Auditory and Vestibular Research

Differential effects of temporal and spectral regularities on auditory streaming

Sanaz Soltanparast[1], Reyhane Toufan [2], Saeed Talebian[3], Akram Pourbakht [2]*

Highlight

- Acoustic regularities of competing auditory streams aid in target detection
- Spectral and temporal regularities have a distinct effect on auditory streaming

Abstract

Background and aim:: The concept that recognizing sound regularities plays a major role in the segregation of auditory streams has lately gained significant interest. Thus, this study was designed to investigate how temporal and spectral regularities incorporated into the background auditory stream affect auditory stream segregation.

Methods: An indirect measure of auditory streaming task (i.e., detecting rare-level targets) was implemented in twenty-five healthy young adults. Participants were presented with two concurrent auditory streams involving foreground and background ones .Participants were instructed to detect rare-level targets in the foreground stream during three experimental conditions. These conditions vary based on the background auditory stream, which contained repeating temporal and spectral patterns alongside elements of randomness.

Results: Temporal and spectral regularities of the background auditory stream significantly increased the hit rate compared to random structure. Notably, this effect of regular cues on target detection and, possibly in turn, stream segregation was significantly greater for temporal compared to spectral regularities.

Conclusion: These findings showed that incorporating temporal or spectral regularities in the background auditory stream facilitated target detection and, possibly in turn, stream segregation. This perceptual regularity benefit was greater for temporal regularities than spectral regularities. These findings might present primary evidence for distinct facilitating effects of various theoretical frameworks of sound feature regularities on auditory streaming.

Keywords: auditory scene analysis, stream segregation, temporal regularity, spectral regularity, regularity encoding

Introduction

In everyday auditory scenarios, a mixture of several acoustic sources constantly impinges on the auditory system [1,2]. Thus the critical task of the auditory system is to organize this mixture into perceptually meaningful units (i.e., auditory objects) and map the relevant foreground and irrelevant background objects [3-5]. In this context, object formation includes integrating sounds from the common source into a single stream

[1.] Department of Audiology, School of Rehabilitation, Babol University of Medical Sciences, Babol, Iran.
Sanaz Soltanparast, email:Sanaz.audiologist@gmail.com. ORCID ID: https://orcid.org/0000-0003-4072-8701

[2.] Rehabilitation Research Center, Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran
Reyhane Toufan, email: r.toufan@ gmail.com. ORCID ID: https://orcid.org/0000-0001-9615-9259

Akram Pourbakht, email: Pourbakht.a@iums.ac.ir ORCID ID: https://orcid.org/0000-0002-9779-3558

[3]. Motor Control Laboratory, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.
Saeed Talebian, email: talebian@tums.ac.ir. ORCID ID: https://orcid.org/0000-0002-4945-8085

*__Corresponding author:__ Akram Pourbakht

Rehabilitation Research Center, Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, at Madadkaran Street, Shahid Shahnazari Street, Madar Square, Mirdamad Boulevard, Tehran 15459-13487, Iran. E-mail address: Pourbakht.a@iums.ac.ir (A. Pourbakht).

and concomitantly segregating those from the competing sound sources into separate streams, a process termed auditory scene analysis (ASA) [6,7]. Several studies have exploited the auditory streaming paradigm to display the acoustical attributes necessary for ASA, including sequences of sequentially repeating tones rendered in an ABA_ ABA or A-B-A-B model. These studies revealed that the segregation of the auditory stream takes advantage of different acoustical attributes (e.g., frequency separation (Δf)) [8,9]. Classically, any large acoustical differences amid A and B sounds may result in stream segregation (segregated percept), while small acoustical differences amid A and B boost stream integration (integrated percept) [6,7].

