



Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Auditory and audio-vocal responses of single neurons in the monkey ventral premotor cortex

Steffen R. Hage^{a, b, *}^a Neurobiology of Vocal Communication, Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Str. 25, 72076 Tübingen, Germany^b Animal Physiology Unit, Institute of Neurobiology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

ARTICLE INFO

Article history:

Received 14 December 2017

Received in revised form

15 March 2018

Accepted 19 March 2018

Available online xxx

Keywords:

Cognitive control

Premotor cortex

Sensory-motor integration

Vocal motor control

Vocalization

Rhesus macaque

ABSTRACT

Monkey vocalization is a complex behavioral pattern, which is flexibly used in audio-vocal communication. A recently proposed dual neural network model suggests that cognitive control might be involved in this behavior, originating from a frontal cortical network in the prefrontal cortex and mediated via projections from the rostral portion of the ventral premotor cortex (PMvr) and motor cortex to the primary vocal motor network in the brainstem. For the rapid adjustment of vocal output to external acoustic events, strong interconnections between vocal motor and auditory sites are needed, which are present at cortical and subcortical levels. However, the role of the PMvr in audio-vocal integration processes remains unclear. In the present study, single neurons in the PMvr were recorded in rhesus monkeys (*Macaca mulatta*) while volitionally producing vocalizations in a visual detection task or passively listening to monkey vocalizations. Ten percent of randomly selected neurons in the PMvr modulated their discharge rate in response to acoustic stimulation with species-specific calls. More than four-fifths of these auditory neurons showed an additional modulation of their discharge rates either before and/or during the monkeys' motor production of the vocalization. Based on these audio-vocal interactions, the PMvr might be well positioned to mediate higher order auditory processing with cognitive control of the vocal motor output to the primary vocal motor network. Such audio-vocal integration processes in the premotor cortex might constitute a precursor for the evolution of complex learned audio-vocal integration systems, ultimately giving rise to human speech.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Primate vocal communication systems are by far inferior to human speech in scope and flexibility (Balter, 2010; Hammerschmidt and Fischer, 2008). Monkey vocalizations were consequently, and quite reasonably, assumed to be largely innate and highly affective for decades (Egnor and Hauser, 2004; Jürgens, 2002). However, recent studies have shown that monkeys have distinct learning mechanisms during vocal development and vocal flexibility allowing them to cognitively control when (Coudé et al., 2011; Hage et al., 2013; Roy et al., 2016), where (Choi et al., 2015), and what to vocalize (Fischer and Price, 2017; Hage et al., 2013;

Seyfarth et al., 1980). Neural correlates for cognitive vocal behaviors have recently been identified within the ventrolateral prefrontal cortex (VLPFC), where single neurons have been reported to specifically predict the preparation of volitional vocal output (Gavrilov et al., 2017; Hage and Nieder, 2013). In addition, the VLPFC contains auditory and audio-vocal neurons, i.e., auditory neurons with an additional modulation of their discharge rates either before and/or during a monkey's motor production of vocalization (Hage and Nieder, 2015; Plakke et al., 2013). These results suggest that the VLPFC underlies the apex of complex audio-vocal integration processes by combining higher order auditory processing (Romanski and Averbeck, 2009) with cognitive control of vocal motor output (Hage and Nieder, 2016).

Recently, two reviews proposed similar dual neural network models that posit a cortical network originating in the VLPFC that is capable of cognitively controlling vocal output via a primary vocal motor network that is predominantly situated in subcortical structures (Hage and Nieder, 2016; Loh et al., 2017). According to

* Neurobiology of Vocal Communication, Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Str. 25, 72076 Tübingen, Germany.

E-mail address: steffen.hage@uni-tuebingen.de.

<https://doi.org/10.1016/j.heares.2018.03.019>

0378-5955/© 2018 Elsevier B.V. All rights reserved.

these models and recent neurophysiological and neuroanatomical results, the VLPFC might take control over the primary vocal motor network either via the anterior cingulate cortex, the apex structure of the latter network, or alternatively via the PMvr. On the one side, PMvr has direct connections to the reticular formation of the brainstem and medulla [the laryngeal PMvr (Jürgens and Ehrenreich, 2007; Simonyan, 2014);], where the vocal pattern generating network and cranial motoneuron pools involved in phonation are located and, on the other side, PMvr is controlling orofacial movements via the primary motor cortex (Coudé et al., 2011; Petrides et al., 2005). Since electrical microstimulation in the VLPFC as well as in the PMvr elicits orofacial and laryngeal responses (Coudé et al., 2011; Hast et al., 1974; Petrides et al., 2005; Simonyan and Jürgens, 2002, 2003) and vocalization-correlated neuronal activity can be recorded in the VLPFC as well as in the PMvr (Coudé et al., 2011; Hage and Nieder, 2013), the pathway leading from the VLPFC via the PMvr to the corticobulbar tract to control phonatory motor neurons seems more likely (Hage and Nieder, 2016).

The PMvr receives projections from auditory structures via the VLPFC (Simonyan and Jürgens, 2002; Yeterian et al., 2012) and deep layers of the superior temporal sulcus (Simonyan and Jürgens, 2005). Single-unit studies revealed vocal motor activity (Coudé et al., 2011; Gavrilov et al., 2017; Hage and Nieder, 2013) and an imaging study observed auditory activation of the PMvr in response to species-specific vocalizations (Gil-da-Costa et al., 2006). Consequently, like the VLPFC, the PMvr has been proposed to be a privileged site for sensory-motor integration (Gerbella et al., 2011). However, it remains unclear whether, and if yes, how audio-vocal integration processes are encoded within the PMvr at the single cell level.

2. Methods

2.1. Subjects

Two five-year-old male rhesus monkeys (*Macaca mulatta*) weighing 4.2 kg and 4.5 kg were used for this study. During the experiments, the monkeys worked under a controlled water intake protocol. All surgical procedures were performed in aseptic conditions under general anesthesia. All procedures were in accordance with the guidelines for animal experimentation and authorized by the national authorities (Regierungspräsidium Tübingen, Germany). All neurons analyzed in this study are a fraction of the cells that were recorded in a previous study (Hage and Nieder, 2013).

