The neurobiology of primate vocal communication

Asif A. Ghazanfar[1,2,*] and **Steven J. Eliades**[4]

[1]Princeton Neuroscience Institute, Department of Psychology, Princeton University, Princeton NJ 08544, USA, Ph. 609 258 9314

[2]Department of Ecology & Evolutionary Biology, Princeton University, Princeton NJ 08544, USA, Ph. 609 258 9314

[4]Department of Otorhinolaryngology: Head and Neck, University of Pennsylvania Perelman School of Medicine, 3400 Spruce St, 5 Ravdin, Philadelphia, PA 19104

Abstract

Recent investigations of non-human primate communication revealed vocal behaviors far more complex than previously appreciated. Understanding the neural basis of these communicative behaviors is important as it has the potential to reveal the basic underpinnings of the still more complex human speech. The latest work revealed vocalization-sensitive regions both within and beyond the traditional boundaries of the central auditory system. The importance and mechanisms of multi-sensory face-voice integration in vocal communication are also increasingly apparent. Finally, studies on the mechanisms of vocal production demonstrated auditory-motor interactions that may allow for self-monitoring and vocal control. We review the current work in these areas of primate communication research.

Introduction

Primates typically live in large groups and maintain cohesion in their groups with moment-to-moment social interactions and using the specialized signaling that such interactions require [1]. In a dynamic social environment, it is essential that primates are well equipped for detecting, learning and discriminating communication signals. Primates need to be able to produce signals accurately (both in terms of signal structure and context) and they need to be able to respond to these signals adaptively.

Many of the signals that primates exchange take the form of vocalizations [1]. Indeed, in anthropoid primates, as group size grows, the complexity of vocal expressions grow as well [2,3]. Despite many years of behavioral studies focused on primate vocal communication, their neurobiology remains ill understood and under-studied. Historically, non-human primate vocal communication has been thought of as mainly under subcortical control, and

*To whom correspondence should be addressed: asifg@princeton.edu.

these subcortical mechanisms have been worked out in some detail [4,5](for review, see [6]). Nevertheless, recent work, and an increasing interest in the set of problems related to the neurobiology of primate communication, has begun to identify some possible neural mechanisms at the level of the cerebral cortex. We review some research efforts examining cortical encoding of primate vocalizations, the multisensory integration of vocal sounds and their associated facial expressions, and self-monitoring during vocal production. We primarily focus on neural mechanisms related to vocal perception in primates, how these might contribute to vocal production and how they might relate to human speech (though we acknowledge that primate vocalizations may have homologies to other human vocal sounds, as well; see, for example,[7]). Vocal perception seems to show greater cross-species homologies than the less well understood, and under-studied, process of vocal production.

Vocalization- and voice-sensitivity in the neocortex

Figure 1 shows the regional divisions of the primate auditory cortex, located in the lower bank of the lateral sulcus (also known as the 'superior temporal plane'); there are certainly other auditory cortical regions both medial and lateral to these regions, but they are not well characterized. We have known for quite some time that neurons in the various regions of superior temporal plane, including the primary auditory cortex (A1), are sensitive to conspecific vocalizations [8–10]. Neuroimaging studies in both humans and monkeys suggest that there is caudal to rostral gradient of vocalization-selectivity (or speech, in the case of humans) within the temporal lobe [11,12]. In neurophysiological investigations of this gradient, single neuron responses to vocalizations and other salient sounds in the macaque monkey auditory cortex are more selective in rostral regions of the superior temporal plane (the 'rostrotemporal polar area', RTp) than in more caudal regions like the primary auditory cortex (A1) [13]. This caudal to rostral gradient is reflected not only in differential firing rate profiles across the stimulus set but also by longer neural response latencies. The longer latencies suggest that there may be a serial, feed-forward pathway related to increasing vocalization-selectivity.

Avoiding some of the limitations of sparse and biased sampling that typifies the single electrode, single neuron approach, a recent study used high density microelectrocorticographic arrays and measured the selectivity of auditory evoked potentials across a number of rhesus macaque auditory cortical fields simultaneously [14*]. A statistical classifier measured differential patterns of neuronal population activity in response to monkey vocalizations and synthetic control sounds that were based on either the spectral or temporal features of the monkey calls. While the information about these classes of sounds was equivalent in the caudal regions, more rostral cortical regions classified real vocalizations better than their synthetic counterparts. Interestingly, breaking down the evoked responses into band-passed local field potentials (LFPs) revealed that the theta frequency band in this rostral region carried the most information about vocalizations. This is consistent with the idea that the structure of monkey vocalizations (like the structure of speech) may exploit neural activity in the theta band [15,16].

