Research Article

Auditory Short-Term Memory Evaluation in Noise in Musicians

Büşra Nur Eser, MSc[1] Mustafa Bülent Şerbetçioğlu, MD, PhD[1]

[1] Department of Audiology, Graduate School of Health Sciences, Istanbul Medipol University, Istanbul, Turkey

Address for correspondence Büşra Nur Eser, MSc, bneser@medipol.edu.tr

J Am Acad Audiol

Abstract

Background Working memory, a short-term memory component, is a multicomponent system that manages attention and short-term memory in speech perception in challenging listening conditions. These challenging conditions cause listening effort that can be objectively evaluated by pupillometry. Studies show that auditory working memory is more developed in musicians for complex auditory tasks.

Purpose This study aims to compare the listening effort and short-term memory in noise between musicians and nonmusicians.

Research Design An experimental research design was adopted for the study.

Study Sample The study was conducted on 22 musicians and 20 nonmusicians between the ages of 20 and 45.

Data Collection and Analysis Participants' effort analysis was measured with pupillometry; performance analysis was measured with short-term memory score by listening to the 15 word lists of Verbal Memory Processes Test. Participants are tested under three conditions: quiet, +15 signal-to-noise ratio (SNR), and +5 SNR.

Results While nonmusicians showed significantly higher short-term memory score (STMS) than musicians in the quiet condition, musicians' STMS were significantly higher in both noise conditions (+15 SNR and +5 SNR). The nonmusician's percentage of pupil growth averages were higher than the musicians for three conditions.

Conclusion As a result, musicians had better memory performance in noise and less effort in the listening task according to lower pupil growth. This study objectively evaluated the differences between participants' listening efforts by pupillometry. It is also observed that the SNR and music training affect memory performance.

Keywords
- working memory
- memory
- short-term
- music
- pupil
- noise

Listening and understanding speech are extraordinarily complex tasks involving a wide variety of sensory and cognitive processes. Cognitive processes include mental functions such as attention, memory, understanding, learning, evaluation, problem solving, and decision making. Cognitive abilities provide short- or long-term storage in our memory by processing auditory signals in daily life.[1] One of the critical factors that enable this process is auditory working memory. Working memory, a short-term memory component, is a multicomponent system that manages attention and short-term memory in speech perception in challenging listening conditions.[2,3] The capacity of the working memory represents the amount of information a person can remember, participate in, and store in a quickly accessible state at a time.[4] Listening in noise increases the effort spent and makes it difficult to keep said information in memory.[5]

Listening effort refers to the cognitive effort invested by the listener to understand the auditory signal. This effort can be objectively monitored by pupil size as a response to

received
October 22, 2021
accepted after revision
July 5, 2022

Thieme Medical Publishers, Inc.,
333 Seventh Avenue, 18th Floor,
New York, NY 10001, USA

DOI https://doi.org/10.1055/a-1896-5129.
ISSN 1050-0545.

allocated cognitive resources.[6] Pupillary responses are used widely to measure listening effort for speech processing in noise.[7]

In studies conducted by presenting auditory stimuli to the participants, it has been shown that the participants' pupil sizes vary with the change of cognitive resources. Significant phasic changes in pupil diameter are observed during listening in accordance with the difficulty of the auditory task.[8] While the mean pupillary growth reflects the average processing load of a particular process, the peak pupillary growth rate reflects the maximum processing load.[9] In our study, we compared the cognitive effort spent by calculating the peak pupil growth amount as a percentage (PGP).