Some researchers recently investigated auditory streaming using a long-lasting sequence of ABA_ABA streaming stimuli that simulate bi-stable auditory perception in many natural acoustic situations. These researchers showed that, despite constant stimulus design, the perception of these sequences alternates between the integrated and segregated percepts [10-13]. Thus, perceptual organizations of bi-stable ABA_ ABA sequences are not completely characterized via these stimuli' acoustical attributes (e.g., Δf) [11,14]. A few recent investigations utilize the bi-stable essence of perception within auditory streaming paradigms to specify the practical role of acoustic regular cues in analyzing complex auditory scenes. According to these psychophysical studies, extracting the regularities within ongoing auditory sequences increased the likelihood of the perception of segregated streams [15-18]. In addition, extracting these regularities facilitated segregating a foreground stream from background even when the regularities were superimposed on background [19, 20]. However, further studies are necessitated to replicate these findings.

Recent studies have proposed a difference between 'temporal' regularities and other sound regularities (e.g., spectral regularities) in auditory processing and neural sources [21-26]. Notably, recent studies showed distinct dedicated neural sources for encoding temporal (e.g., parietal cortex) and spectral (e.g., inferior frontal gyrus and medial frontal gyrus) regularities [21-23]. Concerning distinct auditory processing and neural sources of temporal and spectral-based regular sound features, it can be assumed that the different types of sound regularities might affect stream segregation differently. Therefore, in this study, we investigated whether auditory regularities facilitate stream segregation and whether distinct types of regularities differentially influence this perceptual process. These issues were considered in the present study by comparing temporal and spectral sound regularities with an indirect measure of auditory streaming (a task of detecting rare intensity level target stimuli). Target stimuli were introduced in an irregular stream of sounds presented simultaneously with another auditory stream of either temporal or spectral regularities. This task imitated a natural acoustic scene by small spectral differences and arbitrary temporal overlaps between the two sound streams. The successful task implementation, detecting rare-level targets in foreground irregular stream, necessitated participants segregating the foreground auditory stream from a temporally or spectrally regularly repeating background stream.

Methods

Participants

Twenty-five right-handed healthy young volunteers (twenty females and five males) have participated in this study. Participants' age ranged from 19 to 33 years (mean age=23.24±3.43). All had pure tone thresholds ≤ 25 dB HL from 250 to 8000 Hz, with no well-known neurological or psychiatric conditions and no head injury experience. This study was done in the Rehabilitation School of Iran University of Medical Sciences and was confirmed by the Ethics Committee of Iran University of Medical Sciences, Tehran, Iran (IR.IUMS.REC.1397.1027).

Stimuli and experimental design

Behavioral testing and electroencephalogram recordings were conducted simultaneously in an acoustically and electromagnetically sound-attenuated booth. In testing, the participants were seated gently in a comfortable chair and were asked to recognize the target stimuli. Sounds were generated in MATLAB 2016, b software (The MathWorks, Inc.) at a 44.1 kHz sampling rate with 16-bit resolution. The Cogent toolbox controlled the presentation of the stimuli.

The stimuli consist of two distinct auditory streams: a foreground of "A" tones and a background of "B" tones. To simulate the challenging everyday listening conditions, auditory stimuli of foreground and background streams were created separately and then aggregated. Thus, the resulting stimuli occasionally comprised