Vocalizations also carry voice-content (or "indexical cues") related to the identity and physical characteristics of the speaker [17,18] and this rostral region of the superior

temporal plane has a sector containing neurons that seems particularly sensitive to voices (Figure 1). A functional imaging study using macaques compared auditory responses to their own species-specific calls with control sounds that had the same spectral profile and duration, other animal vocalizations and natural sounds [19]. Macaques, it seems, do have a voice area that is especially sensitive to conspecific vocalizations in the same manner as the human voice region [20] (though the anatomical location is different; [21]). It may be that this voice area is one that is sensitive to the formant structure embedded in vocalizations. Formants are the acoustic products of sound filtering caused by the shape and length of the vocal tract (the oral and nasal cavities above the larynx) [22,23]. Because an individual's vocal tract is uniquely shaped and has a length dependent upon body size [24], formants are acoustic cues to both individual identity and other physical characteristics [25–27]. This idea was indirectly confirmed by data showing that the BOLD response of the voice area, in essence, habituates to different calls (e.g. a grunt and a coo call) from the same individual (and thus similar formant signatures), but does not habituate when two calls of the same category (a coo and a coo) but from different individuals (and thus, different formant signatures) are presented [19].

The neurophysiological responses within this sector support the notion that it is region specialized for voice processing [28]. FMRI-guided placement of electrodes within this sector revealed "voice cells" with selectivity on par with those of face cells elsewhere in the temporal lobe. These neurons responded more strongly to monkey vocalizations than to other animal vocalizations or other environmental sounds. It is unclear how sensory processing of this vocalization- and voice-sensitive rostral region of auditory cortex relates to or differs from processes in the prefrontal cortex that are presumed to relate to higher order signal processing functions. Neurons in the ventrolateral prefrontal cortex are also very selective to vocalizations [29,30], but not to any greater or lesser to degree than neurons in some auditory cortical regions. Presumably, prefrontal cortical neurons also encode voice content, though this remains to be tested.

Though we have learned quite a bit in the last decade about the neurophysiological processing of primate vocalizations, a significant limitation of these studies is that they don't require the monkeys to perform a sophisticated behavioral task, one that could relate neural activity to behavioral performance and meaningful perception of the communication sounds. Although macaques are notoriously difficult to train on some auditory tasks, to do so is not impossible. Indeed, it has been shown that prefrontal cortical activity is related to behavioral choices in an auditory oddball task using speech stimuli [31].

Integration of faces and voices

While facial and vocal expressions are typically treated separately in most studies, in fact, they are often inextricably linked: a vocal expression cannot be produced without concomitant movements of the face. Given that vocalizations are physically linked to different facial expressions, it is perhaps not surprising that many primates recognize the correspondence between the visual and auditory components of vocal signals [25,32–36]. Moreover, a recent vocal detection study of macaque monkeys revealed that, like humans,

monkeys exhibited greater accuracy and faster reaction times to audiovisual vocalizations than to unisensory events [37].

How the brain mediates the behavioral benefits from integrating signals from different modalities is the subject of intense debate and investigation [38]. For face/voice integration, traditional models emphasize the role of association areas embedded in the temporal, frontal and parietal lobes [39]. Although these regions certainly play important roles, numerous recent studies demonstrate that they are not the sole regions for multisensory convergence [40]. The auditory cortex, in particular, has many potential sources of visual input, and an increasing number of studies, in both humans and nonhuman primates, demonstrate that dynamic faces influence auditory cortical activity [41].

However, the relationship between multisensory behavioral performance and neural activity in auditory cortex remains unknown. This is for two reasons. First, methodologies typically used to study auditory cortex in humans are unable to resolve neural activity at the level of action potentials. Second, until recently, all face/voice neurophysiological studies in monkeys-- regardless of the brain areas explored--have not required monkeys to perform a multisensory task. This is true not only for auditory cortical studies [42–46], but for studies of association areas as well [47–49]. All of these physiological studies demonstrated that neural activity in response to faces combined with voices is integrative, exhibiting both enhanced and suppressed changes in response magnitude when multisensory conditions are compared to unisensory ones. It is presumed that such firing rate changes mediate behavioral benefits of multisensory signals, but there is the possibility that integrative neural responses—particularly in the auditory cortex—are epiphenomenal.

To bridge this gap in our knowledge, a recent study combined an audiovisual vocal detection task with auditory cortical physiology in macaque monkeys [50**]. Monkeys were trained to detect visual, auditory and audiovisual presentations of vocalizations in a free- response task in which there are no explicit trial markers (Figure 2A). The task approximates natural face-to-face communication, where the timing of vocalizations is not entirely predictable and in which varying levels of noise degrade the acoustic components of vocalizations but the face and its motion are perceived clearly. This is typical of macaque monkey social interactions. When detecting voices alone, the signal-to-noise ratio systematically influences the monkeys' behavioral performance (Figure 2B), and these same systematic effects are observed in the latency of spiking activity in the lateral belt of auditory cortex (Figure 2C). The addition of a dynamic face leads faster reaction times and better accuracy *and* to audiovisual neural responses that were faster than auditory-only responses. That is, dynamic faces reduce the latency of auditory cortical spiking activity. Surprisingly, dynamic faces did not systematically change firing rate magnitude or variability (Figure 2C). These data suggest a novel latency facilitation role for visual influences on auditory cortex during audiovisual vocal detection. Facial motion speeds up the spiking responses of auditory cortex while leaving firing rate magnitudes largely unchanged [50**].

