Neuroscience 256 (2014) 309–321

RESPONSE CHARACTERISTICS OF PRIMARY AUDITORY CORTEX NEURONS UNDERLYING PERCEPTUAL ASYMMETRY OF RAMPED AND DAMPED SOUNDS

J. WANG, [a] **L. QIN,** [a,b] **S. CHIMOTO,** [a] **S. TAZUNOKI** [a] **AND Y. SATO** [a*]

[a] *Department of Physiology, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Chuo, Yamanashi 409-3898, Japan*

[b] *Department of Physiology, China Medical University, Shenyang 110001, People's Republic of China*

Abstract—Sound envelope plays a crucial role in perception: ramped sounds (*slow attack* and *quick decay*) are louder in strength and longer in subjective duration than damped sounds (*quick attack* and *slow decay*) even if they are equal in intensity and physical duration. To explain the asymmetrical perception, the *perceptual constancy* hypothesis supposes that the listener eliminates the slow decay of damped sounds from the judgment of perception, while the *persistence of perception* hypothesis supposes asymmetrical neural responses after the source has stopped. To understand neural mechanisms underlying the perceptual asymmetry, we explored response properties of the primary auditory cortex (A1) neurons during ramped and damped stimuli in awake cats. We found two distinct types of cells tuned to specific features of the sound envelope: *edge* cells sensitive to the temporal edge, such as quick attack and decay, while *slope* cells sensitive to slow attack and decay. The former needs a *short* (<2.5 ms) period of stimulus duration for evoking maximal peak responses, while the latter needs a *long* (20 ms) period, suggesting that the timescale of processing underlies differential sensitivity between the cell types. The findings suggest that perceptual constancy is not yet be executed at A1 because the specific cells distinguishing the direction of amplitude change (attack or decay) are lacking in A1. On the other hand, there is evidence of persistence of perception: overall response duration during ramped sound reached 1.4 times longer than that during damped sound, originating mainly from the response asymmetry of the edge cell (sensitive to the quick decay of ramped sounds but not to the slow decay of damped sounds), and neuronal *persistence of excitation* after the termination of ramped sounds was substantially longer than that of damped sounds, corresponding to the psychological evidence of persistence of perception.

Key words: auditory cortex, response asymmetry, response duration, sound envelope, integration time, perceptual asymmetry.

INTRODUCTION

Because all sounds begin and end at some time point, amplitude must increase (referred to as *attack*) at the onset and decrease (*decay*) at the offset. Patterson (1994a,b) called sound with an instantaneous attack followed by an exponential slow decay *damped* sound, while the time-reversed sound with an exponential slow attack followed by an instantaneous decay was called *ramped* sound. Ramped and damped sounds were shown to differ in terms of perception of timbre, loudness, and subjective duration: a ramped tone is stronger in *tonal* timbre (Patterson, 1994a,b), louder in strength (Irino and Patterson, 1996), and longer in subjective duration (Schlauch et al., 2001) than a damped tone, even if the two tones are equal in physical duration and intensity.

Attack is primarily determined by forces that drive an acoustic medium to vibrate: *plucking* the acoustic medium produces a sound with a quick attack, while *bowing* the medium produces a sound with a slow attack (Cutting and Rosner, 1974; Rosen and Howel, 1981; Cutting, 1982). Decay, on the other hand, is determined by multiple factors including both physical characteristics of the acoustic medium (Lutfi and Stoelinga, 2010) and the reverberant environment. Stecker and Hafter (2000) proposed that a listener has the ability to parse a sound to recover information about the physical characteristics of the acoustic medium and the reverberant environment. The asymmetrical perception in loudness was interpreted as the phenomenon of *perceptual constancy* related to the parsing of auditory input into direct and reverberant sound (Stecker and Hafter, 2000). The listener eliminates the slow decay portion from the judgment of loudness as a sense of reverberation, resulting in soft perception for damped sound.

For duration matching experiments of ramped and damped sounds in humans, Digiovanni and Schlauch (2007) instructed one group of participants to simply match the duration and the other group to includes all aspects of the sounds: the former yielded longer subjective duration for ramped sounds than damped sounds, by contrast, the latter significantly reduced the size of the asymmetry in subjective duration. They

---

*Corresponding author. Tel: +81-55-273-9482; fax: +81-55-273-6731.

E-mail address: yu_sato@yamanashi.ac.jp (Y. Sato).

Abbreviations: A1, primary auditory cortex; BF, best frequency; PSTHs, peri-stimulus time histograms; RDI, response duration index; SCI, spike count index; SPL, sound pressure level.

0306-4522/13 $36.00 © 2013 IBRO. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuroscience.2013.10.042

suggested that the reducing size of the asymmetry is related to the perceptual constancy and the remaining perceptual difference is due to a sensory process of *persistence of perception*, which is thought of as ‘‘the continuation of a sound’s internal representation in the auditory nervous system after the source has stopped’’ (Digiovanni and Schlauch, 2007). They showed, by temporal masking experiments in humans, the asymmetrical time course of the internal representation of the ramped and damped sounds (Digiovanni and Schlauch, 2007). Ries et al. (2008) showed that the persistence of perception hypothesis explained the loudness matching data.