2.2. Experimental design

Single-cell recordings were conducted in monkeys trained to alternately produce calls and listen to species-specific vocalizations. The behavioral protocol required the monkeys to vocalize in response to an arbitrary visual cue in a go/nogo detection task as described earlier (Gavrilov et al., 2017; Hage et al., 2016, 2013, Hage and Nieder, 2013, 2015). Briefly, the monkey started a trial by grabbing a bar ('ready'-response; see Fig. 1A). A visual waiting-signal ('pre-cue') appeared for a randomized period of 1–5 s (white square, diameter: 0.5 deg of visual angle) during which the monkey was not allowed to vocalize. In 80% of trials, the 'pre-cue'-signal was followed by a colored visual 'go'-signal (red or blue square with equal probability; diameter: 0.5 deg of visual angle) lasting 3 s during which the monkey had to emit a vocalization to receive a liquid reward. To control for random calling behavior, the 'pre-cue' remained unchanged for 3 s in the remaining 20% of the trials and the monkey had to continue withholding vocal output

('catch'-trials).

At the end of each successful 'catch'-trial, i.e., when the monkey did not vocalize, the 'pre-cue'-stimulus remained unchanged for 1 s and the monkey was acoustically stimulated with the type of call it uttered during 'go' trials as described earlier (Hage and Nieder, 2015) (see also below for further details).

One session was recorded per individual per day. Monkeys were head-fixed during the experiment maintaining a constant distance of 5 cm between the monkey's head and the microphone. Eye movements were monitored via an IR-eye tracking system (ISCAN, Woburn, MA, USA), sampled at 1 kHz, and stored using a Plexon system for subsequent analysis. After initial vocal reinforcement training, the monkeys were successfully trained to perform the go/nogo detection task in 5–9 months.

2.3. Behavioral data acquisition

Stimulus presentation and behavioral monitoring was automated on PCs running the CORTEX program (NIH) and recorded using a multi-acquisition system (Plexon Inc., Dallas, TX) as described earlier (Hage and Nieder, 2013, 2015). Briefly, vocalizations were recorded synchronously with the neuronal data by the same system with a sampling rate of 40 kHz via an A/D converter for post-hoc analysis. Vocalizations were detected automatically by a custom-written MATLAB program (MathWorks, Natick, MA) that calculated several temporal and spectral acoustic parameters online and ran on another PC, which monitored the vocal behavior in real-time. All recordings were performed in a double-walled soundproof booth (IAC Acoustics, Niederkrüchten, Germany).

Successful vocalizations during 'go'-trials were defined as 'hits' and calls during 'catch'-trials as 'false alarms' according to the go/nogo detection paradigm. To test whether the monkey was capable of performing the detection task successfully, I computed d'-sensitivity-values derived from signal detection theory (Green and Swets, 1966) by subtracting z-scores (normal deviates) of median 'hit'-rates from z-scores of median 'false alarm'-rates. The detection threshold for d'-values was set to 1.8.

2.4. Auditory stimuli

For auditory stimulation, I used the call type identical to that uttered by the monkeys during the behavioral protocol. The main focus of the present study was to directly compare the activity of single neurons in response to specific self-produced and perceived identical vocalization. Because our behavioral protocol required a sparse presentation of sufficiently often repeated acoustic stimuli, I played back the vocal stimulus only during the catch trial, when the monkeys were not preparing for vocal output. Monkey T produced 'coo' vocalizations and was therefore acoustically stimulated with a high-quality recording of its own 'coo' call (753 ms duration). Monkey C produced 'grunt' vocalizations and was therefore played back a recording of its own 'grunt' call (140 ms duration) (see Fig. 1B). I used one vocalization exemplar as an acoustic stimulus for each. The vocalizations were stored as WAV files (sample rate 44.1 kHz), amplified (Yamaha amplifier A-520), and played back using one broadband speaker (Visaton), which was positioned 55 cm centered in front and 45° above the animal's head. The system was calibrated using a measuring amplifier (Brüel & Kjær, 2606 with condenser microphone 4135 and preamplifier 2633) to ensure a flat response of sound presentation (± 5 dB) between 0.1 kHz and 18 kHz. Vocal stimuli were presented at intensities of 80 dB SPL for the coo and 75 dB SPL for the grunt call.

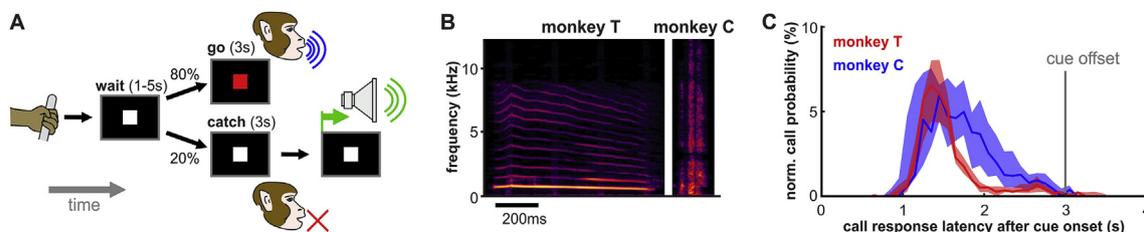


Fig. 1. Behavioral and auditory stimulation protocol, and behavioral performance. A Monkeys were trained to vocalize in response to a colored cue and were stimulated acoustically at the end of catch trials (20% of all trials) that required the monkey to withhold vocalization. B Spectrograms of the 'coo' and 'grunt' vocalizations that were used as acoustic stimuli. Spectral intensity is represented by different shades of color with black as the lowest and yellow as the highest intensity. C Call probability of both monkeys during go trials normalized for 31 (monkey T) and 27 recording sessions (monkey C); bin width 100 ms, shaded areas indicate 1st and 3rd quartiles.

2.5. Neurophysiological recordings

Extracellular single-unit activity was recorded with arrays of two to six glass-coated tungsten microelectrodes of 1 M Ω impedance (Alpha Omega, Alpharetta GA) within the left PMvr. Microelectrode arrays were inserted each recording day using a grid with 1-mm spacing that was mounted inside the recording chamber. Several marks on the grid and chamber ensured that the grid was positioned identically on each day. Neurons were randomly selected; no attempts were made to preselect units for specific discharge properties. Signal acquisition, amplification, filtering, and offline spike-sorting was accomplished using the Plexon system. Data analysis was carried out using MATLAB (MathWorks, Natick, MA). I analyzed all well-isolated neurons with mean discharge rates above 1 Hz that were recorded for at least seven hit trials, three miss trials, and seven auditory stimulation trials. Neurons showing eye-movement- and fixation-correlated activity were excluded from data analyses as described earlier (Gavrilov et al., 2017; Hage and Nieder, 2013).