temporal convergence between "A" and "B" tones similar to real-world auditory scenes (Fig.1). Tones in each stream were pure tones of 75 ms in duration, including 10 ms rise/fall time, and were introduced binaurally via ER-3A insert earphones (Etymotic ER-3A). The "A" tones within the foreground stream were set at 630 Hz and were introduced at the intensity level of 70 dB SPL. The tones in the "A" stream had random Inter Stimulus Intervals (ISIs) of 100 to 250 ms in 50ms steps. Rare-level target tones that blend quasi-randomly in the "A" stream were introduced at higher intensity of 80 dB SPL. In all experimental blocks, target tones were pseudo-randomly interlaced within the sequence. To achieve this, no target tones were presented in the first 30 seconds of the 'A' stream, and the two rare target tones were never presented consecutively. The minimum and the average inter-target times were 2 and 10 seconds, respectively**.** Tones within the "B" stream had a random level of 65-85 dB SPL with 1dB steps. Indeed, for the participants to perceive 80 dB, A sounds as targets, all 70 dB sounds of this stream should have been separated from intervening stimuli of the B stream, which had random intensity changes. The "B" tones had a mean frequency value of 529 Hz (B1 =510 Hz, B2 = 529 Hz, B3 = 548 Hz). The frequency separation between the "A" and "B" streams was three semitones (i.e., 101 Hz). This frequency separation was specified based on previous investigation [20], which established the most challenging listening condition where the segregation of two streams cannot be assigned to frequency separation. Tones within the "B" stream were arranged in three experimental conditions: temporal regularity, spectral regularity, and random conditions. Tones within temporal regularity condition had an irregular frequency pattern like B2B1B2B2B1B1B2B1B3, while they had constant and regular ISI (175 ms). In the spectral regularity condition, tones were presented with an ascending regular frequency pattern (i.e., B1B2B3B1B2B3) but with randomly changed ISIs of 100 to 250 ms with 50ms steps. In the random experimental condition, tones within the background "B" stream were presented with no temporal or spectral regular patterns.
A training block including all experimental conditions was introduced before the main experiment to familiarize the participants with the task. In a training block, the participants had to detect target stimuli in three experimental conditions in a directed context. The occurrence of the auditory target was emphasized via illuminating a green light at the center of the computer screen. Afterward, participants had to accomplish at least one practice block of all experimental conditions without visual support. Participants were given further training blocks until they felt comfortable with the task. Participants were provided feedback regarding their execution following each training block as well as every block all over the experiment. Subsequently, the primary experiment was carried out. In the primary experiment, each condition was introduced in five-minute blocks, and each block was randomly repeated 4 times. In each block, target stimuli appeared between 25 and 36 times in the "A" stream. Participants were instructed to press a button whenever they detected the rare-level targets in the foreground stream. After each randomly repeated block, break times were given to the participants to get rid of fatigue. Across the experiment, participants were not informed about regularity manipulation within the background stream. The full experimental session—including training, rest periods, and behavioral and electrophysiological testing—lasted approximately 4 hours.

Data analysis
To determine the participant's ability to detect the rare level target, hit rate (HR) and false alarm rate (FAR) were computed for each experimental condition. All responses occurring in a 300-1000 ms following the onset of target were judged as hits (this interval was defined according to the pilot study). Others were considered false alarms. Because the current stimulus framework required sustained stimulus introduction, that all non-targets in both streams can appear throughout the response window, the amount of false alarms is highly dependent on the response window employed in the analysis. We utilized a procedure that considers the response window. A condition total presentation time was divided into time bins with durations corresponding to the range of 300-900 ms response window (i.e., 600 ms). For every participant and condition , the hit rate was determined by dividing the whole duration of time bins where a hit response was expected (amount of hits by time bin) by the sum of the time bins where hits were feasible (amount of targets multiplied by time bin). Since the time bin variables in the numerator as well as denominator cancelled, this may be condensed to: HR = n hits/n targets. Independently for each subject and condition, the false alarm rate was determined as the sum of time bins that had false alarms divided by the sum of time bins where a false alarm was feasible [19].

HR = n hits/n targets

$$\text{False alarm rate} = \frac{nfalse\ alarms\ \times time\ bin}{overall\ presentation\ time - (ntargets \times time\ bin\ )}$$

Statistical analysis

All Statistical analyses were carried on via the SPSS version 17 software package (SPSS Inc., Chicago, IL) at significance values of ($p < 0.05$). To assess the effects of regularity manipulation, a repeated measure ANOVA was done on each behavioral measure with experimental condition (spectral, temporal, or random) as a factor, and post hoc comparisons were carried out using a Bonferroni correction.

Results

Table 1 displays the average and standard deviation of behavioral performance in all experimental conditions. A repeated measure ANOVA showed statistically significant main effects of sound regularities on Hit rate ($df = 1.830$, $f\ (23) = 16.740$, $p < 0.001$, $\eta^2_P = 0.411$). Corresponding to post hoc comparisons the Hit rate was significantly different between the temporal and spectral conditions ($p = 0.004$), the temporal and random conditions ($p < 0.001$) as well as between the spectral and random conditions ($p = 0.010$) (Fig.2). Furthermore, a repeated measure ANOVA result revealed no significant effect of sound regularities on false alarm rate ($df = 1.827$, $f\ (23) = 1.188$, $p = 0.311$, $\eta^2_P = 0.047$). In conjunction with the information provided in the paper, case information can be accessed by submitting a written request to the associated author.

