Research Article

Different Effects of Temporal and Spectral Regularities on Auditory Stream Segregation

Sanaz Soltanparast[1], Reyhane Toufan[2,3], Saeed Talebian[4], Akram Pourbakht[2,3*]

1. Department of Audiology, School of Rehabilitation, Babol University of Medical Sciences, Babol, Iran
2. Rehabilitation Research Center, Iran University of Medical Sciences, Tehran, Iran
3. Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran
4. Motor Control Laboratory, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

Use your device to scan and read the article online

Citation: Soltanparast S, Toufan R, Talebian S, Pourbakht A. Different Effects of Temporal and Spectral Regularities on Auditory Stream Segregation. Aud Vestib Res. 2026;35(3):270-6.

https://doi.org/10.18502/avr.v35i3.22184

Highlights

- Acoustic regularities of competing auditory streams help with target detection
- Spectral and temporal regularities have different effects on auditory streaming

Article info:
Received: 04 Aug 2025
Revised: 23 Sep 2025
Accepted: 08 Oct 2025

* **Corresponding Author:**
Rehabilitation Research Center, Iran University of Medical Sciences, Tehran, Iran.
pourbakht.a@iums.ac.ir

ABSTRACT

Background and Aim: The fact that recognizing sound regularities plays a crucial role in segregating auditory streams has recently garnered significant interest. This study aimed to investigate how temporal and spectral regularities incorporated into the background auditory stream affect auditory stream segregation.

Methods: Twenty-five healthy young adults performed an indirect measure of the auditory streaming task (i.e. detecting rare-level targets). They were presented with two concurrent auditory streams involving foreground and background streams, and were asked to detect rare-level targets in the foreground stream during three experimental conditions (spectral, temporal, or random regularities).

Results: Temporal and spectral regularities of the background auditory stream significantly increased the hit rate compared to random regularities. This effect of regular cues on target detection and possibly stream segregation was significantly greater under temporal regularity conditions than under spectral regularity conditions.

Conclusion: The findings suggest that incorporating temporal or spectral regularities into the background auditory stream can facilitate target detection and possibly stream segregation. This perceptual regularity benefit is greater for temporal regularities than for spectral regularities. These findings may provide primary evidence for the distinct facilitating effects of various theoretical frameworks on sound feature regularities and auditory streaming.

Keywords: Auditory scene analysis; stream segregation; temporal regularity; spectral regularity; regularity encoding

Introduction

In everyday auditory scenarios, a set of several acoustic sources constantly affects the auditory system [1, 2]. Thus, the critical task of the auditory system is to organize this set into perceptually meaningful units (i.e. auditory objects) and map the relevant foreground and irrelevant background objects [3-5]. In this context, auditory object formation involves integrating sounds from a common source into a single stream and, simultaneously, segregating those from competing sound sources into separate streams, a process known as Auditory Scene Analysis (ASA) [6, 7]. Several studies have used the auditory streaming paradigm to show the acoustical attributes necessary for ASA, including sequences of repeating tones organized in an ABA-ABA or A-B-A-B model. These studies have revealed that the segregation of the auditory stream utilizes different acoustical attributes (e.g. frequency separation) [8, 9]. Classically, any large acoustical differences between A and B tones may result in stream segregation (segregated percept), while small acoustical differences between A and B tones boost stream integration (integrated percept) [6, 7].

Some studies have 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. According to these studies, despite constant stimulus design, the perception of these sequences alternates between the integrated and segregated percepts [10-13]. Therefore, perceptual organizations of bi-stable ABA-ABA sequences are not completely characterized by acoustical attributes of these stimuli, including frequency separation [11, 14]. A few recent studies utilized 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 increases the probability of perceiving them as segregated streams [15-17]. In addition, extracting these regularities facilitates the segregation of a foreground stream from the background, even when the regularities are superimposed on the background [18, 19]. However, further studies are necessary 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 [20-25]. Notably, they have shown distinct dedicated neural sources for encoding temporal (e.g. parietal cortex) and spectral (e.g. inferior frontal gyrus and medial frontal gyrus) regularities [20-22]. Regarding distinct auditory processing and neural sources of temporal and spectral regularities, it can be assumed that different types of sound regularities may affect stream segregation differently. Therefore, in this study, we aimed to investigate whether auditory regularities facilitate stream segregation and whether distinct types of regularities differentially influence this perceptual process. We compared temporal and spectral sound regularities using an indirect measure of auditory streaming (by detecting the rare intensity level targets). 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 implementation of the task requires participants to segregate the foreground auditory stream from a temporally or spectrally regular repetition of the background stream, thereby detecting rare intensity level targets in the foreground irregular stream.

