Available online at www.sciencedirect.com

ScienceDirect

Hearing Research 229 (2007) 186–203

Hearing Research

www.elsevier.com/locate/heares

Research paper

Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1?

Jonathan B. Fritz *, Mounya Elhilali, Stephen V. David, Shihab A. Shamma

Centre for Auditory and Acoustic Research, University of Maryland, College Park, MD 20742, USA

Received 28 September 2006; received in revised form 27 November 2006; accepted 3 January 2007
Available online 16 January 2007

Abstract

Acoustic filter properties of A1 neurons can dynamically adapt to stimulus statistics, classical conditioning, instrumental learning and the changing auditory attentional focus. We have recently developed an experimental paradigm that allows us to view cortical receptive field plasticity on-line as the animal meets different behavioral challenges by attending to salient acoustic cues and changing its cortical filters to enhance performance. We propose that attention is the key trigger that initiates a cascade of events leading to the dynamic receptive field changes that we observe. In our paradigm, ferrets were initially trained, using conditioned avoidance training techniques, to discriminate between background noise stimuli (temporally orthogonal ripple combinations) and foreground tonal target stimuli. They learned to generalize the task for a wide variety of distinct background and foreground target stimuli. We recorded cortical activity in the awake behaving animal and computed on-line spectrotemporal receptive fields (STRFs) of single neurons in A1. We observed clear, predictable task-related changes in STRF shape while the animal performed spectral tasks (including single tone and multi-tone detection, and two-tone discrimination) with different tonal targets. A different set of task-related changes occurred when the animal performed temporal tasks (including gap detection and click-rate discrimination). Distinctive cortical STRF changes may constitute a “task-specific signature”. These spectral and temporal changes in cortical filters occur quite rapidly, within 2 min of task onset, and fade just as quickly after task completion, or in some cases, persisted for hours. The same cell could multiplex by differentially changing its receptive field in different task conditions. On-line dynamic task-related changes, as well as persistent plastic changes, were observed at a single-unit, multi-unit and population level. Auditory attention is likely to be pivotal in mediating these task-related changes since the magnitude of STRF changes correlated with behavioral performance on tasks with novel targets. Overall, these results suggest the presence of an attention-triggered plasticity algorithm in A1 that can swiftly change STRF shape by transforming receptive fields to enhance figure/ground separation, by using a contrast matched filter to filter out the background, while simultaneously enhancing the salient acoustic target in the foreground. These results favor the view of a nimble, dynamic, attentive and adaptive brain that can quickly reshape its sensory filter properties and sensori-motor links on a moment-to-moment basis, depending upon the current challenges the animal faces. In this review, we summarize our results in the context of a broader survey of the field of auditory attention, and then consider neuronal networks that could give rise to this phenomenon of attention-driven receptive field plasticity in A1.

Keywords: Auditory cortex; Receptive field; STRF; Plasticity; Attention; Acoustic salience; Ferret; A1

1. Introduction

How is the cortical representation of sound influenced by attention? Since the work of Hubel, Galambos and colleagues (Hubel et al., 1959), it has been known that the responses of single neurons in auditory cortex can be strongly modulated by attention. In their pioneering study in awake cat auditory cortex, a brief but prescient paper in

---

Abbreviations: A1, primary auditory cortex; MTD, multiple-tone detection task; STRF, spectrotemporal receptive field; $\mathrm{STRF}_{\mathrm{diff}}$, “difference” between two STRFs from the same neuron, measured in two behavioral states, such as active task and prior “passive” listening condition; TORC, temporally orthogonal ripple combination; NB, nucleus basalis; MMN, mismatch negativity; PFC, prefrontal cortex

* Corresponding author. Tel.: +1 301 405 6596; fax: +1 301 314 9920.
E-mail address: ripple@isr.umd.edu (J.B. Fritz).

0378-5955/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2007.01.009

the style of anecdotal neurophysiology, they observed that the responses of some cells (~10%) were highly dependent upon whether or not the cat was aroused by the presented sounds, or attended to the acoustic stimuli. As Hubel wrote (Personal communication 2006): "One day I entered the triple soundproof room to see if the cat was still alive, and discovered that rattling the doorknob, or keys, produced clear and lively responses… I found that almost anything I did that made a noise elicited firing as long as the cat appeared interested". Some of their "attention" units were located in A1, others in higher auditory cortical areas. In their study, they noted several characteristics of sounds that elicited an attentive state in their cats and lead to enhanced neural responses: (i) novelty (i.e. novel sounds were better than repeated sounds), (ii) meaning (i.e. natural sounds were better than clicks or tones), (iii) multi-sensory spatial coherence (i.e. acoustic stimuli presented simultaneously with a matched visual source were better than sounds without a matched visual counterpart). Although the cats were fully awake in the experiments of Hubel and colleagues, they were not behaviorally trained on any auditory task, so it was not possible in this early study to systematically explore the role of goal-directed attention in modulating sensory processing, a challenge left for future research. One caveat, noted by Hubel, is that this study did not control for pinnae movement, nor measure neuronal directionality tuning, thus leaving open the question of whether the observed effects were truly the results of either spatial or feature-based attention.