Recording sites, i.e., recording wells and craniotomies, were localized using stereotaxic reconstructions from the individuals' magnetic resonance (MR) images by combining stereotaxic coordinates measured during surgery and the individuals' MR images taken prior to implantation. The identical stereotaxic frame was used for MR imaging and all surgeries (MRI stereotaxic frame 9-YST-35-P-O, all plastic materials, ear tips style: A, Eye/ear bars: offset, guide rails: 19 mm, Crist Instrument Company, Inc., Hagerstown, USA). Recordings from the PMvr were taken with the recording well and craniotomy centered on the inferior arcuate sulcus. The recording wells were implanted parallel to the cortical surface in both monkeys, i.e., the electrodes entered perpendicular to the cortical surface. Recordings of the cortical surface were all centered around the PMvr, posterior to the lower arcuate sulcus. This area contains neurons controlling orofacial movements via the primary motor cortex (Coudé et al., 2011; Petrides et al., 2005) as well as the laryngeal PMvr region that has direction connections to the lateral reticular formation of the brainstem via the corticobulbar tract (Jürgens and Ehrenreich, 2007; Simonyan, 2014).

2.6. Auditory neurons

Monkey T was trained to produce coo vocalizations and monkey C to utter grunt calls. Therefore, I acoustically stimulated monkey T with a coo vocalization and monkey C with a grunt vocalization, which gave us the unique ability to investigate both the neuron's activity in response to the auditory stimuli and its activity prior to the production of the equivalent vocalization. I determined a significant auditory response by comparing the neuronal activity during auditory stimulation with the neuron's baseline activity measured in a 500 ms window prior to stimulus onset.

In monkey T, which was stimulated with a 753 ms-coo vocalization, auditory responses were determined within two response periods: an early 500-ms interval starting 50 ms after stimulus onset and a subsequent late 500-ms interval. Auditory responses in monkey C, which was stimulated with a 140-ms grunt vocalization, were identified within a single 500-ms interval starting 50 ms after stimulus onset. The second analysis window in monkey T and the analysis window in monkey C, respectively, reach out beyond the offset of the used stimuli, which ensured the detection of potential auditory offset responses. Neurons with significant differences in firing rate in at least one of the analysis windows compared to that of the baseline were determined as auditory neurons (paired *t*-test, $p < 0.05$).

In addition, I detected phasic onset responses by dividing the time after stimulus onset into two intervals of equal duration (0–200 ms and 200–400 ms) and comparing the neuronal activity within these windows with the neuron's baseline activity within a 200-ms window prior to stimulus onset. Responses were defined as 'phasic' if the statistical analysis revealed significant differences within these three windows (Kruskal-Wallis-test, $p < 0.05$) and the neuronal activity in the first window after stimulus onset was higher than in the other two windows.

2.7. Pre-vocal and peri-vocal activity in auditory neurons

I further investigated the auditory neurons *post hoc*, examining their pre-vocal and peri-vocal activity prior to and during self-produced vocalizations. I tested for pre-vocal activity by comparing the neuronal activity in a 500-ms window just before vocal onset with baseline activity (500-ms analysis window prior to 'go'-stimulus onset; paired *t*-test, $p < 0.05$). Peri-vocal activity was evaluated by comparing the neuronal activity in a 500-ms window right after vocal onset with baseline activity (500-ms window prior to 'go'-stimulus onset; paired *t*-test, $p < 0.05$). Neurons with significant rate differences were determined as pre-vocal and peri-vocal neurons, respectively.

2.8. Modulation index

Neurons that showed statistically significant changes in neuronal discharge rates, either in response to auditory stimulation or in pre-vocal or peri-vocal activity, were further investigated to evaluate the direction of the changes relative to baseline. Similar to earlier studies (Eliades and Wang, 2013, 2008, 2003), I calculated a normalized measure, the modulation index, for auditory responses, pre-vocal responses, and peri-vocal responses as $(R_{stimulus} - R_{baseline}) / (R_{stimulus} + R_{baseline})$. Here, a modulation index of 0 indicates identical discharge rates during baseline and stimulus epoch, i.e., auditory response, pre-vocal, or peri-vocal activity. A modulation index of -1 indicates complete suppression during the stimulus

epoch, whereas a value of 1 indicates a neuron with a very low spontaneous rate and/or a very strong response during the stimulus epoch. To get an overview of significant pre- and peri-vocal activity of auditory neurons in the PMvr, I combined pre-vocal and peri-vocal activity by setting one premotor/motor modulation index per neuron. For neurons showing a significant modulation in both phases, prior to and during vocal output, the strongest modulation index was used.

2.9. Population analysis and normalization

Normalized activity was calculated by subtracting the mean neuronal baseline activity from the neuronal responses and dividing the outcome by the standard deviation (SD) of the baseline activity. Spike density histograms for single neurons were smoothed using a Gaussian kernel (bin width 100 ms; step size 1 ms) for illustrative purposes only. For baseline normalization, I analyzed auditory responses and vocalization-correlated activity relative to baseline activation in a preceding trial phase. This way, I could measure changes in discharge rates directly related to auditory stimuli and vocal output, respectively.

3. Results

The activity of single neurons during the playback of vocalizations (auditory condition), motor preparation of vocalizations (pre-vocal condition), and production of vocalizations (peri-vocal condition) was recorded in two monkeys trained to vocalize on command in a go/nogo task. (see Fig. 1A for experimental design) (Gavrilov et al., 2017; Hage and Nieder, 2013, 2015). I trained one monkey (monkey T) to utter 'coo' vocalizations, the second monkey (monkey C) was taught to emit 'grunts' (see Fig. 1B). To have identical vocalizations during listening and production, monkey T was therefore acoustically stimulated with a 'coo' vocalization, whereas 'grunts' were played back to monkey C. Neuronal recordings from the PMvr were performed during 15 daily sessions for monkey T and 17 sessions for monkey C. On average, each recording session consisted of 165 ± 11 (monkey T) and 157 ± 8

(monkey C) vocalizations per session. Mean 'hit rates' during go-trials were $29.4 \pm 2.5\%$ (monkey T) and $41.6 \pm 2.0\%$ (monkey C) and 'false alarm rates' during catch-trials $0.08 \pm 0.02\%$ (monkey T) and $0\% \pm 0\%$ (monkey C), respectively, resulting in values of the sensitivity index d' that were above the threshold of 1.8 (mean d' : 2.9 ± 0.1 in monkey T; 4.0 ± 0.1 in monkey C) in all sessions for both monkeys. This showed that the monkeys produced calls very reliably and almost exclusively in response to the visual go-cues. The median call latencies were similar in both monkeys (1.48 s in monkey T; 1.70 s in monkey C; see Fig. 1C).

