Neuropsychologia 64 (2014) 218–229

Contents lists available at ScienceDirect

Neuropsychologia

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

Effects of task-switching on neural representations of ambiguous sound input

Elyse S. Sussman [a,*], Albert S. Bregman [b], Wei-Wei Lee [a]

[a] Albert Einstein College of Medicine, Bronx, NY, USA
[b] McGill University, Montreal, Quebec, Canada

ARTICLE INFO

Article history:
Received 18 April 2014
Received in revised form
13 September 2014
Accepted 22 September 2014
Available online 30 September 2014

Keywords:
Auditory scene analysis
Stream segregation
Mismatch negativity (MMN)
Auditory perception
Task-switching

ABSTRACT

The ability to perceive discrete sound streams in the presence of competing sound sources relies on multiple mechanisms that organize the mixture of the auditory input entering the ears. Many studies have focused on mechanisms that contribute to integrating sounds that belong together into one perceptual stream (integration) and segregating those that come from different sound sources (segregation). However, little is known about mechanisms that allow us to perceive individual sound sources within a dynamically changing auditory scene, when the input may be ambiguous, and heard as either integrated or segregated. This study tested the question of whether focusing on one of two possible sound organizations suppressed representation of the alternative organization. We presented listeners with ambiguous input and cued them to switch between tasks that used either the integrated or the segregated percept. Electrophysiological measures indicated which organization was currently maintained in memory. If mutual exclusivity at the neural level was the rule, attention to one of two possible organizations would preclude neural representation of the other. However, significant MMNs were elicited to both the target organization and the unattended, alternative organization, along with the target-related P3b component elicited only to the designated target organization. Results thus indicate that both organizations (integrated and segregated) were simultaneously maintained in memory regardless of which task was performed. Focusing attention to one aspect of the sounds did not abolish the alternative, unattended organization when the stimulus input was ambiguous. In noisy environments, such as walking on a city street, rapid and flexible adaptive processes are needed to help facilitate rapid switching to different sound sources in the environment. Having multiple representations available to the attentive system would allow for such flexibility, needed in everyday situations to maintain stable auditory percepts, and to allow rapid scanning of interesting events in a busy environment.

1. Introduction

In natural situations, the auditory environment is dynamically changing, with multiple sound sources overlapping in time. Imagine yourself walking on a busy city street as noises from the environment constantly change: a jet flies overhead; a jackhammer makes repeated sound bursts; cars drive by; people are talking as they walk past you. The neural representation or model of the auditory environment must account for the dynamics of such situations as the population of sounds in the auditory scene can change from moment to moment. Most studies investigating the neurobiological basis of auditory scene analysis (the process that allows us to hear distinct auditory sound events in noisy environments) have focused on how the brain disentangles two fixed sets of sound that differ in frequency (Bregman, 1978; Brochard et al., 1999; Carlyon et al., 2001; Cusack et al., 2004; Micheyl et al., 2005; Müller et al., 2005; Rahne et al., 2007; Rahne and Sussman, 2009; Shamma et al., 2011; Sussman, 2005; Sussman et al., 2007a, 2007b; Sussman-Fort and Sussman, 2014; Sussman et al., 1999; Sussman and Steinschneider, 2006; Szalárdy et al., 2013). However, most of the input to our ears overlaps dynamically in time and is often not clearly disambiguated. The current study investigated the problem of how ambiguous input is resolved by the auditory system, and what systems mediate perception of individual sound events when there are competing sound streams. We tested the hypothesis that the attended sound organization 'wins' out for neural representation when the input

* Correspondence to: Departments of Neuroscience and Otorhinolaryngology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461 USA. Tel.: +1 718 430 3313.
E-mail address: elyse.sussman@einstein.yu.edu (E.S. Sussman).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.09.039
0028-3932/

is ambiguous for hearing a single integrated stream or two segregated streams. Thus, we asked the question of whether focusing on one of two possible organizations (integrated or segregated) results in suppression of the alternative organization, or instead whether both organizations can be represented even when only one organization appears in perception.

Previously, we have used the mismatch negativity component (MMN) of event-related brain potentials (ERPs) as an index of sound organization (Sussman et al., 1998; Sussman et al., 1999; Sussman, 2013; Sussman, 2007). The MMN is elicited by discriminability of a deviation from a detected sound regularity maintained in the neural trace of the input (Näätänen et al., 2001; Näätänen et al., 2014). Any deviant can elicit MMN as long as the memory trace of the regularity is represented (Sussman, 2007). Our previous paradigms were set up such that within-stream regularities would only emerge when the sounds were neurophysiologically segregated, permitting detection of within-stream deviants (e.g., intensity or pattern deviants) (Rahne and Sussman, 2009; Sussman and Steinschneider, 2006; Sussman et al., 2005; Sussman, 2005). These studies allowed us to index sound organization, but the results could not distinguish whether both possible organizations were present simultaneously. This was because the MMN was elicited by deviant tones only when the sounds were segregated into distinct streams. We had no direct assessment of the integrated organization. For the current study, we modified a previous paradigm (Sussman and Steinschneider, 2006) to provide two distinct indices, one evoked only by integration and the other only by segregation (Fig. 1). Attention was cued to switch between performing one of two different tasks: a loudness detection task that required segregation of the higher from the lower tones (Fig. 1B), and a pattern identification task that required integration of the entire sequence (Fig. 1C). Thus, the demands of the task required either segregation of the sounds to perceive an oddball in one of the streams or integration of the sounds to perceive patterns of the stimuli. The task dictated which of the two possible organizations was to be active in perception. The MMN response was used to index whether a loudness change (e.g., intensity deviant) and/or a pattern deviant were detected based on different regularity representations from the same input.