Discussion

The current study aimed to investigate whether: 1) the auditory system's capacity to discover the regularly repeating sound features incorporated in a background auditory stream affect stream segregation and 2) whether auditory stream segregation is distinctively affected by the different types of regularly repeating sound features (i.e., spectral and temporal regularities). The present study showed that regularly repeating sound features incorporated into a background auditory stream enhanced target detection performance. This finding suggests that regularity modulates auditory stream segregation processes. Additionally, different types of regularly repeating sound features (spectral and temporal regularities) modulated target detection differently.

This study showed that in temporal and spectral regularity conditions, participants' target detection was higher than in random conditions. To solve the current task, detecting rare level targets, the participants must segregate two concurrent auditory streams. Since the frequency separation between streams was too small to induce segregation, enhancement of target detection in regular conditions can pertain solely to the effect of manipulating regular cues. This finding indicates that incorporating regularities in a background auditory stream enhances target detection in a concurrent irregular stream, demonstrating that regularity influences stream segregation. Present study findings are in line with previous psychophysical studies on the role of regularly repeating sound features in auditory stream segregation during a subjective reports task [16-18]. For instance, Bendixen et al. (2010) examined the role of regularity on stream segregation using bi-stable ABA-ABA streaming stimulus with or without regular patterns (frequency and/or intensity) [18]. They asked their listeners to indicate their perception of sound sequences constantly. These authors reported that independently introduced regular patterns in the A tones, B tones, or both stabilized the perception of two streams for longer durations. Thus their findings demonstrated that regular patterns increased the likelihood of perceiving segregated auditory streams. The effect of regularly repeating sound features on the segregation of streams has also been found in objective listening measure of auditory stream segregation (a within-stream deviant detection task). For instance, Andreou et al. (2011) studied the role of regularity in auditory stream segregation as a function of different Δfs during an objective measure of a within-stream target detection task [20]. They introduced temporal regularity in background stream while tones in foreground stream were arranged irregularly. Their chief finding was that manipulating regularity throughout the background stream enhanced foreground stream segregation, primarily when Δfs alone was insufficient. Similarly, Rimmele et al. (2012) demonstrated that stream segregation depends on the auditory system's regular patterns detection ability, whether the regular patterns occur in foreground or background streams during an intensity-deviation detection task [19].

Current study showed that temporal and spectral regularities in the background auditory streams had different effects on stream segregation, as measured by the hit rate of rare level targets occurred within an irregular foreground stream. In fact, it's possible to conclude that participants experienced greater regularity benefits for the stream segregation when temporal regular patterns were embedded in the background stream than spectral regular patterns. It probably means that, the regularity benefit was related to the type of regularly repeating features embedded in the background stream. In fact, empirical evidence increasingly indicates that listeners utilize sophisticated mechanisms to track the rhythmic properties, including micro-timing differences, of acoustic sources. This enables the rapid formation of temporal expectancies, thereby optimizing behavioral responses [20]. On the other hand, previous studies have proposed that the distinctiveness of neural sources is associated with different functional roles[27,28]. Therefore, the finding that temporal and spectral regularities differentially affect stream segregation can be explained by distinct neural mechanisms dedicated to encoding temporal [22] and spectral regularities[21,23]. This result is in contrast with the finding of [18] study showing that introduced regularly repeating frequency and intensity patterns into the either one or both streams of bi-stable ABA streaming stimulus had equal effects on the ratio of segregated percept. The disagreement can be explained by methodological differences like the frequency separation between the A and B streams and the type of inspected regularities. Moreover, in the current study, the task employed to derive auditory streaming aimed to imitate real-world auditory scenes (e.g., busy street or busy restaurant) in contrast to the [27] study which used temporally non-converging regularly repeating ABA patterns which were far from complex natural auditory scenes.