While the monkeys performed the task, I recorded 233 neurons from the left PMvr; see Fig. 2A) of these two male Rhesus monkeys (*Macaca mulatta*). To avoid potential confounding effects caused by eye movements, I excluded neurons showing eye movement- and eye fixation-related activity from the data set with *post hoc* tests because the monkeys were not required to maintain fixation during the task. Nine percent of the neurons (20 out of 233 neurons) showed saccadic eye movement-related activity, and 5% of neurons (12/233) exhibited eye fixation correlated activity. After exclusion of these neurons, a total of 201 neurons were used for further analyses.

3.1. Auditory response properties

First, I examined the activity of PMvr neurons in response to external acoustic stimulation with vocalizations. Ten percent of the recorded neurons (20/201) showed significant modulation in response to external acoustic stimulation (paired *t*-test and Kruskal Wallis-test, respectively, $p < 0.05$; Fig. 2A). I observed auditory responses with increasing discharge rates as well as suppressed neuronal activity after stimulus onset (Fig. 2B).

To quantify the effect of auditory stimulation on the neurons' discharge rate, I calculated the auditory modulation index for every single responsive neuron (Fig. 3A). Overall, half of the neurons (10/20) showed positive modulation indices, indicating increased neuronal firing in response to auditory stimulation, while the other responsive neurons were characterized by negative modulation indices indicating that the neurons' firing rates were suppressed by

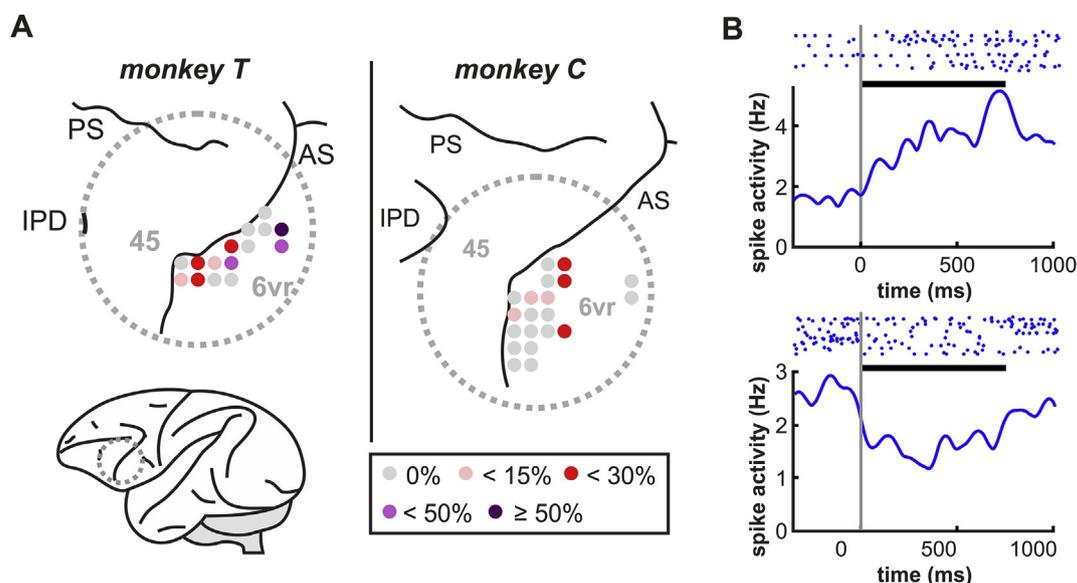


Fig. 2. Recording area and examples for acoustic responses. A Lateral view of the left hemisphere indicating the circular recording area encompassing the PMvr. B Example neurons showing increased and suppressed responses to monkey vocalizations in monkey T. In the top panels, the neuronal responses are plotted as dot-raster histograms (each dot represents an action potential). Bottom panels show spike density functions (activity averaged over all trials and smoothed by a 100 ms Gaussian kernel). Black horizontal bars between dot-raster histograms indicate the duration of the 'coo' vocalization used as acoustic stimulus for monkey T.

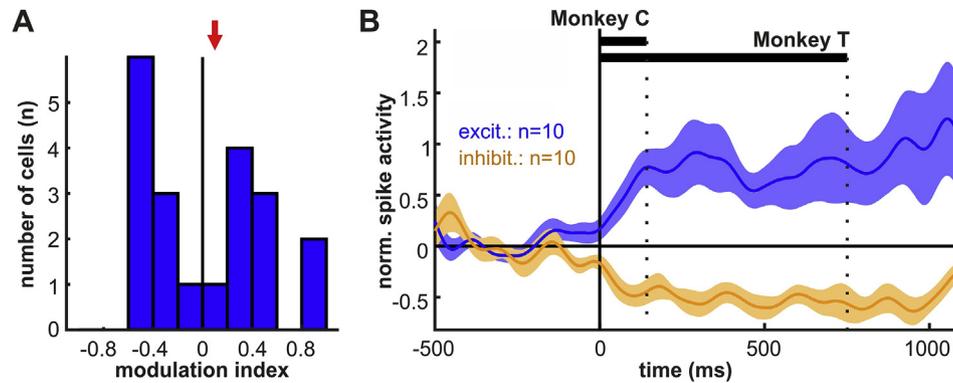


Fig. 3. Modulation indices and population responses for excited and inhibited auditory activity. A Distribution of the modulation indices indicating the magnitude of excitation and suppression during auditory stimulation (red arrow: median modulation index). B Averaged and normalized activity (population responses) for auditory neurons showing increased and suppressed responses in PMvr; black horizontal bars indicate the duration of the 'coo' and 'grunt' vocalization used as acoustic stimuli for monkey T and C, respectively.

auditory stimulation. Auditory neurons showed a median auditory modulation index of 0.03. I separated the auditory neurons into two populations regarding modulation response for each monkey. No significant differences were found between the absolute (rectified) modulation indices exhibited during inhibited and excited auditory responses ($p > 0.5$, $n = 20$, Wilcoxon rank sum test). Fig. 3B shows the mean population responses of acoustically excited (modulation index > 0) and inhibited neurons (modulation index < 0) for each monkey.

3.2. Interaction of auditory and vocal activity

Next, I investigated the activity of the auditory neurons before and during conditioned vocalizations. I were not able to analyze changes of discharge rates during spontaneous calls due to the small number of auditory neurons recorded during spontaneous vocalizations. Fig. 4A shows an auditory neuron with excitatory responses to auditory stimulation. At the same time, this neuron shows a significant decrease in pre-vocal activity, i.e., prior to vocalization, that persisted during conditioned vocal output (suppressed peri-vocal activity). Forty percent (8/20) of the auditory neurons showed significantly modulated pre-vocal activity (paired t -test, $p < 0.05$). Peri-vocal discharge rates were modulated in 70% (14/20) of the auditory cells. Fig. 4B shows the population responses for auditory neurons showing significant excited (5 neurons) and suppressed (12 neurons) pre- and/or peri-vocal activity for cued vocal behavior.