Thus, it is well known that sound envelope plays a crucial role in perception. Nevertheless, little is known about the neural response characteristics during ramped and damped stimuli. We investigated (1) response characteristics of primary auditory cortex (A1) neurons of awake cats during ramped and damped stimuli, and (2) whether or not the response characteristics support hypotheses of perceptual constancy and persistence of perception.

EXPERIMENTAL PROCEDURES

Experiments were performed in accordance with the Guidelines for Animal Experiments, University of Yamanashi, and the Guiding Principles for the Care and Use of Animals approved by the Council of the Physiological Society of Japan.

Animal preparation, recording, and histology

Three cats were chronically prepared for single-unit recordings from both hemispheres of the auditory cortex in a manner similar to that described previously (Qin et al., 2007, 2008a,b, 2009). Under pentobarbital sodium anesthesia (initial dose 40 mg/kg) and aseptic conditions, cats had an aluminum cylinder (inner diameter 12 mm) implanted into the bilateral temporal bone for microelectrode access, at an angle of 10–20° from the sagittal plane. A metal block was embedded in the dental acrylic cap to immobilize the head. After at least 1 week of postoperative recovery, the cats were acclimated to the experimental conditions. Each cat’s body was gently wrapped in a cloth bag and the head was restrained with holding bars for several minutes. In successive daily sessions, the period was lengthened, and they were familiarized to sitting in an electrically shielded, sound-attenuated chamber. The animals were given food and water during the sessions. The conditioning procedure lasted at least 2 weeks. When recording experiments began, they sat with no sign of discomfort or restlessness. One day before the recording session, the bone (diameter 1–2 mm) at the bottom of the cylinder was removed, leaving the dura intact under ketamine anesthesia (initial dose 15 mg/kg).

The recording session began the following day. The dura was pierced with a sharpened probe, and an epoxylite-insulated tungsten microelectrode (impedance: 2–5 MΩ at 1 kHz; FHC Inc.) was advanced into the A1 with a remote-controlled micromanipulator (MO-951; Narishige). Extracellular single-unit activities were recorded and discriminated using a template-matching discriminator (ASD, Alpha-Omega Engineering). Search stimuli were tone bursts of variable single frequency and sound pressure level (SPL, in dB re: 20 μPa). The cat’s face, particularly the eyes, was continuously observed on a monitor connected to a camera in front of the cat. Saccadic eye movements and eye fixation were judged as signs of an awake state (Chimoto et al., 2002). Rapid eye movements of paradoxical sleep were easily identified by their characteristic appearance of half-opened eyelids, and were judged as signs of sleep (Chimoto et al., 2002). When drowsiness was suspected, the cat was alerted by gently tapping the body with a remote-controlled tapping tool, or by briefly opening and closing the door. The cats sometimes moved during recording sessions, producing artifacts in the recording. By carefully checking the monitors and the spike train, artifacts were marked in the computer in real time while recordings were in progress. Data with artifacts could therefore be excluded. Daily recording sessions lasted 3–5 h for 2–6 months for each animal. At the end of each session, the recording chamber was rinsed with sterile saline and antibiotic fluid (sulfamethoxazole, Taisho Pharmaceutical Co, Saitama, Japan), and sealed with Exafine (GC Corporation, Tokyo, Japan) and an aluminum cap. The animal was returned to its cage. The animals remained healthy throughout the experimental period.

At the termination of the experiment, some of the recording sites were marked with electrolytic lesions by passing cathodal current (25 μA, 10 s). Each animal was deeply anesthetized with sodium pentobarbital and perfused with 10% formalin before the brain was removed. The brain surface was photographed. The cerebral cortex was cut into transverse sections and stained with neutral red. Each section was captured using a scanner. On the basis of the lesion locations and electrode tracks, the recording sites were reconstructed on the image of the section.

Sound delivery

Sound delivery system was controlled by a custom-made program written in MATLAB (MathWorks). Digitally generated waveforms of sound stimuli were fed into a 16-bit digital-to-analog converter (PCI-6052E; National Instruments, Austin, USA) at a sampling frequency of 100 kHz and to an 8-pole Chebyshev filter (P-86; NF Electric Instruments, Yokohama, Japan) with a high cut-off frequency of 20 kHz. The outputs were sent to a low-output-impedance power amplifier (PMA2000III; Denon, Kawasaki, Japan) and played through a speaker (K1000; AKG, New York, USA) placed 2 cm away from the auricle contralateral to the recording site. We calibrated the sound delivery system between 128 and 16,000 Hz at frequency steps of 8 Hz, and the output varied by ± 5 dB. Harmonic distortion was less than −60 dB.

Stimulus paradigm and data analysis

While recording a single neuron, we first presented tone burst (5 ms in rise/fall time; 100 ms in stimulus duration)