The question of how ambiguous input is represented in auditory cortex has been of interest in the last decade (Sterzer et al., 2009; Nelken, 2004; Pressnitzer and Hupé, 2006; Sussman and Steinschneider, 2006; Rahne and Sussman, 2009; Denham et al., 2014; Winkler et al., 2009; Sussman, 2010). Much of the work has studied the temporal dynamics of a phenomenon called auditory bi-stability, in which spontaneous switching between two or more perceptual states occurs when the input is ambiguous. This is akin to the visual face-vase illusion (or binocular rivalry) in which perception of one excludes perception of the other at any given time even though switching back and forth between the two arises spontaneously. With long presentations of sounds, participants record the points at which they hear integrated, segregated, or some other organization of the input (e.g., Pressnitzer and Hupé, 2006; Denham et al., 2013; Denham et al., 2014; Winkler et al., 2012). The measures of the spontaneous switching between perceptual states across time indicate that whereas multiple alternative organizations of the input are proposed to be simultaneously maintained, one dominates in perception for a period of time (before switching to another) depending on the stimulus-driven characteristics of the input (Denham et al., 2013; Denham et al., 2014; Sterzer et al., 2009; Pressnitzer and Hupé, 2006). These results suggest that multiple sound organizations are available when the input is ambiguous and that spontaneous perceptual switching occurs as a result of low-level competition between the potential sound organizations (Denham et al., 2013; Horváth et al., 2001; Nelken, 2004; Pressnitzer and Hupé, 2006; Sterzer et al., 2009; Winkler et al., 2012). Although most of these studies report perception arising from spontaneous switching behavior, when participants were instructed to volitionally group (integrate) or segregate the sounds, it did not increase the duration of the designated perceptual state across trials but rather decreased the proportion of reported trials for the irrelevant organization (Pressnitzer and Hupé, 2006). Thus, it is not fully clear how attention modulates the ambiguous case to facilitate task performance. The results of the Pressnitzer and Hupé (2006) study, for example, suggest this may occur by reducing the intrusion of the unattended sound organization rather than by strengthening the attended one.

The current study tested effects of top-down processing on the storage of neural representations for ambiguous input. Attention was manipulated to control which organization was required to perform the task, while the elicitation of the MMN component was used as an index of which organization(s) was present during task performance. Thus, we assessed, using unique triggers for integration and segregation, whether one or both of the potential organizations were neurophysiologically represented while participants actively integrated or actively segregated the sound input to perform a task. This study thus asks the question of whether task performance modulates neural representations of ambiguous sound input toward the organization that is used to perform the task, and minimizes the irrelevant organization. What happens to the alternative, unattended organization when attention is focused on one set of the sounds?

If mutual exclusivity at the neural level was the rule, attention to one of two possible organizations would preclude neural representation of the other. This would predict that when attention biased the neurophysiological response toward the organization needed to perform the task, MMN would be elicited only by the deviants produced by the organization that was induced by the task, suppressing the alternative, unattended organization. Alternatively, if both organizations were represented, regardless of the task performed, then MMNs would be elicited by unattended non-target deviants of the organization that was not required by the task. Thus, we determine if both organizations are represented in the neural trace or if attention ‘resolves’ the ambiguity toward the organization that is used to perform the task.

2. Experiment 1: ambiguous (5 ST) condition

2.1. Methods and materials

2.1.1. Participants

Fifteen healthy young adults (10 male) aged 18–33 years old ($M=26$ years, $SD=4$) participated in the study. All participants passed a hearing screen (20 dB HL for pure tones at 500, 1000, 2000, and 4000 Hz, bilaterally). Participants gave written consent and were paid for their participation. Three participants could not perform at least one of the tasks. Their data were excluded from analyses. The remaining 12 participants' data are included in this report.

2.1.2. Stimuli

Stimuli were created using Neuroscan software (STIM, Compumedics, Corp., Charlotte, NC). Two pure tones (50 ms. duration, 5 ms. rise/fall time) were presented bilaterally through insert earphones (E-A-R-tone® 3A, Indianapolis, IN). Sounds were calibrated for peak-to-peak equivalent sound pressure level (ppe SPL) using a Brüel & Kjær sound level meter (2209) with an artificial ear (4152). There was a 5 semitone [ST] distance between tones: the higher frequency tone (H) was 1397 Hz and the lower frequency tone (L) was 1046 Hz (Fig. 1A). Tones were presented in three patterns: HLHH (Pattern1), HHLH (Pattern2), and HHHL (Pattern3), randomly distributed, with Pattern 1 occurring 40%, Pattern 2 occurring 40% and Pattern 3 occurring 20% of the time (Fig. 1A). Stimulus onset asynchrony (SOA) from high tone to high tone was 220 ms; SOA from high tone to low tone was 110 ms; and 440 ms separated the four-tone patterns (Fig. 1A). Thus, the inter-stimulus interval (ISI) for the high tone stream (the interval between successive