Conclusion

The present study's data expanded the prior restricted literature regarding the facilitating effects of regularly repeating acoustic cues in the segregation of auditory streams. Specifically, this is the first study that reveals regularities carried out by distinct stimulus features probably result in differential modulations of auditory stream segregation. On the other hand, Gaining insights into how sound regularities influence auditory perception may provide valuable information for upcoming studies on music and speech auditory processing.

Acknowledgments: This study was a part of an approved Ph.D. dissertation project in audiology by the Iran University of Medical Sciences (grant number: 13320). The authors thank Amin Asgharzade for aid in stimulus production and data analysis.

Conflict of interest: the authors declare that they have no competing interest.

Author Contributions; SS: Study design, acquisition of data, interpretation of the results, and drafting the manuscript; RT: Study design and interpretation of the results; ST: Study design and statistical analysis; AP: Interpretation of the results and drafting the manuscript.

References

1. Power AJ, Lalor EC, Reilly RB. Endogenous auditory spatial attention modulates obligatory sensory activity in auditory cortex. Cereb Cortex. 2011; 21(6):1223-30. [DOI:10.1093/cercor/bhq233]
2. Alain C, Arnott SR. Selectively attending to auditory objects. Front Biosci. 2000; 5:D202-12. [DOI:10.2741/Alain]
3. Chakrabarty D, Elhilali M. A Gestalt inference model for auditory scene segregation. PLoS Comput Biol. 2019; 15(1):e1006711. [DOI:10.1371/journal.pcbi.1006711]
4. Snyder JS, Gregg MK, Weintraub DM. Attention, awareness, and the perception of auditory scenes. Front Psychol. 2012; 3:15. [DOI:10.3389/fpsyg.2012.00015]
5. Deike S, Denham SL, Sussman E. Probing auditory scene analysis. Front Neurosci. 2014; 8:293. [DOI:10.3389/fnins.2014.00293]
6. Paredes-Gallardo A, Innes-Brown H, Madsen SM, Dau T, Marozeau J. Auditory stream segregation and selective attention for cochlear implant listeners: Evidence from behavioral measures and event-related potentials. Front Neurosci. 2018; 12:581. [DOI:10.3389/fnins.2018.00581]
7. Gutschalk A, Micheyl Ch, Melcher JR, Rupp A, Scherg M, Oxenham AJ. Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci. 2005; 25(22):5382-8. [DOI:10.1523/JNEUROSCI.0347-05.2005]
8. Xiang J, Simon J, Elhilali M. Competing streams at the cocktail party: Exploring the mechanisms of attention and temporal integration. J Neurosci.2010; 30(36):12084-93. [DOI:10.1523/JNEUROSCI.0827-10.2010]
9. Moore BC, Gockel H. Factors influencing sequential stream segregation. Acta Acustica United with Acust. 2002; 88(3):320-33.
10. Szalárdy O, Bőhm TM, Bendixen A, Winkler I. Event-related potential correlates of sound organization: Early sensory and late cognitive effects. Biol Psychol. 2013; 93(1):97-104. [DOI:10.1016/j.biopsycho.2013.01.015]
11. Denham SL, Gyimesi K. Perceptual bistability in auditory streaming: How much do stimulus features matter? Learn Percept. 2013; 5(Supplement 2):73-100. [DOI:10.1556/LP.5.2013.Suppl2.6]
12. Denham SL, Gyimesi K, Stefanics G, Winkler I. Stability of perceptual organisation in auditory streaming. In: Lopez-Poveda E, Palmer A, Meddis R, editors. The neurophysiological bases of auditory perception. New York: Springer; 2010. [DOI:10.1007/978-1-4419-5686-6_44]