How different auditory cortical areas relate to different task requirements is largely a mystery. Thus, while facial motion may speed up lateral belt neurons in a vocal detection task, it is not clear how such neurons may respond under different task conditions. There is

evidence to suggest that the neurophysiology of face/voice integration is not only different between primary and lateral belt auditory cortical regions [43], but also between auditory cortical regions and the superior temporal sulcus (STS) and prefrontal cortex [46,47,49**], and even within the different neural frequency bands of a single area [49**]. By comparing it to the upper bank of the STS, a recent multisensory study of the macaque "voice" area (see above) revealed just how the differential sensitivity of one region over another may influence patterns of face/voice integration [46]. There were many more bimodal neurons in the STS than in the voice area, as expected given that the STS is convergence zone of sensory inputs from different modalities. Other differences were unexpected. Though the STS is considered a higher-order association area, neurons in this region were not as sensitive to voice identity, vocalization type, caller species or familiarity as neurons in the voice region. Conversely, neurons in the STS were more sensitive to face/voice congruency than were voice region neurons. How these differences manifest themselves in relation to vocal behavior is unknown.

Sensory-motor mechanisms and self-monitoring

Successful vocal communication also depends on subjects ensuring the accuracy and fidelity of their communication signals. Because vocalizations project both to the intended recipient as well as back to the individual producing them, such vocal feedback can be monitored by the producer and used to adjust acoustic structure according to context. Our knowledge of this feedback-dependent vocal control is well-established in humans [51], but non-human primate vocalization have traditionally been thought of as less dependent upon feedback or self-monitoring [52]. Recent evidence obligates us to reconsider this assumption.

Primates exhibit a robust increase in vocal intensity in the presence of masking noise [53,54], a fundamental auditory-vocal behavior known as the Lombard effect [55]. New World primates have also been shown to adjust the timing of their vocalizations to avoid interference from noise [54,56,57], or even terminate ongoing vocal production in the presence of overwhelming masking [58]. More recently, marmoset monkeys were observed to take turns vocally in a manner very similar to human conversations (albeit on a different timescale) [59*]. These vocal exchanges (like human social interactions) exhibit coupled oscillator dynamics. Whether or not there is on-line control of other parameters of vocalization, such as the frequency spectrum, is not yet known. Other evidence demonstrated that nonhuman primates they do not seem to learn their vocalizations, they do seem capable of adjusting their vocal acoustics to social peers [60,61]. While this vocal acoustic flexibility does not necessarily require the on-line self-monitoring needed during feedback vocal control, it does suggest a more general self-monitoring or awareness of an animal's own vocal sounds compared to those around it.

Despite this behavioral evidence for self-monitoring, the sensory-motor neural mechanisms underlying this process remain unclear. Previous electrophysiological and lesion studies showed the involvement of the brainstem in both the vocal motor output pathway [62] as well as in the encoding of vocal sounds by the ascending auditory system [63]. Recent investigations focused on the role that cortical or cortico-thalamic systems may be playing in self-monitoring by interfacing between these two systems. Neural recordings in vocalizing

marmoset monkeys revealed a prominent vocalization-induced suppression of most auditory cortex neurons [64,65]. This suppression often begins prior to the onset of vocal production, suggesting a sensory-motor rather than a purely sensory origin. Despite this suppression, auditory cortical neurons remain sensitive to feedback during vocalization, including experimentally induced changes in vocal frequency or masking [66**,67]. Such neurons are actually more sensitive than predicted from purely sensory responses, and are modulated by Lombard effect-like vocal compensations [67] (Figure 3A,B), making them a possible node in the neural network for auditory-vocal self-monitoring.

The theoretical work on motor control and applied to understanding human speech self-monitoring and feedback vocal control [51] may have parallel implications for non-human primate vocalizations. Under such models, self-monitoring requires an animal to both know what they actually hear of their vocalization (i.e. feedback), and also what they expected to hear. Such predictions are termed efference copy or corollary discharge signals [68–70], and may be the origin of vocalization-induced suppression of the auditory cortex. This mechanism may also facilitate vocal turn-taking; that is, being able to distinguish one's own vocalization from that of another conspecific. For example, the marmoset monkey vocal turn-taking described above requires self-monitoring [59*]. Prefrontal and premotor cortical areas in primates are connected with temporal lobe auditory areas [71,72] and are involved in both natural [73*,74*] and trained vocal production [75,76]. Such regions are also implicated in the encoding of vocal sounds during passive listening, making them another possible site involved in the audio-motor interface circuit [30]. Additionally, activity in frontal cortex is capable of modulating activity in auditory cortex in mice [77], but whether or not this frontal-temporal pathway exists in primates and is responsible for auditory suppression and self-monitoring remain open questions.

Many sensory cues are generated during the production of vocalizations [78]. Self-monitoring for vocal communication is not limited to the auditory modality--somatosensory-proprioceptive self-monitoring may also be taking place. This pathway likely plays a greater role in the modification-via-articulation change in vocal acoustics, and perhaps to a lesser extent in voicing. While feedback control of human speech has been found to depend on somatomotor self-monitoring [79,80], the role of articulation in non-human primates is less well understood [81]. Macaques, in particular, exhibit prototypical facial movements associated with specific types of vocalizations [82]. Both the pre-frontal mirror neuron system [83] and anterior cingulate cortex [84] exhibit neural activities that reflect an animal's current facial gestures, making them two possible sites for non-auditory vocal self-monitoring. Further work will be needed to elucidate the role of these pathways in vocalization-associated somatomotor control and in the integration of multiple sensory pathways for feedback vocal control.