Overall, 85% (17/20) of the auditory neurons showed a significant modulation of their discharge rates either prior to vocal onset (pre-vocal), during vocal output (peri-vocal), or in both situations (Fig. 4C). Specifically, a quarter of the neurons exhibiting changes in pre-vocal activity showed an additional modulation in firing rates during goal-directed vocalizations (25%, 5/20). About half of the auditory neurons (9/20) showed significant peri-vocal responses. Fifteen percent (3/20) of the auditory neurons showed only pre-vocal activity changes.

Finally, I calculated the modulation indices for the observed pre-/peri-vocal activities to quantify the premotor-/motor-related neuronal responses. Then, I compared the strengths of the neuronal modulation of the auditory neurons during the pre-/peri-vocal and auditory stimulation phases by plotting the respective values of the modulation indices of individual neurons against each other (Fig. 5A). The distribution of pre-/peri-vocal and corresponding auditory modulation indices in relation to the bisecting lines indicate that modulation rates observed prior to and during vocal

output were higher than during auditory stimulation. More than two-thirds of audio-vocal neurons (70.6%, 12/17) showed a negative modulation index indicating suppressed activity prior to and/or during vocal output. Five neurons (29.4%) exhibited increased pre- and/or peri-vocal neuronal activity. Notably, about half of the neurons showed contrary modulation indices between auditory stimulation and pre- and/or peri-vocal epochs (47.1%, 8/17). For quantification, I calculated the absolute pre-/peri-vocal and auditory modulation indices (Fig. 5B). Consequently, absolute pre-/peri-vocal modulation indices (median = 0.68) were significantly higher than those during auditory stimulation (median = 0.36, $p < 0.01$, $n = 17$, Wilcoxon sign rank test).

4. Discussion

The present study shows that a fraction of neurons in the PMvr exhibit responses to auditory stimulation with species-specific calls. More than four-fifths of these auditory neurons show a significant modulation of their discharge rates either before and/or during vocalization. These audio-vocal interactions suggest the PMvr as another candidate region for combining auditory processing mechanisms and vocal motor control in the primate cortex.

4.1. Auditory activity in monkey premotor cortex

Auditory neurons have been recorded in several areas of the monkey frontal lobe such as in areas BA 44, 45, and 12/47 of the VLPFC (Hage and Nieder, 2015; Romanski, 2004; Romanski and Goldman-Rakic, 2002). Our findings of auditory responses in the PMvr are in agreement with a previous imaging study that observed auditory activation within the PMvr in response to species-specific vocalizations (Gil-da-Costa et al., 2006). The fraction of auditory neurons recorded within the pool of randomly selected neurons was similar to that recorded in the VLPFC (10% PMvr vs. 12% VLPFC) (Hage and Nieder, 2015). However, Coudé et al. (2011) recently also recorded from the PMvr of vocalizing monkeys and did not find auditory responses in vocalization-correlated neurons. This discrepancy is most likely due to the small sample size of vocalization-correlated neurons in the latter study, which may have led to missing single auditory or audio-vocal neurons from the overall low fraction of auditory and audio-vocal neurons. The PMvr seems to most likely receive auditory information from the belt and para-belt regions of the auditory cortex via the VLPFC (Simonyan and Jürgens, 2002; Yeterian et al., 2012) and via the deep layers of the superior temporal sulcus (Simonyan and Jürgens,

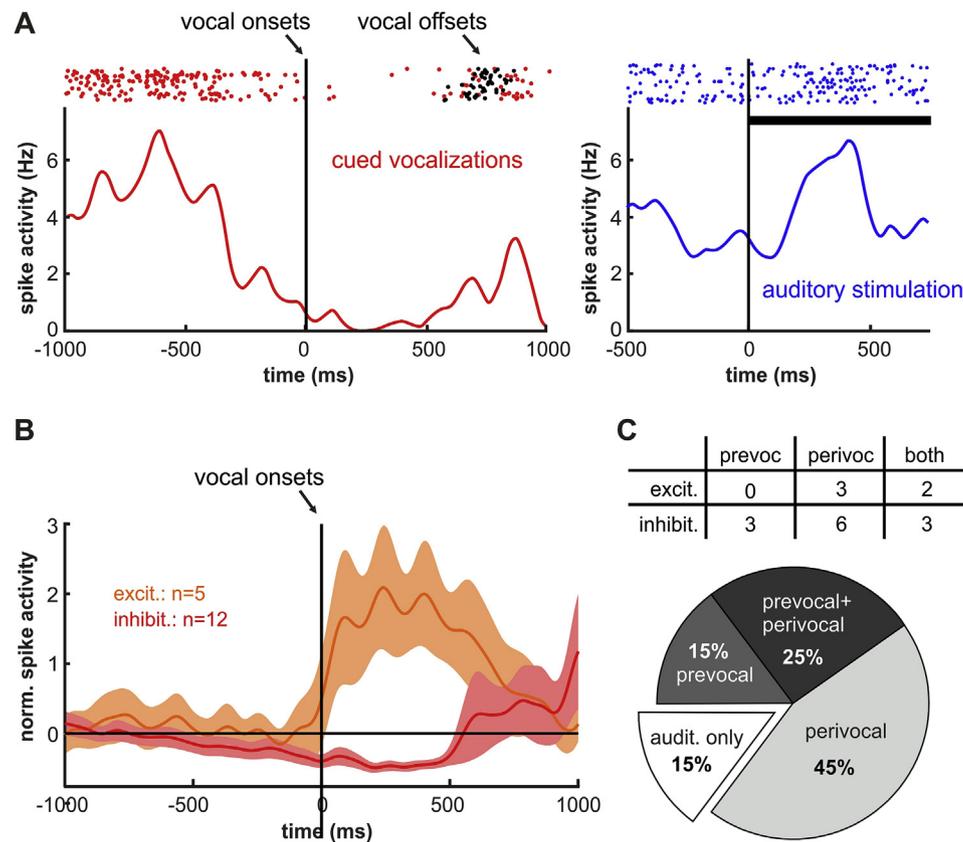


Fig. 4. Audio-vocal activity in PMvr neurons. An Example neuron showing excitatory response during auditory stimulation and suppressed neuronal activity starting prior to cued vocalizations and lasting throughout vocal output. Upper panels show raster plots; lower panels represent the corresponding spike density histograms averaged and smoothed with a Gaussian kernel (100 ms) for illustration. B Averaged and normalized population responses of auditory neurons subdivided into neurons showing additional significantly increased (orange curve) and inhibited responses (red curve) during self-produced cued vocalizations. C Proportions of auditory neurons that also showed pre-vocal, peri-vocal, or both types of activation.

