Hearing Research xxx (2018) 1–8

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.

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/

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.

Please cite this article in press as: Hage, S.R., Auditory and audio-vocal responses of single neurons in the monkey ventral premotor cortex, Hearing Research (2018), https://doi.org/10.1016/j.heares.2018.03.019