Acknowledgments

References and Recommended Reading

*of special interest

**of outstanding interest

1. Ghazanfar AA, Santos LR. Primate brains in the wild: The sensory bases for social interactions. Nature Reviews Neuroscience. 2004; 5:603–616.

2. Gustison ML, le Roux A, Bergman TJ. Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. Philosophical Transactions of the Royal Society B: Biological Sciences. 2012; 367:1847–1859.

3. McComb K, Semple S. Coevolution of vocal communication and sociality in primates. Biology Letters. 2005; 1:381–385. [PubMed: 17148212]

4. Jürgens U, Maurus M, Ploog D, Winter P. Vocalization in the squirrel monkey (Saimiri sciureus) elicited by brain stimulation. Exp Brain Res. 1967; 4:114–117. [PubMed: 4970751]

5. Myers RE. Comparative neurology of vocalization and speech: proof of a dichotomy. Ann N Y Acad Sci. 1976; 280:745–760. [PubMed: 827963]

6. Jürgens U. The neural control of vocalization in mammals: a review. J Voice. 2009; 23:1–10. [PubMed: 18207362]

7. Davila Ross M, Owren MJ, Zimmerman E. Reconstructing the evolution of laughter in great apes and humans. Current Biology. 2009; 19:1106–1111. [PubMed: 19500987]

8. Rauschecker JP, Tian B, Hauser M. Processing of Complex Sounds in the Macaque Nonprimary Auditory-Cortex. Science. 1995; 268:111–114. [PubMed: 7701330]

9. Recanzone GH. Representation of Con-Specific Vocalizations in the Core and Belt Areas of the Auditory Cortex in the Alert Macaque Monkey. Journal of Neuroscience. 2008; 28:13184–13193. [PubMed: 19052209]

10. Wang XQ, Merzenich MM, Beitel R, Schreiner CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. Journal of Neurophysiology. 1995; 74:2685–2706. [PubMed: 8747224]

11. Petkov CI, Logothetis NK, Obleser J. Where are the human speech and voice regions and do other animals have anything like them? The Neuroscientist. 2009; 15:419–429. [PubMed: 19516047]

12. Scott SK, Johnsrude IS. The neuroanatomical and functional organization of speech perception. Trends in Neurosciences. 2003; 26:100–107. [PubMed: 12536133]

13. Kikuchi Y, Horwitz B, Mishkin M. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. Journal Of Neuroscience. 2010; 30:13021–13030. [PubMed: 20881120]

*14. Fukushima M, Saunders RC, Leopold DA, Mishkin M, Averbeck BB. Differential coding of conspecific vocalizations in the ventral auditory cortical stream. Journal Of Neuroscience. 2014; 34:4665–4676. This study used chronic, high density electrode arrays to characaterize gradients of vocalization selectivity in the macaque monkey auditory cortex. It also demonstrated a putative role for the theta rhythm in vocal encoding. [PubMed: 24672012]

15. Ghazanfar, AA.; Poeppel, D. The neurophysiology and evolution of the speech rhythm. In: Gazzaniga, MS., editor. The cognitive neurosciences V. The MIT Press; 2014. vol In press

16. Ghazanfar AA, Takahashi DY. Facial expressions and the evolution of the speech rhythm. Journal of Cognitive Neuroscience. 2014 In press.

17. Belin P, Fecteau S, Bedard C. Thinking the voice: neural correlates of voice perception. Trends In Cognitive Sciences. 2004; 8:129–135. [PubMed: 15301753]

18. Weiss DJ, Garibaldi BT, Hauser MD. The production and perception of long calls by cotton-top tamarins (Saguinus oedipus): acoustic analyses and playback experiments. J Comp Psychol. 2001; 115:258–271. [PubMed: 11594495]

19. Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK. A voice region in the monkey brain. Nat Neurosci. 2008 In press.

20. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature. 2000; 403:309–312. [PubMed: 10659849]

21. Ghazanfar AA. Language evolution: neural differences that make a difference. Nature Neuroscience. 2008; 11:382–384.

22. Fitch WT, Hauser MD. Vocal production in nonhuman primates - acoustics, physiology, and functional constraints on honest advertisement. American Journal of Primatology. 1995; 37:191–219.

23. Ghazanfar AA, Rendall D. Evolution of human vocal production. Curr Biol. 2008; 18:R457–R460. [PubMed: 18522811]

24. Fitch WT. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J Acoust Soc Am. 1997; 102:1213–1222. [PubMed: 9265764]

25. Ghazanfar AA, Turesson HK, Maier JX, van Dinther R, Patterson RD, Logothetis NK. Vocal tract resonances as indexical cues in rhesus monkeys. Current Biology. 2007; 17:425–430. [PubMed: 17320389]

26. Rendall D, Owren MJ, Rodman PS. The role of vocal tract filtering in identity cueing in rhesus monkey (Macaca mulatta) vocalizations. Journal of the Acoustical Society of America. 1998; 103:602–614. [PubMed: 9440345]