13. Winkler I, Takegata R, Sussman E. Event-related brain potentials reveal multiple stages in the perceptual organization of sound. Brain Res Cogn Brain Res. 2005; 25(1):291-9. [DOI:10.1016/j.cogbrainres.2005.06.005.]
14. Bendixen A. Predictability effects in auditory scene analysis: A review. Front Neurosci. 2014; 8:60. [DOI:10.3389/fnins.2014.00060]
15. Bendixen A, Denham SL, Winkler I. Feature predictability flexibly supports auditory stream segregation or integration. Acta Acust united Acust. 2014; 100(5):888-99. [DOI:10.3813/AAA.918768]
16. Bendixen A, Bőhm TM, Szalárdy O, Mill R, Denham SL, Winkler I. Different roles of similarity and predictability in auditory stream segregation. Learn Percept. 2013; 5(Supplement 2):37-54. [DOI:10.1556/LP.5.2013.Suppl2.4]
17. Bendixen A, Denham SL, Winkler I. Sound predictability as a higher-order cue in auditory scene analysis. In: AIA-DAGA 2013 Conference on Acoustics, 18-21 March 2013, Merano, Italy. DOI ندارد
18. Bendixen A, Denham SL, Gyimesi K, Winkler I. Regular patterns stabilize auditory streams. J Acoust Soc Am. 2010. 128(6):3658-66. [DOI:10.1121/1.3500695]
19. Rimmele J, Schröger E, Bendixen A. Age-related changes in the use of regular patterns for auditory scene analysis. Hear Res. 2012; 289(1-2):98-107. [DOI:10.1016/j.heares.2012.04.006]
20. Andreou LV, Kashino M, Chait M. The role of temporal regularity in auditory segregation. Hear Res. 2011; 280(1-2):228-35. [DOI:10.1016/j.heares.2011.06.001]
21. Soltanparast S, Toufan R, Talebian S, Pourbakht A. Regularity of background auditory scene and selective attention: A brain oscillatory study. Neurosci Lett. 2022; 772:136465. [DOI:10.1016/j.neulet.2022.136465]
22. Sohoglu E, Chait M. Detecting and representing predictable structure during auditory scene analysis. Elife. 2016; 5:e19113. [DOI:10.7554/eLife.19113]
23. Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M. Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc Natl Acad Sci U S A. 2016; 113(5):E616-25. [DOI:10.1073/pnas.1508523113]
24. Schwartze M, Farrugia N, Kotz SA. Dissociation of formal and temporal predictability in early auditory evoked potentials. Neuropsychologia. 2013; 51(2):320-5. [DOI:10.1016/j.neuropsychologia.2012.09.037]
25. Hughes G, Desantis A, Waszak F. Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. Psychol Bull. 2013; 139(1):133-51. [DOI:10.1037/a0028566]
26. Sperduti M, Tallon-Baudry C, Hugueville L, Pouthas V. Time is more than a sensory feature: Attending to duration triggers specific anticipatory activity. Cogn Neurosci. 2011; 2(1):11-8. [DOI:10.1080/17588928.2010.513433]
27. Shalgi S, Deouell LY. Direct evidence for differential roles of temporal and frontal components of auditory change detection. Neuropsychologia. 2007; 45(8):1878-88. [DOI:10.1016/j.neuropsychologia.2006.11.023]
28. Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage. 2002; 15(1):167-74. [DOI:10.1006/nimg.2001.0970]