Methods

Participants

Twenty-five right-handed healthy young volunteers (20 females and 5 males) participated in this study. Participants' age ranged from 19 to 33 years (mean age=23.24±3.43). All participants had pure tone thresholds ≤25 dB HL at 250–8000 Hz, with no known neurological or psychiatric conditions and no history of head injury.

Stimuli and experimental design

Behavioral testing and Electroencephalogram (EEG) recordings were conducted simultaneously in an acoustically and electromagnetically sound-attenuated booth. During the recording, the participants were seated comfortably in a chair and asked to recognize the target stimuli. Sounds were generated using MATLAB 2016b 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 comprised temporal convergence between A and B tones similar to real-world auditory scenes (Figure 1).

Tones in each stream were pure tones with 75 ms in duration, including 10 ms rise/fall time, and were introduced binaurally via ER-3A insert earphones (ER-3A, Etymotic Co.). The A tones in 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 100–250 ms in 50-ms steps. Rare-level target tones that blend quasi-randomly in the A stream were introduced at a 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 1 dB steps. Indeed, for the participants to perceive 80 dB A sounds as targets, all 70 dB sounds of the A stream had to be separated from the intervening stimuli from the B stream, which had random intensity changes. The B tones had a mean frequency value

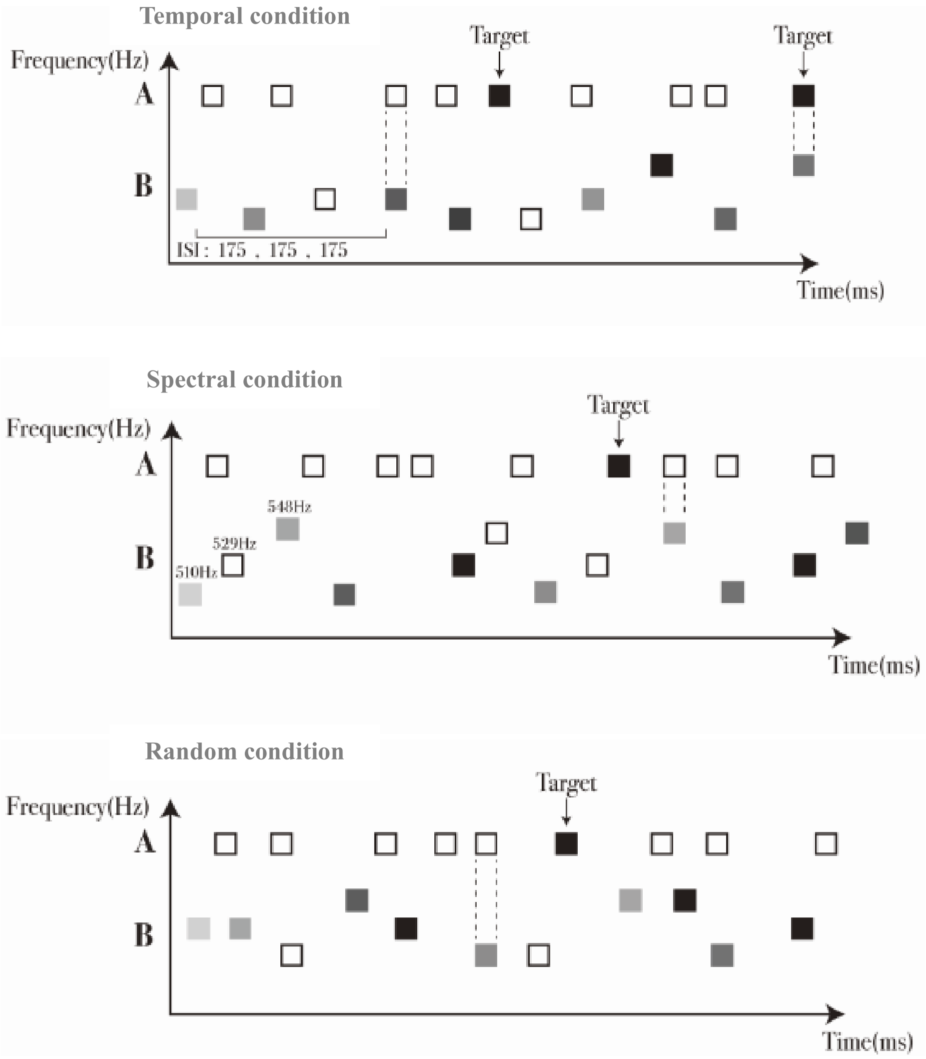

Figure 1. A schematic presentation of the experimental design. Stimuli were introduced in three distinct conditions: temporal regularity condition in which tones within the "B" stream had constant and regular inter-stimulus interval of 175 ms; spectral regularity condition in which tones within the "B" 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. ISI; inter-stimulus interval

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 a previous study [19], 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 under the temporal regularity condition had an irregular frequency pattern like B2B1B2B2B1B1B2B1B3, but had constant and regular ISI (175 ms). Under the spectral regularity condition, tones were presented in an ascending order (i.e. B1B2B3B1B2B3) but had randomly changed ISIs of 100–250 ms at 50-ms steps. Under the random experimental condition, tones in the background B stream were presented with no temporally or spectrally regular patterns.