27. Fitch WT, Fritz JB. Rhesus macaques spontaneously perceive formants in conspecific vocalizations. Journal of the Acoustical Society of America. 2006; 120:2132–2141. [PubMed: 17069311]

28. Perrodin C, Kayser C, Logothetis NK, Petkov CI. Voice cells in the primate temporal lobe. Curr Biol. 2011; 21:1408–1415. [PubMed: 21835625]

29. Averbeck BB, Romanski LM. Probabilistic encoding of vocalizations in macaque ventral lateral prefrontal cortex. Journal of Neuroscience. 2006; 26:11023–11033. [PubMed: 17065444]

30. Romanski LM, Averbeck BB, Diltz M. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophysiol. 2005; 93:734–747. [PubMed: 15371495]

31. Russ BE, Orr LE, Cohen YE. Prefrontal neurons predict choices during an auditory same-different task. Curr Biol. 2008; 18:1483–1488. [PubMed: 18818080]

32. Ghazanfar AA, Logothetis NK. Facial expressions linked to monkey calls. Nature. 2003; 423:937–938. [PubMed: 12827188]

33. Parr LA. Perceptual biases for multimodal cues in chimpanzee (Pan troglodytes) affect recognition. Animal Cognition. 2004; 7:171–178. [PubMed: 14997361]

34. Sliwa J, Duhamel JR, Pascalis O, Wirth S. Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. Proc Natl Acad Sci U S A. 2011; 108:1735–1740. [PubMed: 21220340]

35. Adachi I, Hampton RR. Rhesus monkeys see who they hear: spontaneous crossmodal memory for familiar conspecifics. PLoS One. 2011; 6:e23345. [PubMed: 21887244]

36. Habbershon HM, Ahmed SZ, Cohen YE. Rhesus macaques recognize unique multimodal face-voice relations of familiar individuals and not of unfamiliar ones. Brain, Behavior and Evolution. 2013; 81:219–225.

37. Chandrasekaran C, Lemus L, Trubanova A, Gondan M, Ghazanfar AA. Monkeys and humans share a common computation for face/voice integration. PLoS Comput Biol. 2011; 7:e1002165. [PubMed: 21998576]

38. Ghazanfar, AA.; Chandrasekaran, C. Non-human primate models of audiovisual communication. In: Stein, BE., editor. The new handbook of multisensory processes. The MIT Press; 2012. p. 407-420.

39. Calvert GA. Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cereb Cortex. 2001; 11:1110–1123. [PubMed: 11709482]

40. Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? Trends in Cognitive Sciences. 2006; 10:278–285. [PubMed: 16713325]

41. Ghazanfar, AA.; Chandrasekaran, C. Non-human primate models of audiovisual communication. In: Stein, BE., editor. The New Handbook of Multisensory Processes. MIT Press; 2012. p. 407-420.

42. Ghazanfar AA, Chandrasekaran C, Logothetis NK. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. Journal Of Neuroscience. 2008; 28:4457–4469. [PubMed: 18434524]

43. Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci. 2005; 25:5004–5012. [PubMed: 15901781]

44. Kayser C, Logothetis NK, Panzeri S. Visual enhancement of the information representation in auditory cortex. Curr Biol. 2010; 20:19–24. [PubMed: 20036538]

45. Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. Cereb Cortex. 2008; 18:1560–1574. [PubMed: 18180245]

**46. Perrodin C, Kayser C, Logothetis NK, Petkov CI. Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. Journal Of Neuroscience. 2014; 34:2524–2537. An elegant study comparing the auditory and audiovisual vocalization-elicited neuronal responses in both the voice area and the superior temporal sulcus. The two regions seem to process vocalizations in very different ways, with the former primarily dedicated to encoding indexical cues related to the caller and the latter primarily dedicated to integrating multisensory inputs. [PubMed: 24523543]

47. Sugihara T, Diltz MD, Averbeck BB, Romanski LM. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. J Neurosci. 2006; 26:11138–11147. [PubMed: 17065454]

48. Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci. 2005; 17:377–391. [PubMed: 15813999]

49. Chandrasekaran C, Ghazanfar AA. Different neural frequency bands integrate faces and voices differently in the superior temporal sulcus. Journal Of Neurophysiology. 2009; 101:773–788. [PubMed: 19036867]

**50. Chandrasekaran C, Lemus L, Ghazanfar AA. Dynamic faces speed-up the onset of auditory cortical spiking responses during vocal detection. Proc Natl Acad Sci U S A. 2013; 110:E4668–4677. The first study to directly link audiovisual vocal detection to neurophysiological activity. The results suggest that previous assumptions about the role of firing rate changes in mediating multisensory integration were incorrect. In auditory cortex, dynamic faces consistently speed up auditory cortical responses without systematically changing firing rate response profiles. [PubMed: 24218574]

51. Hickok G, Houde J, Rong F. Sensorimotor integration in speech processing: computational basis and neural organization. Neuron. 2011; 69:407–422. [PubMed: 21315253]

52. Janik VM, Slater PJ. The different roles of social learning in vocal communication. Anim Behav. 2000; 60:1–11. [PubMed: 10924198]

53. Brumm H, Voss K, Kollmer I, Todt D. Acoustic communication in noise: regulation of call characteristics in a New World monkey. J Exp Biol. 2004; 207:443–448. [PubMed: 14691092]

54. Egnor SE, Hauser MD. Noise-induced vocal modulation in cotton-top tamarins (Saguinus oedipus). Am J Primatol. 2006; 68:1183–1190. [PubMed: 17096420]

55. Lombard E. Le signe de l'elevation de la voix. Annales des Maladies de l'Oreille et du Larynx. 1911; 37:101–119.

56. Egnor SE, Wickelgren JG, Hauser MD. Tracking silence: adjusting vocal production to avoid acoustic interference. J Comp Physiol A. 2007; 193:477–483.