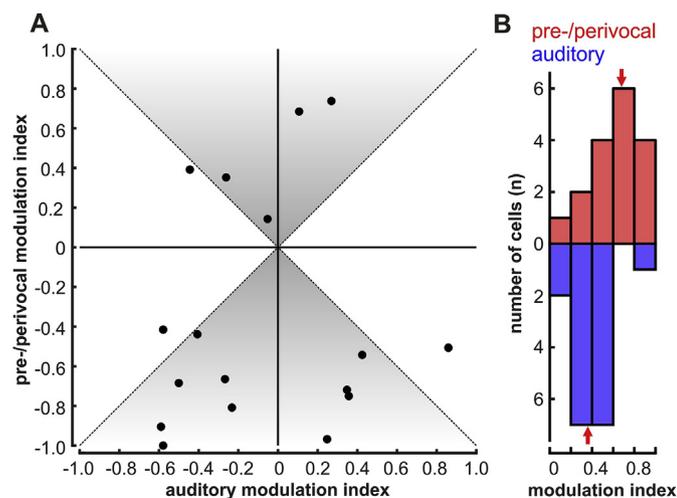


Fig. 5. Modulation indices of audio-vocal neurons. A Scatter plots of auditory neurons as a function of their modulation indices. Each dot represents one neuron with its corresponding modulation index in response to auditory stimulation as well as prior to (pre-vocal) or during cued vocalization (peri-vocal). Grey shaded area indicated regions with neurons showing a higher absolute modulation rate prior to or during vocal output compared to their respective modulation rate during auditory stimulation. B Distribution of the absolute modulation indices of auditory neurons exhibiting significant modulation prior to or during vocal output. Red arrows mark the median modulation indices.

2005). Neurons in the PMvr, as well as in the VLPFC, and belt and para-belt regions of the auditory cortex, show similar responses to auditory stimulation with species-specific vocalizations (Hage and Nieder, 2015; Romanski, 2004; Tian et al., 2001). In contrast to primary areas of the auditory cortex, single neurons that have been recorded in the belt and para-belt regions of the auditory cortex of macaque monkeys were activated by complex sounds, such as vocalizations, rather than by simple stimuli (Kikuchi et al., 2010; Rauschecker et al., 1997, 1995). Furthermore, the supra-temporal plane of macaque monkeys contains a vocalization area in which neurons preferred monkey calls over other auditory stimuli (Petkov et al., 2008) and sometimes even preferred individual voices (Perrodin et al., 2014, 2011) or one call type over others (Fukushima et al., 2014).

4.2. Audio-vocal interaction in PMvr

In recent studies, the PMvr has been suggested to be involved in connecting the cortical articulatory vocal motor network with the primary vocal motor network, which is predominantly situated within the brainstem (Hage and Nieder, 2016; Loh et al., 2017). The PMvr encompasses the cortical larynx area, which projects via the corticobulbar pathway to areas in the lower brainstem containing motoneurons and premotor neurons, as well as pattern-generating structures that are involved in vocal production (Simonyan, 2014; Simonyan and Jürgens, 2003). Furthermore, electrical stimulation within the PMvr in awake and anesthetized macaques elicited orofacial and laryngeal responses (Coudé et al., 2011; Hast et al.,

1974; Petrides et al., 2005; Simonyan and Jürgens, 2003, 2002). In addition, single neurons have recently been found within the PMvr that exhibited a change in activity prior to and during vocal output (Coudé et al., 2011; Hage and Nieder, 2013; Roy et al., 2016). These findings support the direct involvement of the PMvr in vocal motor processes.

Several studies have revealed that the PMvr is a central hub for sensory-motor integration processes involved in a variety of motor acts such as grasping, reaching or placing objects (Caggiano et al., 2016; Rizzolatti et al., 2002). However, it is virtually unknown how the PMvr is involved in audio-vocal integration processes. In the current study, such audio-vocal integration processes could be observed at the single neuron level. Most recorded auditory neurons showed changes in the neuronal modulation rate prior to and/or during vocal onset. Neurons with such audio-vocal response properties seem to be a general phenomenon of auditory-vocal interactions and have already been found at several levels in the primate brain, such as at lower brainstem (Hage et al., 2006), midbrain (Pieper and Jürgens, 2003), and auditory cortical levels (Eliades and Wang, 2013, 2008; 2003; Müller-Preuss and Ploog, 1981), as well as in the VLPFC (Hage and Nieder, 2015). Similar to the ventral premotor level, audio-vocal neurons in the lower brainstem, auditory cortex, and prefrontal cortex showed each combination of excited or suppressed activity in response to self-produced vocalizations and auditory stimulation (Eliades and Wang, 2013, 2003; Hage et al., 2006; Hage and Nieder, 2015). Similar to audio-vocal neurons in the auditory cortex and in contrast to the VLPFC, auditory neurons were predominantly suppressed around self-produced vocalizations (Eliades and Wang, 2008, 2003; Hage and Nieder, 2015). While a predominant suppression in the auditory cortex might be directly linked to vocal output to increase the dynamic range of the auditory system and maintain the hearing sensitivity to the external acoustic environment during vocal output (Eliades and Wang, 2003), in the PMvr it might be related to complex audio-vocal integration processes.

4.3. Potential role of audio-vocal neurons in PMvr

Neuronal modulation rates of single neurons were analyzed to investigate the possible functions of audio-vocal activity in the PMvr. Like in the VLPFC, the absolute modulation rates of most neurons prior to and/or during self-produced vocalizations were significantly higher than those during auditory stimulation (Hage and Nieder, 2015). These results indicate that in PMvr neurons, similar to the audio-vocal neurons of the VLPFC, pre-motor functions are dominant, i.e., that these neurons receive modulatory auditory input but seem to be primarily involved in motor control. However, while audio-vocal neurons in the VLPFC predominantly exhibited excitatory responses prior to and/or during vocal output, most of the PMvr neurons were suppressed before and during calling. Furthermore, about half of the audio-vocal neurons in the PMvr show inverse modulation indices during auditory and pre- and/or peri-vocal epochs. These results indicate that audio-vocal neurons in the PMvr are involved in complex audio-vocal integration processes rather than just responding to self-produced and externally perceived vocalizations in a similar way. Since the PMvr is both directly connected to the reticular formation of the brainstem and medulla (Jürgens and Ehrenreich, 2007; Simonyan, 2014), where the vocal pattern generating network and cranial motoneuron pools involved in phonation are located, as well as connected via the primary motor cortex to motoneuron pools being involved in orofacial movements (Hage and Nieder, 2016), it might be able to directly modulate ongoing vocalizations during external acoustic events. In accordance, recent work observed that external acoustic events are capable of rapidly modulating ongoing

vocalizations (Eliades and Tsunada, 2017; Pomberger et al., 2018). Future studies will now have to decipher, if and how, premotor neurons are involved in such rapid, complex audio-vocal integration processes.