A training block, encompassing all experimental conditions, was used before the main experiment to familiarize participants with the task. In this training block, the participants were asked to detect target stimuli in three experimental conditions. The occurrence of the auditory target was emphasized by displaying a green light at the center of the computer screen. Afterward, participants had to accomplish at least one training block of all experimental conditions without visual support. Participants received additional training blocks until they felt confident with the task. Participants received feedback regarding their execution following each training block as well as every block throughout the experiment. Subsequently, the main experiment was carried out. Each condition was introduced in five-minute blocks, each randomly repeated four times. In each block, target stimuli appeared 25–36 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, participants were given break times to alleviate fatigue. During the experiment, participants were not informed about regularity manipulation within the background stream. The entire experiment —including training, resting, behavioral testing and EEG recording — lasted approximately 4 hours.

Data analysis

To determine the participant's ability to detect the rare-level targets, the Hit Rate (HR) and False Alarm Rate (FAR) were computed for each experimental condition. All responses occurring in 300–1000 ms following the onset of the target were considered as hits (defined according to a pilot study) and others were considered as false alarms. Since the current stimulus framework requires the introduction of sustained stimuli so that all non-targets in both streams can appear throughout the response window, the number of false alarms is highly dependent on the response window employed in the analysis. We employed a procedure that takes into account the response window. A condition's total presentation time was divided into time bins with durations corresponding to the 300–900 ms response window (i.e. 600 ms). For each participant and condition, the HR was determined by dividing the whole duration of time bins where a hit response was expected (the number of hits by time bin) by the sum of the time bins where hits were feasible (the number of targets multiplied by time bin). Since the time bin variables in the numerator as well as denominator cancelled, the formula may be condensed to: HR=n hits/n targets. For each participant and condition, the FAR was calculated independently as the sum of time bins that had false alarms was divided by the sum of time bins where a false alarm was feasible, as shown below: [18].

$$FAR = \frac{\text{Number of false alarms} \times \text{time bin}}{\text{Overall presentation time} - (\text{number of targets} \times \text{time bin})}$$

Statistical analysis

All Statistical analyses were carried out in SPSS v.17 (SPSS Inc., Chicago, IL), considering a significance level set at 0.05. To assess the effects of regularity manipulation, repeated-measure ANOVA was used for each behavioral measure under experimental conditions (spectral, temporal, or random) as a factor, followed by post hoc comparisons using Bonferroni's test.

Results

Table 1 shows the mean and standard deviation of behavioral test parameters (HR and FAR) for all experimental conditions. The results of repeated-measures ANOVA showed the significant main effect of sound regularities on HR ($F_{(2,23)}$=16.740, p<0.001, η^2=0.411). The post hoc comparisons showed that the HR was significantly different between the temporal and spectral conditions (p=0.004), between the temporal and random conditions (p<0.001), and between the spectral and random conditions (p=0.010) (Figure 2). The repeated-measures ANOVA result showed no significant effect of sound regularities on FAR ($F_{(2,23)}$=1.188, p=0.311, η^2=0.047).

Table 1. 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

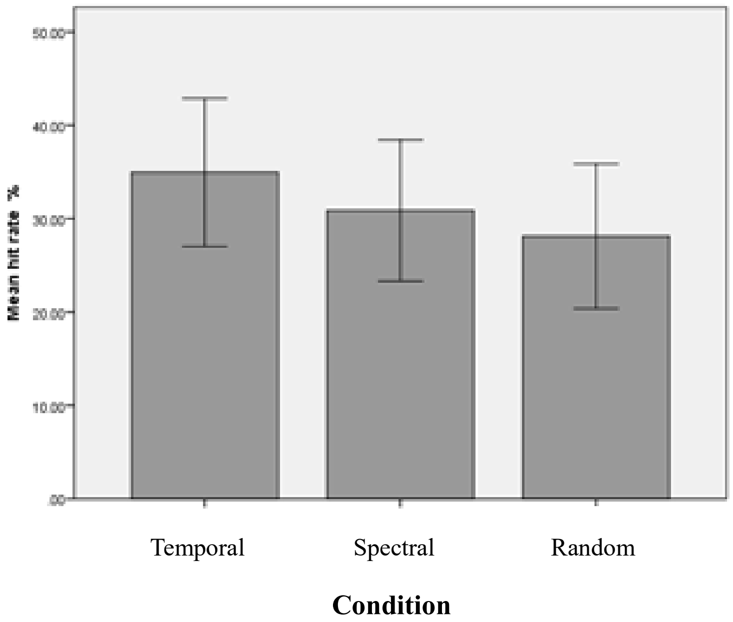

Figure 2. The impact of experimental conditions on the hit rate.