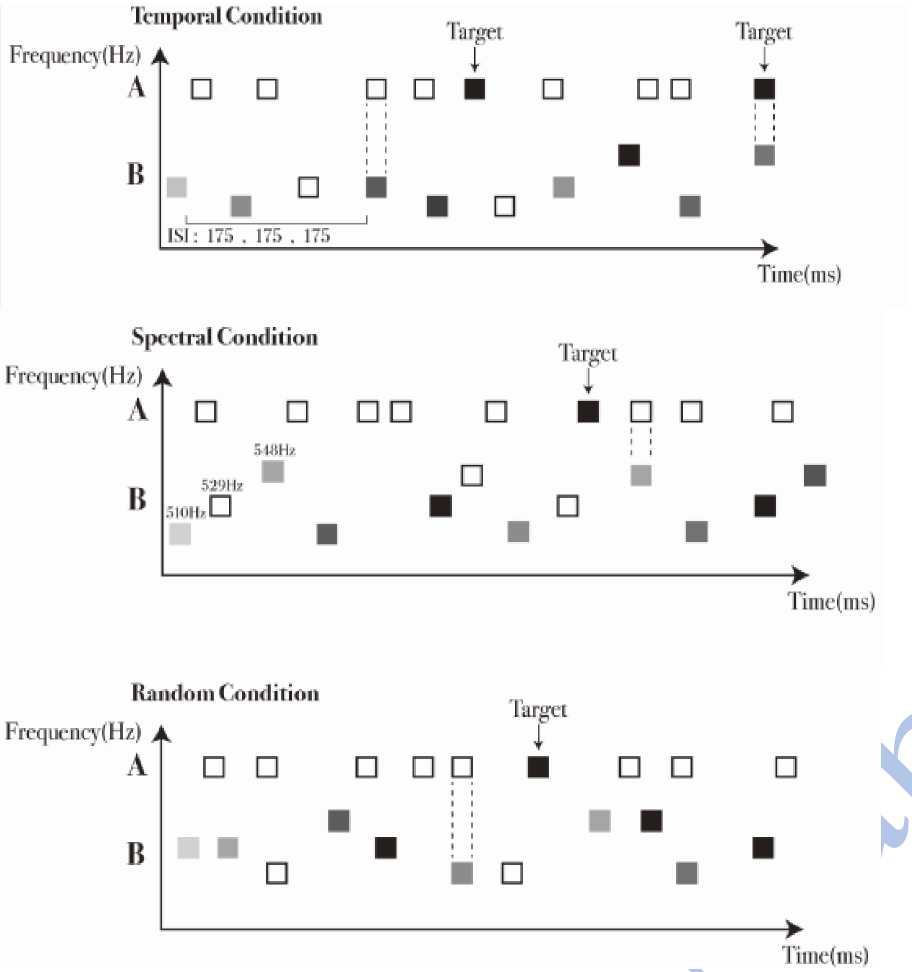

Fig 1. A schematic presentation of the experimental design. Stimuli were introduced in three distinct conditions: temporal regularity condition in which tones within the "B" stream had constant and regular ISI of 175 ms; Spectral regularity condition in which tones within the "B" stream had ascending regular frequency pattern and random condition in which tones within the "B" stream had no temporal or spectral regular patterns. The random intensity level of the B tones is specified with the different colors of the square. Temporal convergence between "A" and "B" tones was exhibited via the dashed line. In all experimental conditions, the rare level targets were exhibited via arrows.

Table1. Descriptive statistics of behavioral performance in all experimental conditions.

	Temporal condition		Spectral condition		Random condition	
	Mean	SD	Mean	SD	Mean	SD
Hit rate	34.94	19.80	30.86	18.89	28.11	19.33
False alarm rate	3.06	1.29	3.22	1.23	3.24	1.49

Note. SD = Standard Deviation.

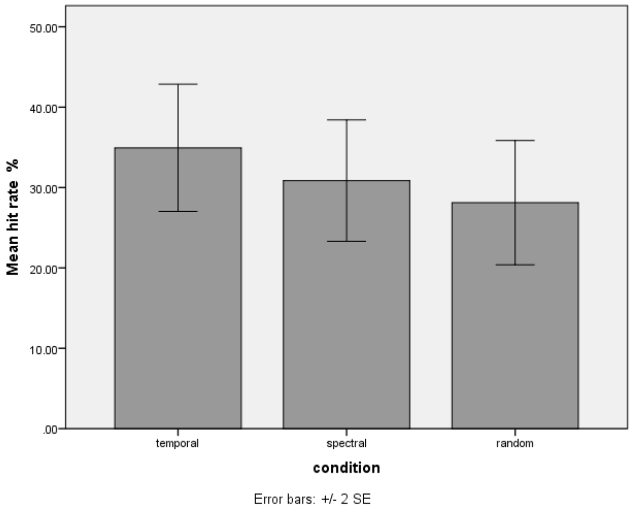

Fig 2. The impact of experimental conditions on the hit rate.