BRAIN RESEARCH 1265 (2009) 80–92

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Differential representation of spectral and temporal information by primary auditory cortex neurons in awake cats: Relevance to auditory scene analysis

Masashi Sakai, Sohei Chimoto, Ling Qin, Yu Sato*

Department of Physiology, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Yamanashi 409-3898, Japan

ARTICLE INFO

Article history:
Accepted 29 January 2009
Available online 10 February 2009

Keywords:
Auditory cortex
Neuron
Temporal coding
Synchronization
Adaptation

ABSTRACT

We investigated how the primary auditory cortex (AI) neurons encode the two major requisites for auditory scene analysis, i.e., spectral and temporal information. Single-unit activities in awake cats AI were studied by presenting 0.5-s-long tone bursts and click trains. First of all, the neurons (n=92) were classified into 3 types based on the time-course of excitatory responses to tone bursts: 1) phasic cells (P-cells; 26%), giving only transient responses; 2) tonic cells (T-cells; 34%), giving sustained responses with little or no adaptation; and 3) phasic-tonic cells (PT-cells; 40%), giving sustained responses with some tendency of adaptation. Other tone–response variables differed among cell types. For example, P-cells showed the shortest latency and smallest spiking jitter while T-cells had the sharpest frequency tuning. PT-cells generally fell in the intermediate between the two extremes. Click trains also revealed between-neuron-type differences for the emergent probability of excitatory responses (P-cells>PT-cells>T-cells) and their temporal features. For example, a substantial fraction of P-cells conducted stimulus-locking responses, but none of the T-cells did. f_r-dependency characteristics of the stimulus locking resembled that reported for "comodulation masking release," a behavioral model of auditory scene analysis. Each type neurons were omnipresent throughout the AI and none of them showed intrinsic oscillation. These findings suggest that: 1) T-cells preferentially encode spectral information with a rate-place code and 2) P-cells preferentially encode acoustic transients with a temporal code whereby rate-place coded information is potentially bound for scene analysis.

* *Corresponding author*. Fax: +81 55 273 6731.
E-mail address: msakai@yamanashi.ac.jp (M. Sakai).

Abbreviations: AI, primary auditory cortex; ANOVA, analysis of variance; BF, best frequency; #clicks$_{onset}$, number of clicks contributing to response onset; CMR, comodulation masking release; DR$_{max}$, maximum driven rate; f_r, repetition rate; FRF, frequency receptive field; P-cell, phasic cell; pe-SPL, peak equivalent sound pressure level; PSTH, peri-stimulus time histogram; PT-cellphasic-tonic cell; SD, standard deviation of the mean; SPL, sound pressure level; T-cell, tonic cell; tMTF, temporal modulation transfer function; rMTF, rate modulation transfer function

0006-8993/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2009.01.064

1. Introduction

In the real-world environment, energy from different sound sources often overlaps in the spectral and temporal domains. Nonetheless, our audition can promptly segregate or group such mixtures of information into their original sources (i.e., auditory scene analysis [for review, Bregman, 1990]). A behavioral model that chiefly captures spectral-cue-dependent aspects of auditory scene analysis is the "streaming effect" of repeating ABA— triplets, where "A" (low) and "B" (high) are pure tones of different frequencies and "—" is a silent gap (van Noorden, 1975; Snyder et al., 2008). Depending on the frequency separation between the tone A and B (ΔF), the signals can evoke two qualitatively different percepts: for relatively small ΔF, listeners usually perceive a single stream of tones in a galloping rhythm (ABA—ABA—); for large ΔF, galloping rhythm is lost and listeners usually report hearing 2 streams of tones (A—A—A—A— and B——B——). Another behavioral model, which chiefly captures its temporal-cue-dependent aspects, is "comodulation masking release (CMR)" (Hall et al., 1984; Schooneveldt and Moore, 1989). The CMR enables the detection of an otherwise masked tone signal by adding envelope fluctuation of background noise.