57. Roy S, Miller CT, Gottsch D, Wang X. Vocal control by the common marmoset in the presence of interfering noise. J Exp Biol. 2011; 214:3619–3629. [PubMed: 21993791]

58. Miller CT, Flusberg S, Hauser MD. Interruptibility of long call production in tamarins: implications for vocal control. J Exp Biol. 2003; 206:2629–2639. [PubMed: 12819269]

*59. Takahashi DY, Narayanan DZ, Ghazanfar AA. Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. Curr Biol. 2013; 23:2162–2168. This study shows that marmoset monkeys use vocal turn-taking rules similar to humans and that they engage in long sequences of vocal exchanges. The temporal properties of such sequences exhibit the characteristics of coupled oscillators: marmosets become phase-locked and entrained to each other. [PubMed: 24139740]

60. Lemasson A, Ouattara K, Petit EJ, Zuberbuhler K. Social learning of vocal structure in a nonhuman primate? BMC Evol Biol. 2011; 11:362. [PubMed: 22177339]

61. Elowson AM, Snowdon CT. Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. Anim Behav. 1994; 47:1267–1277.

62. Jürgens U. The neural control of vocalization in mammals: a review. J Voice. 2009; 23:1–10. [PubMed: 18207362]

63. Tammer R, Ehrenreich L, Jürgens U. Telemetrically recorded neuronal activity in the inferior colliculus and bordering tegmentum during vocal communication in squirrel monkeys (Saimiri sciureus). Behav Brain Res. 2004; 151:331–336. [PubMed: 15084450]

64. Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol. 2003; 89:2194–2207. [PubMed: 12612021]

65. Eliades SJ, Wang X. Comparison of auditory-vocal interactions across multiple types of vocalizations in marmoset auditory cortex. J Neurophysiol. 2013; 109:1638–1657. [PubMed: 23274315]

**66. Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature. 2008; 453:1102–1106. This study showed that neurons in the marmoset monkey auditory cortex were sensitive to auditory feedback during vocal production, and that changes in the feedback altered the coding properties of these neurons. Moreover, cortical suppression during vocalization actually increased the sensitivity of these neurons to vocal feedback. This heightened sensitivity to vocal feedback suggests that these neurons may have an important role in auditory self-monitoring. [PubMed: 18454135]

67. Eliades SJ, Wang X. Neural correlates of the lombard effect in primate auditory cortex. J Neurosci. 2012; 32:10737–10748. [PubMed: 22855821]

68. Crapse TB, Sommer MA. Corollary discharge circuits in the primate brain. Curr Opin Neurobiol. 2008; 18:552–557. [PubMed: 18848626]

69. von Holst E, Das HM. Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. Naturwissenschaften. 1950; 37:464–476.

70. Bridgeman B. A review of the role of efference copy in sensory and oculomotor control systems. Ann Biomed Eng. 1995; 23:409–422. [PubMed: 7486348]

71. Hackett TA, Stepniewska I, Kaas JH. Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Res. 1999; 817:45–58. [PubMed: 9889315]

72. Romanski LM, Bates JF, Goldman-Rakic PS. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol. 1999; 403:141–157. [PubMed: 9886040]

*73. Miller CT, Dimauro A, Pistorio A, Hendry S, Wang X. Vocalization Induced CFos Expression in Marmoset Cortex. Front Integr Neurosci. 2010; 4:128. [PubMed: 21179582]

*74. Simoes CS, Vianney PV, de Moura MM, Freire MA, Mello LE, Sameshima K, Araujo JF, Nicolelis MA, Mello CV, Ribeiro S. Activation of frontal neocortical areas by vocal production in marmosets. Front Integr Neurosci. 2010:4. The two studies referenced in 73 and 74 both used immediate-early gene expression to identify cortical areas related to natural, untrained vocal production in the marmoset areas. A number of areas seemed to be engaged, including those in the prefrontal cortex. This is at odds with long-held assumptions that neocortical involvement in vocal production was absent in monkeys. [PubMed: 20300470]

75. Coudé G, Ferrari PF, Rodà F, Maranesi M, Borelli E, Veroni V, Monti F, Rozzi S, Fogassi L. Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. PLoS One. 2011; 6:e26822. [PubMed: 22073201]

76. Hage SR, Nieder A. Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. Nat Commun. 2013; 4:2409. [PubMed: 24008252]

77. Nelson A, Schneider DM, Takatoh J, Sakurai K, Wang F, Mooney R. A circuit for motor cortical modulation of auditory cortical activity. J Neurosci. 2013; 33:14342–14353. [PubMed: 24005287]