Other contemporary experimentalists working during this time period on the awake cat or monkey auditory cortex, such as Katsuki, did not mention the presence of any such "attention" units. However, in a fairly thorough study of responses in awake cat auditory cortex (Evans and Whitfield, 1964), the authors wrote: "About one third of the units responding to sound could be stimulated only by clicks or 'odd' sounds, such as the jangling of keys. Many of them gave inconsistent responses unless the attention of the cat was attracted to the source of the sound. These resemble the 'attention' units reported by Hubel et al. (1959). Some of these units had very low thresholds, but most required loud 'startling' sounds for consistent stimulation". Curiously, these researchers said almost nothing more about these "attention" neurons in the rest of this paper, or in two subsequent publications on the awake cat auditory cortex, perhaps because "…all of those units which required 'odd' sounds to stimulate them, or where much ingenuity and experiment were necessary to obtain the 'attention' of the unit, were obtained from cortex which was relatively inactive…". In the following 20 years, a handful of neurophysiological studies continued the investigation of the effects of auditory attention on cortical processing in the context of behavior (including Hocherman et al., 1976; Pfingst et al., 1978; Benson and Hienz, 1978; Miller et al., 1980). These studies demonstrated increases in cell evoked discharge for an attended stimulus compared to an identical non-attended stimulus and showed that these effects could occur with remarkably short-latency.

However, as Hubel and colleagues (1959) had ruefully noted: "Unfortunately attention is an elusive variable that no one has as yet been able to quantify". It remains so today. Although there has been considerable research on auditory attention over the past fifty years, using a variety of approaches (psychoacoustic, behavioral, neurophysiological (single unit and EEG), MEG, functional fMRI neuroimaging) the underlying neural mechanisms remain mysterious. Moreover, to make the problem even more challenging, there is clear evidence that attention itself, defined as a top-down selection process that focuses cortical processing resources on the most relevant sensory information in order to maintain goal-directed behavior in the presence of multiple, competing distractions, is hardly a unitary phenomenon, but may be comprised of several distinct behavioral and neural processes (Posner and Peterson, 1990; Desimone and Duncan, 1990; Parasuraman, 1998; Ahveninen et al., 2006; Johnson and Zatorre, 2006).

1.1. Brief review of auditory attention

So, what do we currently know about auditory attention? We know that auditory attention allows us to rapidly direct our acoustic focus towards sounds of interest in our acoustic environment. Attention can be bottom-up (sound-based) or top-down (task-dependent), and top-down control can trump involuntary attention switching to task-irrelevant distractor sounds (Sussman et al., 2003) perhaps through top-down attentional modulation by the prefrontal cortex of the deviance detection system in the auditory cortex (Doeller et al., 2003). Attention provides a top-down salience filter (Hafter et al., 2007) that in conjunction with bottom-up "pop-out" auditory salience (Kayser et al., 2005) is thought to pass only a small part of the incoming acoustic information to higher auditory areas. Attentional mechanisms can modulate neural activity encoding the spatial location and/or the acoustic attributes of the selected targets and the early sensory representation of attended stimuli (Ahveninen et al., 2006). This is illustrated by one of the best-known examples of auditory attention – the "cocktail party effect" (Cherry, 1953; Haykin and Chen, 2005) where we can easily selectively eavesdrop on different speakers in a crowded room brimming with multiple conversations. Cherry speculated on possible cues to its solution, including location, lip-reading, mean pitch differences, different speaking speeds, male/female speaking voices, or accents. However, whatever the cues, or the exact mechanisms involved in deciphering them, it is clear that in order to accomplish this feat of selective attention to a single stream in a natural environment with multiple sound sources, we must already be highly proficient at auditory scene segregation (or ASA). As Bregman's influential studies emphasized (Bregman, 1990), listeners have to solve the ASA problem in order to extract one or more relevant auditory streams from the mixture of sources that typify