These results help to elucidate the role of frontal brain structures, such as the PMvr, in audio-vocal integration processes. It has been a matter of debate, whether the activation of the motor cortex during auditory tasks in humans might just be an artifact of brain imaging studies or not (Zatorre et al., 1996). The results of this study provide evidence that such an auditory-motor interaction in the primate premotor cortex does indeed exist. Such circuits in the frontal lobe that integrate auditory information with premotor activity could have readily been adopted during the course of primate evolution to establish complex audio-vocal behaviors known to be crucial for proper speech and language communication in humans (Rauschecker and Scott, 2009).

Competing interests

The author has no competing financial interests.

Acknowledgements

I thank Andreas Nieder for valuable comments and fruitful discussion and John Holmes for proofreading. This work was supported by the Werner Reichardt Centre for Integrative Neuroscience (CIN) at the Eberhard Karls University of Tübingen (CIN is an Excellence Cluster funded by the Deutsche Forschungsgemeinschaft within the frame-work of the Excellence Initiative EXC 307) and the Deutsche Forschungsgemeinschaft Grant HA5400/3-1.).

References

- Balter, M., 2010. Evolution of language. Animal communication helps reveal roots of language. *Science* 328, 969–971. <https://doi.org/10.1126/science.328.5981.969>.
- Caggiano, V., Fleischer, F., Pomper, J.K., Giese, M.A., Thier, P., 2016. Mirror neurons in monkey premotor area F5 show tuning for critical features of visual causality perception. *Curr. Biol.* 26, 3077–3082. <https://doi.org/10.1016/j.cub.2016.10.007>.
- Choi, J.Y., Takahashi, D.Y., Ghazanfar, A.A., 2015. Cooperative vocal control in marmoset monkeys via vocal feedback. *J. Neurophysiol.* 114, 274–283. <https://doi.org/10.1152/jn.00228.2015>.
- Coudé, G., Ferrari, P.F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., Monti, F., Rozzi, S., Fogassi, L., 2011. Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One* 6 e26822. <https://doi.org/10.1371/journal.pone.0026822>.
- Egnor, S.E.R., Hauser, M.D., 2004. A paradox in the evolution of primate vocal learning. *Trends Neurosci.* 27, 649–654. <https://doi.org/10.1016/j.tins.2004.08.009>.
- Eliades, S.J., Tsunada, J., 2017. Auditory cortical activity predicts feedback-dependent vocal control in marmoset monkeys. In: *Neuroscience Meeting Planner. Society for Neuroscience, Washington, DC. Program No. 586.07. Online*.
- Eliades, S.J., Wang, X., 2013. Comparison of auditory-vocal interactions across multiple types of vocalizations in marmoset auditory cortex. *J. Neurophysiol.* 109, 1638–1657. <https://doi.org/10.1152/jn.00698.2012>.
- Eliades, S.J., Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106. <https://doi.org/10.1038/nature06910>.
- Eliades, S.J., Wang, X., 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 2194–2207. <https://doi.org/10.1152/jn.00698.2012>.
- Fischer, J., Price, T., 2017. Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* 82, 22–31. <https://doi.org/10.1016/j.neubiorev.2016.10.014>.
- Fukushima, M., Saunders, R.C., Leopold, D. a., Mishkin, M., Averbeck, B.B., 2014. Differential Coding of Conspecific Vocalizations in the Ventral Auditory Cortical Stream. *J. Neurosci.* 34, 4665–4676. <https://doi.org/10.1523/JNEUROSCI.3969-13.2014>.
- Gavrilov, N., Hage, S., Nieder, A., 2017. Functional specialization of the primate frontal lobe during cognitive control of vocalizations. *Cell Rep.* 21, 2393–2406. <https://doi.org/10.1016/j.celrep.2017.10.107>.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., Luppino, G., 2011. Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Struct. Funct.* 216, 43–65. <https://doi.org/10.1007/s00429-010-0293-6>.
- Gil-da-Costa, R., Martin, A., Lopes, M. a, Muñoz, M., Fritz, J.B., Braun, A.R., 2006.

- Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. *Nat. Neurosci.* 9, 1064–1070. <https://doi.org/10.1038/nn1741>.
- Green, D.M., Swets, J., 1966. *Signal Detection Theory and Psychophysics*. Wiley, New York.
- Hage, S.R., Gavrilov, N., Nieder, A., 2016. Developmental changes of cognitive vocal control in monkeys. *J. Exp. Biol.* 219, 1744–1749. <https://doi.org/10.1242/jeb.137653>.
- Hage, S.R., Gavrilov, N., Nieder, A., 2013. Cognitive control of distinct vocalizations in rhesus monkeys. *J. Cognit. Neurosci.* 25, 1692–1701. https://doi.org/10.1162/jocn_a.00428.
- Hage, S.R., Jürgens, U., Ehret, G., 2006. Audio-vocal interaction in the pontine brainstem during self-initiated vocalization in the squirrel monkey. *Eur. J. Neurosci.* 23, 3297–3308. <https://doi.org/10.1111/j.1460-9568.2006.04835.x>.
- Hage, S.R., Nieder, A., 2016. Dual neural network model for the evolution of speech and language. *Trends Neurosci.* 39, 813–829. <https://doi.org/10.1016/j.tins.2016.10.006>.
- Hage, S.R., Nieder, A., 2015. Audio-vocal interaction in single neurons of the monkey ventrolateral prefrontal cortex. *J. Neurosci.* 35, 7030–7040. <https://doi.org/10.1523/JNEUROSCI.2371-14.2015>.
- Hage, S.R., Nieder, A., 2013. Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* 4 (2409). <https://doi.org/10.1038/ncomms3409>.
- Hammerschmidt, K., Fischer, J., 2008. Constraints in primate vocal production. In: Oller, D.K., Griebel, U. (Eds.), *Evolution of Communicative Flexibility*. MIT Press, Cambridge, pp. 93–121.
- Hast, M.H., Fischer, J.M., Wetzel, A.B., Thompson, V.E., 1974. Cortical motor representation of the laryngeal muscles in *Macaca mulatta*. *Brain Res.* 73, 229–240. [https://doi.org/10.1016/0006-8993\(74\)91046-4](https://doi.org/10.1016/0006-8993(74)91046-4).
- Jürgens, U., 2002. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* 26, 235–258.
- Jürgens, U., Ehrenreich, L., 2007. The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Res.* 1148, 90–95. <https://doi.org/10.1016/j.brainres.2007.02.020>.
- Kikuchi, Y., Horwitz, B., Mishkin, M., 2010. Hierarchical auditory processing directed rostrally along the Monkey's supratemporal plane. *J. Neurosci.* 30, 13021–13030. <https://doi.org/10.1523/JNEUROSCI.2267-10.2010>.
- Loh, K.K., Petrides, M., Hopkins, W.D., Procyk, E., Amiez, C., 2017. Cognitive control of vocalizations in the primate ventrolateral-dorsomedial frontal (VLF-DMF) brain network. *Neurosci. Biobehav. Rev.* 82, 32–44. <https://doi.org/10.1016/j.neubiorev.2016.12.001>.
- Müller-Preuss, P., Ploog, D., 1981. Inhibition of auditory cortical neurons during phonation. *Brain Res.* 215, 61–76. [https://doi.org/10.1016/0006-8993\(81\)90491-1](https://doi.org/10.1016/0006-8993(81)90491-1).
- 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. *J. Neurosci.* 34, 2524–2537. <https://doi.org/10.1523/JNEUROSCI.2805-13.2014>.
- Perrodin, C., Kayser, C., Logothetis, N.K., Petkov, C.I., 2011. Voice cells in the primate temporal lobe. *Curr. Biol.* 21, 1408–1415. <https://doi.org/10.1016/j.cub.2011.07.028>.
- Petkov, C.I., Kayser, C., Studel, T., Whittingstall, K., Augath, M., Logothetis, N.K., 2008. A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374. <https://doi.org/10.1038/nn2043>.
- Petrides, M., Cadoret, G., Mackey, S., 2005. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* 435, 1235–1238. <https://doi.org/10.1038/nature03628>.
- Pieper, F., Jürgens, U., 2003. *N* Euronal Activity in the Inferior Colliculus and Bordering Structures during Vocalization in the Squirrel Monkey, vol. 979, pp. 153–164.
- Plakke, B., Diltz, M.D., Romanski, L.M., 2013. Coding of vocalizations by single neurons in ventrolateral prefrontal cortex. *Hear. Res.* 305, 135–143. <https://doi.org/10.1016/j.heares.2013.07.011>.
- Pomberger, T., Risueno-Segovia, C., Löschner, J., Hage, S.R., 2018. Precise motor control enables rapid flexibility in vocal behavior of marmoset monkeys. *Curr. Biol.* 28, 788–794. <https://doi.org/10.1016/j.cub.2018.01.070>.
- Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. <https://doi.org/10.1038/nn.2331>.
- Rauschecker, J.P., Tian, B., Hauser, M.D., 1995. Processing of complex sounds in nonprimary auditory cortex of the rhesus monkey. *Science (80-)* 268, 111–114.
- Rauschecker, J.P., Tian, B., Pons, T., Mishkin, M., 1997. Serial and parallel processing in rhesus monkey auditory cortex. *J. Comp. Neurol.* 382, 89–103. [https://doi.org/10.1002/\(SICI\)1096-9861\(19970526\)382:1<89::AID-CNE6>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1096-9861(19970526)382:1<89::AID-CNE6>3.0.CO;2-G).
- Rizzolatti, G., Fogassi, L., Gallese, V., 2002. Motor and cognitive functions of the ventral premotor cortex. *Curr. Opin. Neurobiol.* 12, 149–154. [https://doi.org/10.1016/S0959-4388\(02\)00308-2](https://doi.org/10.1016/S0959-4388(02)00308-2).
- Romanski, L.M., 2004. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J. Neurophysiol.* 93, 734–747. <https://doi.org/10.1152/jn.00675.2004>.
- Romanski, L.M., Averbeck, B.B., 2009. The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu. Rev. Neurosci.* 32, 315–346. <https://doi.org/10.1146/annurev.neuro.051508.135431>.
- Romanski, L.M., Goldman-Rakic, P., 2002. An auditory domain in primate prefrontal cortex. *Nat. Neurosci.* 5, 15–16. <https://doi.org/10.1038/nn781An>.
- Roy, S., Zhao, L., Wang, X., 2016. Distinct neural activities in premotor cortex during natural vocal behaviors in a new world primate, the common marmoset (*Callithrix jacchus*). *J. Neurosci.* 36, 12168–12179. <https://doi.org/10.1523/JNEUROSCI.1646-16.2016>.
- Seyfarth, R., Cheney, D., Marler, P., 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science (80-)* 210, 801–803. <https://doi.org/10.1126/science.7433999>.
- Simonyan, K., 2014. The laryngeal motor cortex: its organization and connectivity. *Curr. Opin. Neurobiol.* 28, 15–21. <https://doi.org/10.1016/j.conb.2014.05.006>.
- Simonyan, K., Jürgens, U., 2005. Afferent cortical connections of the motor cortical larynx area in the rhesus monkey. *Neuroscience* 130, 133–149. <https://doi.org/10.1016/j.neuroscience.2004.08.031>.
- Simonyan, K., Jürgens, U., 2003. Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Res.* 974, 43–59.
- Simonyan, K., Jürgens, U., 2002. Cortico-cortical projections of the motorcortical larynx area in the rhesus monkey. *Brain Res.* 949, 23–31. [https://doi.org/10.1016/S0531-5131\(03\)00895-1](https://doi.org/10.1016/S0531-5131(03)00895-1).
- Tian, B., Reser, D., Durham, A., Kustov, A., Rauschecker, J.P., 2001. Functional specialization in rhesus monkey auditory cortex. *Science (80-)* 292, 290–293. <https://doi.org/10.1126/science.1058911>.
- Veterian, 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. <https://doi.org/10.1016/j.cortex.2011.03.004>.
- Zatorre, R.J., Meyer, E., Gjedde, A., Evans, A.C., 1996. PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cerebr. Cortex* 6, 21–30. <https://doi.org/10.1093/cercor/6.1.21>.