The single- and multi-unit studies in un-anesthetized animals revealed that several important features of stream segregation of ABA— triplets, including its dependence on ΔF, can be in principle accounted for in terms of frequency selectivity (Fishman et al., 2004; Micheyl et al., 2005), forward suppression (Fishman et al., 2001) and stimulus-selective adaptation (Ulanovsky et al., 2004, Micheyl et al., 2005) of the primary auditory cortex (AI) neurons. A possible neural correlate of CMR has been found also in single AI neurons of anesthetized cat: the neurons tended to lock to the amplitude envelope of slowly fluctuating noise whereas addition of a tone burst suppressed such stimulus-locking responses, a phenomenon referred to as "locking suppression" (Nelken et al., 1999; Las et al., 2005). Human studies using electroencephalography (EEG), magnetoencephalography (MEG) or functional magnetic resonance imaging (fMRI) concurrent with psychophysical tasks have revealed the neural activities consistent with the above described microelectrode data (for review, Snyder and Alain, 2007; Micheyl et al., 2007). Most of those studies, however, qualify the microelectrode data by revealing similar neural activities in higher-cortical areas as well (Harms et al., 2005; Wilson et al., 2007). Collectively, it was suggested that the AI is involved in both spectral- and temporal-cue-based auditory scene analysis.

It has been well documented that single AI neurons exhibit more diverse stimulus–response properties in un-anesthetized conditions as compared to anesthetized conditions (for review, Wang, 2007). For example, many single-unit studies in un-anesthetized animals AI have distinguished onset and sustained discharges in response to even simple acoustic stimuli like pure tones (Gerstein and Kiang, 1964; Wang et al., 2005), with the latter response feature further subdivided into distinct subtypes (Evans and Whitefield, 1964; Recanzone, 2000; Chimoto et al., 2002). This remarks a high contrast to the results in anesthetized animals that the AI neurons typically generate only onset responses to unmodulated acoustic stimuli, if at all (Zurita et al., 1994; Gaese and Ostwald, 2001; Heil and Neubauer, 2003). It also appears that the difference among the AI neurons is larger in un-anesthetized conditions than anesthetized conditions for shape (Gaese and Ostwald, 2003) and width (Recanzone et al., 2000; Qin and Sato, 2004) of frequency-receptive fields. However, it remains unclear whether spectral and temporal information, necessarily for auditory scene analysis, are integrated in identical AI neurons or differentially encoded by distinct AI neurons in un-anesthetized conditions. The present study addresses this issue by analyzing single-unit activities of un-anesthetized cat AI.

2. Results

2.1. *Responses in the pure-tone paradigm*

The results reported here are based on 92 single unit activities of cats AI, which exhibited excitatory responses in the pure-tone paradigm (see Experimental procedures for details). During the recording, the animal was kept awake (see Experimental procedures for the definition). By applying our previous criterion of the response time-course (Chimoto et al., 2002), the units were classified into three types: phasic cells (P-cell; $n=24$), tonic cell (T-cells; $n=31$) and phasic-tonic cell (PT-cells; $n=37$).

Fig. 1A-1 is a rater dot display of spike occurrence of a representative P-cell responding to 0.5-s-long tone bursts (horizontal bar at the bottom; same as follows) of variable frequencies. The stimulus intensity was set at the best SPL. The P-cell typically exhibited only phasic responses at the stimulus onset. Fig. 1A-2 illustrates the isointensity contour which plots driven rate as a function of tone frequency. The excitatory frequency-response field (excitatory-FRF; vertical bar) was identified as the frequency range at which the contour exceeded the threshold for excitatory responses (broken line). The best frequency (BF; arrow) was defined as the frequency where the contour reached the peak. Then, we calculated the "frequency-tuning bandwidth" as BF divided by the width of the excitatory-FRF; the larger values indicate sharper frequency tuning. Fig. 1A-3 is a peri-stimulus time histogram (2-ms bin width) over the excitatory-FRF, which revealed an onset latency of 10 ms. This histogram was converted into the post-latency time (PLT) histogram (50-ms bin width; A-4). As for the definition of P-cell, the driven rate declined below the threshold for excitatory responses (broken line) by the end of the stimuli. Autocorrelograms revealed no oscillations before (A-5), during (A-6) and after the stimuli (A-7) as was generally the case for the present study.

Fig. 1B displays responses of a representative T-cell. The definition of T-cell was of such a small adaptation that the driven rate remained above not only the threshold for excitatory responses but also remained half of the maximum driven rate ($1/2^{*}DR_{max}$; dotted line, B-4) throughout the stimuli. PT-cells correspond to the AI neurons that showed excitatory responses to pure tones, but did no meet the criterion for P-cells and T-cells. Fig. 1C displays responses of a representative PT-cell: the driven rate remained above the threshold for an excitatory response but declined below $1/2^{*}DR_{max}$ by the end of the stimuli (C-4). As long as the animal