78. Smotherman MS. Sensory feedback control of vocalizatons. Behav Brain Res. 2007; 182:315–326. [PubMed: 17449116]

79. Jones JA, Munhall KG. Learning to produce speech with an altered vocal tract: The role of auditory feedback. J Acoust Soc Am. 2003; 113:532–543. [PubMed: 12558289]

80. Tremblay S, Shiller DM, Ostry DJ. Somatosensory basis of speech production. Nature. 2003; 423:866–869. [PubMed: 12815431]

81. Riede T, Bronson E, Hatzikirou H, Zuberbuhler K. Vocal production mechanisms in a non-human primate: morphological data and a model. J Hum Evol. 2005; 48:85–96. [PubMed: 15656937]

82. Ghazanfar AA, Takahashi DY, Mathur N, Fitch WT. Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. Curr Biol. 2012; 22:1176–1182. [PubMed: 22658603]

83. Ferrari PF, Gallese V, Rizzolatti G, Fogassi L. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. European Journal of Neuroscience. 2003; 17:1703–1714. [PubMed: 12752388]

84. Livneh U, Resnik J, Shohat Y, Paz R. Self-monitoring of social facial expressions in the primate amygdala and cingulate cortex. Proc Natl Acad Sci U S A. 2012; 109:18956–18961. [PubMed: 23112157]

Highlights

- There is a caudal to rostral gradient of increasing vocalization-selectivity in auditory cortex

- Dynamic faces differentially influence voice processing in different auditory regions.

- In the context of vocal detection, dynamic faces speed up auditory cortex.

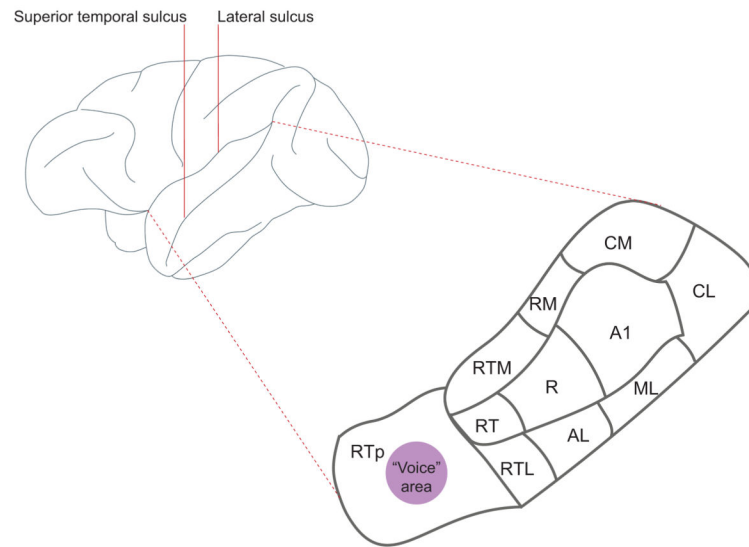- Monkeys exhibit a circuit for self-monitoring during vocal production.