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 affects stream segregation, and 2) the auditory stream segregation is distinctively affected by the different types of regularly repeating sound features (i.e. spectral and temporal regularities). The results 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 (i.e. detecting rare-level targets), the participants were asked to 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 into a background auditory stream enhances target detection in a concurrent irregular stream, demonstrating that regularity influences stream segregation. The present study findings are in line with the findings of previous studies on the role of regularly repeating sound features in auditory stream segregation during a subjective reports task [16, 17]. For instance, Bendixen et al. examined the role of regularity on stream segregation using a bi-stable ABA-ABA streaming stimulus with or without regular patterns (frequency and/or intensity). They asked their listeners to indicate their perception of sound sequences constantly. They reported that independently introduced regular patterns in the A tones, B tones, or both stabilized the perception of two streams for longer durations [17]. Thus, their findings demonstrated that regular patterns increased the probability of perceiving segregated auditory streams. The effect of regularly repeating sound features on the stream segregation has also been found in objective listening measures of auditory stream segregation (under a within-stream deviant detection task). For instance, Andreou et al. studied the role of regularity in auditory stream segregation as a function of different frequency separations during a deviant detection task. They introduced temporal regularity in the background stream while tones in the foreground stream were arranged irregularly. Their main finding was that manipulating regularity throughout the background stream enhanced foreground stream segregation, primarily when frequency separations alone were insufficient [19]. Similarly, Rimmele et al. demonstrated that stream segregation depends on the auditory system's ability to detect regular patterns, whether these patterns occur in foreground or background streams during an intensity-deviation detection task [18].

The current study showed that temporal and spectral regularities in the background auditory streams had different effects on stream segregation, as measured by the HR of rare-level targets that occurred within an irregular foreground stream. Participants experienced greater regularity benefits for the stream segregation when temporal regular patterns were embedded in the background stream than when spectral regular patterns were embedded. It probably means that the benefit was related to the type of regularly repeating features embedded in the background stream. Empirical evidence increasingly suggests that listeners employ sophisticated mechanisms to track the rhythmic properties, including microtiming differences of acoustic sources. This enables the rapid formation of temporal expectancies, thereby optimizing behavioral responses [19]. On the other hand, previous studies have proposed that the distinctiveness of neural sources is associated with different functional roles [26, 27]. Therefore, the finding that temporal and spectral regularities differentially affect stream segregation can be explained by distinct neural mechanisms dedicated to encoding temporal [21] and spectral [20, 22] regularities. This result is against the finding of Bendixen et al. [17], showing that introducing regularly repeating frequency and intensity patterns into either one or both streams of bi-stable ABA streaming stimulus had equal effects on the ratio of segregated percept. The discrepancy can be explained by methodological differences, such as the difference in frequency separation between the A and B streams and the type of examined regularities. Moreover, in our 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 findings of Shalgi and Deouell [26], who used temporally non-converging regularly repeating ABA patterns, which were far from complex natural auditory scenes.

Conclusion

This is the first study to reveal that regularities carried out by distinct stimulus features probably result in differential modulations of auditory stream segregation. Our findings expand the prior restricted literature regarding the facilitating effects of regularly repeating acoustic cues in the segregation of auditory streams. Gaining insights into how sound regularities influence auditory perception may provide valuable information for upcoming studies on music and speech auditory processing.

Ethical Considerations

Compliance with ethical guidelines

This study was conducted in the Rehabilitation School of Iran University of Medical Sciences (Ethical approval code: IR.IUMS.REC.1397.1027).

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' 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.

Conflict of interest

The authors declare that they have no competing interest.

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.

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, Gyimesi K, Winkler I. Regular patterns stabilize auditory streams. J Acoust Soc Am. 2010;128(6):3658-66. [DOI:10.1121/1.3500695]

18. 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]

19. 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]

20. 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]

21. Sohoglu E, Chait M. Detecting and representing predictable structure during auditory scene analysis. Elife. 2016;5:e19113. [DOI:10.7554/eLife.19113]

22. 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]

23. 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]

24. 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]

25. 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]

26. 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]

27. 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]