interaction in single neurons of the monkey ventrolateral prefrontal cortex. *J Neurosci* 35:7030–7040.
73. Hage SR (2018): Auditory and audio-vocal responses of single neurons in the monkey ventral premotor cortex. *Hear Res* 366:82–89.
74. Ford JM, Gray M, Faustman WO, Heinks TH, Mathalon DH (2005): Reduced gamma-band coherence to distorted feedback during speech when what you say is not what you hear. *Int J Psychophysiol* 57:143–150.
75. Ford JM, Mathalon DH, Whitfield S, Faustman WO, Roth WT (2002): Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biol Psychiatry* 51:485–492.
76. Flinker A, Korzeniewska A, Shestuyk AY, Franaszczuk PJ, Dronkers NF, Knight RT, *et al.* (2015): Redefining the role of Broca’s area in speech. *Proc Natl Acad Sci U S A* 112:2871–2875.
77. Nelson A, Schneider DM, Takatoh J, Sakurai K, Wang F, Mooney R (2013): A circuit for motor cortical modulation of auditory cortical activity. *J Neurosci* 33:14342–14353.
78. Schneider DM, Nelson A, Mooney R (2014): A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513:189–194.
79. Henson OW Jr (1965): The activity and function of the middle-ear muscles in echo-locating bats. *J Physiol* 180:871–887.
80. Salomon B, Starr A (1963): Electromyography of middle ear muscles in man during motor activities. *Acta Neurol Scand* 39:161–168.
81. Suga N, Jen PH (1975): Peripheral control of acoustic signals in the auditory system of echolocating bats. *J Exp Biol* 62:277–311.
82. Suga N, Schlegel P (1972): Neural attenuation of responses to emitted sounds in echolocating bats. *Science* 177:82–84.
83. Suga N, Shimozawa T (1974): Site of neural attenuation of responses to self-vocalized sounds in echolocating bats. *Science* 183:1211–1213.
84. Kirzinger A, Jurgens U (1991): Vocalization-correlated single-unit activity in the brain stem of the squirrel monkey. *Exp Brain Res* 84:545–560.
85. Pieper F, Jurgens U (2003): Neuronal activity in the inferior colliculus and bordering structures during vocalization in the squirrel monkey. *Brain Res* 979:153–164.
86. Tammer R, Ehrenreich L, Jurgens U (2004): Telemetrically recorded neuronal activity in the inferior colliculus and bordering tegmentum during vocal communication in squirrel monkeys (*Saimiri sciureus*). *Behav Brain Res* 151:331–336.
87. Muller-Preuss P (1981): Acoustic properties of central auditory pathway neurons during phonation in the squirrel monkey. In: Syka J, Aitkin L, editors. *Neuronal Mechanisms of Hearing*. New York: Plenum, 311–315.
88. Rummell BP, Klee JL, Sigurdsson T (2016): Attenuation of responses to self-generated sounds in auditory cortical neurons. *J Neurosci* 36:12010–12026.

89. Poulet JF, Hedwig B (2002): A corollary discharge maintains auditory sensitivity during sound production. *Nature* 418:872–876.
90. Schneider DM, Sundararajan J, Mooney R (2018): A cortical filter that learns to suppress the acoustic consequences of movement. *Nature* 561:391–395.
91. Daliri A, Max L (2015): Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain Lang* 143:59–68.
92. Behroozmand R, Oya H, Nourski KV, Kawasaki H, Larson CR, Brugge JF, *et al.* (2016): Neural correlates of vocal production and motor control in human Heschl's gyrus. *J Neurosci* 36:2302–2315.
93. Greenlee JD, Behroozmand R, Larson CR, Jackson AW, Chen F, Hansen DR, *et al.* (2013): Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLoS One* 8:e60783.
94. Luo J, Hage SR, Moss CF (2018): The Lombard effect: From acoustics to neural mechanisms. *Trends Neurosci* 41:938–949.
95. von Holst E, Mittelstaedt H (1950): Das Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften* 37:464–476.
96. Wolpert DM, Ghahramani Z, Jordan MI (1995): An internal model for sensorimotor integration. *Science* 269:1880–1882.
97. Wolpert DM, Miall RC (1996): Forward models for physiological motor control. *Neural Netw* 9:1265–1279.
98. Niziolek CA, Nagarajan SS, Houde JF (2013): What does motor efference copy represent? Evidence from speech production. *J Neurosci* 33:16110–16116.
99. Liu H, Meshman M, Behroozmand R, Larson CR (2011): Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clin Neurophysiol* 122:951–957.
100. Behroozmand R, Larson CR (2011): Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neurosci* 12:54.
101. Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM (2005): Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42:180–190.
102. Burnett TA, Freedland MB, Larson CR, Hain TC (1998): Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am* 103:3153–3161.
103. Houde JF, Jordan MI (1998): Sensorimotor adaptation in speech production. *Science* 279:1213–1216.
104. Larson CR, Burnett TA, Bauer JJ, Kiran S, Hain TC (2001): Comparison of voice F0 responses to pitch-shift onset and offset conditions. *J Acoust Soc Am* 110:2845–2848.
105. Espadaler J, Rogic M, Deletis V, Leon A, Quijada C, Conesa G (2012): Representation of cricothyroid muscles at the primary motor cortex (M1) in healthy subjects, mapped by navigated transcranial magnetic stimulation (nTMS). *Clin Neurophysiol* 123:2205–2211.
106. Dichter BK, Breshears JD, Leonard MK, Chang EF (2018): The control of vocal pitch in human laryngeal motor cortex. *Cell* 174:21–31.e29.
107. Haghghi SS, Estrem SA (1991): Comparison of evoked electromyography of the larynx to electrical and magnetic stimulation of the motor cortex of the dog. *Laryngoscope* 101:68–70.
108. Kort NS, Nagarajan SS, Houde JF (2014): A bilateral cortical network responds to pitch perturbations in speech feedback. *Neuroimage* 86:525–535.
109. Jurgens U, Zwirner P (2000): Individual hemispheric asymmetry in vocal fold control of the squirrel monkey. *Behav Brain Res* 109:213–217.
110. Znamenskiy P, Zador AM (2013): Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497:482–485.
111. Chen L, Wang X, Ge S, Xiong Q (2019): Medial geniculate body and primary auditory cortex differentially contribute to striatal sound representations. *Nat Commun* 10:418.
112. Jurgens U (2002): Neural pathways underlying vocal control. *Neurosci Biobehav Rev* 26:235–258.
113. Kirzinger A (1985): Cerebellar lesion effects on vocalization of the squirrel monkey. *Behav Brain Res* 16:177–181.
114. Konopka G, Roberts TF (2016): Animal models of speech and vocal communication deficits associated with psychiatric disorders. *Biol Psychiatry* 79:53–61.
115. Larson CR, Sutton D, Lindeman RC (1978): Cerebellar regulation of phonation in rhesus monkey (*Macaca mulatta*). *Exp Brain Res* 33:1–18.