Figure 1. Lateral view of the macaque monkey brain
Embedded within the lateral sulcus, on the lower bank, is the superior temporal plane. This region is magnified and the arrangement of auditory cortical areas are labeled: A1, primary auditory cortex; AL, anterolateral; CM, caudomedial; CL, caudolateral; ML, middle lateral; R, rostral; RM, rostromedial; RT, rostrotemporal; RTL, lateral rostrotemporal; RTM, medial rostrotemporal; RTp, rostrotemporal pole.
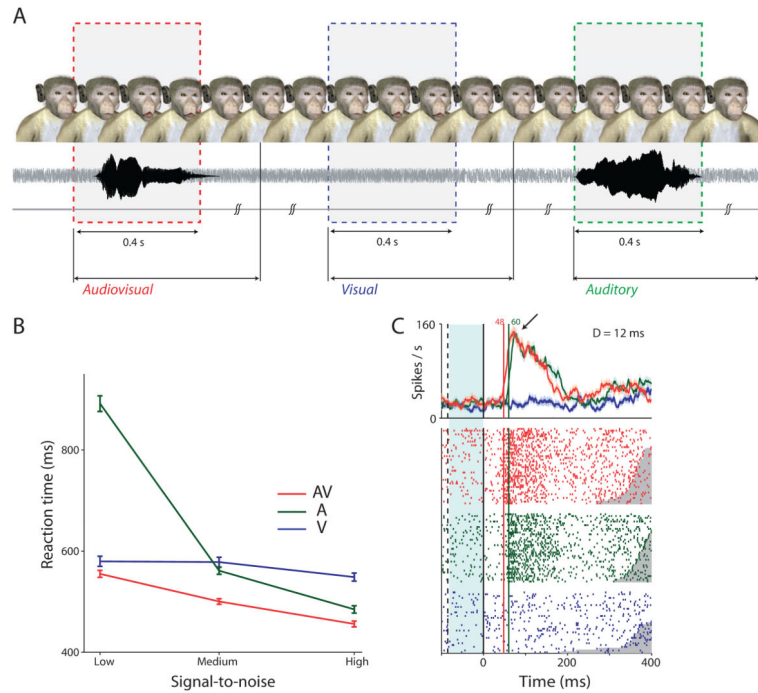
Figure 2. Dynamic faces speed up vocal detection and auditory cortical responses
(A) Task structure for monkeys. An avatar face was always on the screen. Visual, auditory and audiovisual stimuli were randomly presented with an inter stimulus interval of 1–3 seconds drawn from a uniform distribution. Responses within a 2 second window after stimulus onset were considered to be hits. Responses in the inter-stimulus interval are considered to be false alarms and led to timeouts. (B) Average RTs for the three different SNRs for the unisensory and multisensory conditions. Error bars denote standard error of mean across sessions. X-axes in denote SNR; Y-axis denotes RT in ms. (C) Peri-stimulus time histograms (PSTH) of a neuron in auditory cortex responding to AV, auditory and visual-only components of coo 1 at the highest SNR. X-axes depict time in ms. Y-axes the firing rate in spikes/s. Solid line is auditory-onset. Dashed line is visual onset. Blue shading denotes time period when only visual input was present. Green and red numbers indicate neuronal response latency; D=difference between response latencies. Adapted with permission from [50]
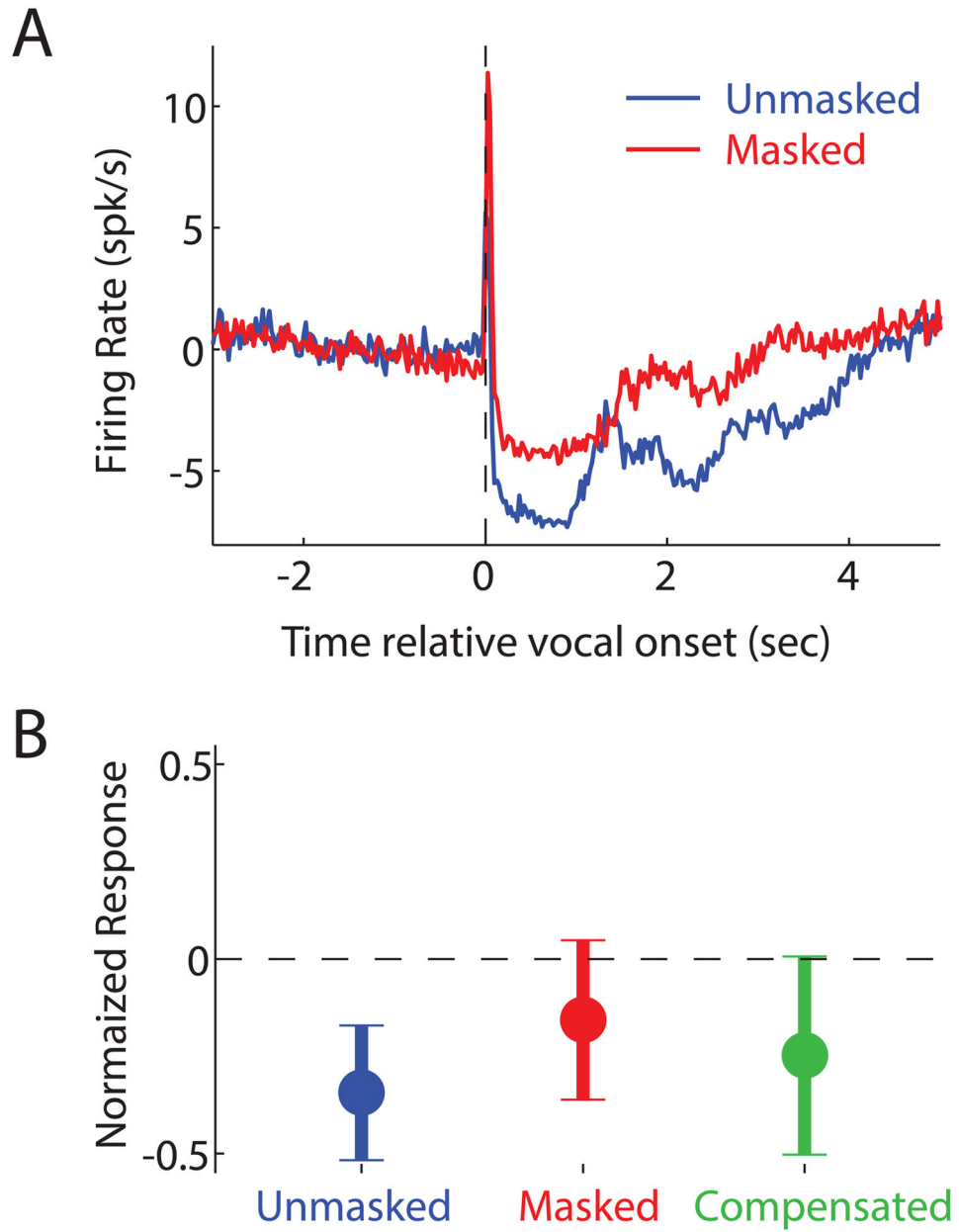
Figure 3. The auditory cortex monitors self-produced vocalizations
(A) A population PSTH shows responses of suppressed neurons in the auditory cortex during vocalization, and a reduction in suppression during masking. (B) Suppression was partially restored when marmosets compensated for masking by increasing their vocal intensity. Adapted with permission from [67]