Zatorre et al has shown that playing a musical instrument is a very complex task.[10] Because of this complex interaction, the music effect has been studied in recent years with different models when examining the functional organization of the brain and brain plasticity. Studies have shown that receiving music education has significant effects on the functional and structural plasticity of the human brain.[11] The study results of Kraus et al, evaluating the hearing performance of individuals with and without musical education, showed that musical experience improves the ability to distinguish speech in challenging listening environments. It has been shown that working memory is a part of this ability. Besides, a relationship is found between working memory ability and years of musical practice.[12]

This study aimed to compare the listening effort and auditory short-term memory in noise in musicians and nonmusicians. Participants' effort was evaluated using both pupillometry and the Öktem Verbal Memory Processes Test (VMPT) word scores. VMPT is part of the Neuropsychological Test Battery, which was developed by Öktem based on the Rey Auditory Verbal Learning Test. The validity and reliability study was also done by Öktem. Öktem-VMPT is a test developed for multifactorial investigation of verbal learning and memory. A list of 15 words is read to the participant with an unstressed tone of voice, leaving 1 second between each word.[13] In this study, we applied the first phase of the VMPT which assesses verbal working memory. It was aimed to show whether music education affects auditory memory function in noise.

Material and Methods

Participants

There are two groups, musicians and nonmusicians in the study. The study was approved by the Istanbul Medipol University Ethics Committee with the number of 10840098-604.01.01-E.1422. Twenty-two musicians who professionally played at least one instrument in an orchestra, and 20 nonmusicians, who had not played any instrument before, participated in the study. All nonmusicians were graduated from at least associates or bachelor's degree. The ages of individuals ranged between 20 and 45 years (mean age 26 ± 5 years). Musicians were all graduated from conservatory or students at conservatory where they have been selected to attend via a musical talent examination. The selected musicians had begun their music training before the age of 18 and they have been practicing for minimum 3 days a week in the last 5 years. Musicians' (20–40 years) average age was 24.63. Their musical experience years' average is 11.72 ± 5.53 (min 6, max 20 years).

Pure-tone audiometry average scores of 500, 1,000, 2,000, and 4,000 Hz being higher than 25 dB, with no history of otological or neurological disorders, and speech discrimination scores over 88% were accepted as audiological inclusion criteria. The study was conducted with the understanding and full informed consent of the participants.

Word Lists and Sound Recordings

Word lists were determined as the A, B, and C lists of the VMPT consisting of 15 words each, which is used for memory assessment in the Neuropsychological Test Battery. The words that individuals were able to recall during the test were tracked and marked on a wordlist, which than generates the VMPT short-term memory scores. This score is specified as the "short-term memory score (STMS)" in the VMPT.

Sound recordings were created for a balanced signal-to-noise ratio (SNR) and a more homogeneous presentation. Each word was read unstressed by a female speaker, 1 second in between. Recordings were taken in a silent cabin with the Shure SV100 dynamic microphone and Tascam DR-05 voice recorder. The noise added to the back of the recordings and the SNR adjustments were made using the Audacity program. Recording parameters were 44,100 sampling rates with 16-bit resolution. Babble noise added to the background was developed by Moore et al using speech mask sounds.[14]

First, each word in the main voice recording was drawn to an equal dB level. Then the background noise originating from the device behind the words was cleaned. Words were placed on the babble noise creating a SNR of +5 and +15 dB. Since baseline recording could be made in each sound recording, words were placed 5 seconds later.

Before the presentation of the stimuli, the level was calibrated at the participants' ear level. Measurements were made with a Class I PCE 430 Sound Level Meter in the free field with a dB A filter at 1,000 Hz at 0° azimuth, at a distance of 1 m from the listener's ear level. The comfortable hearing signal strength level was set at 62 dB sound pressure level as the American National Standards Institute recommended.[15]

Memory Task and Pupillometry

Forty-two participants have listened to three different word lists in quiet condition and two different background noises: +15 SNR and +5 SNR. Participants were asked to listen to these words, keep them in their memory, and repeat them with the cue. Word scores that the participants repeat during recall were marked on the word lists. The background sound levels were randomly presented to the listeners to prevent the learning effect and fatigue. After the completion of each list, the participants were given a 5-minute resting period.

During listening and recall, pupil size changings were recorded with a pupillometer (In the **►Video 1**, one of the pupillometry recordings during the listening task is shown). Pupillometry was performed with Micromedical

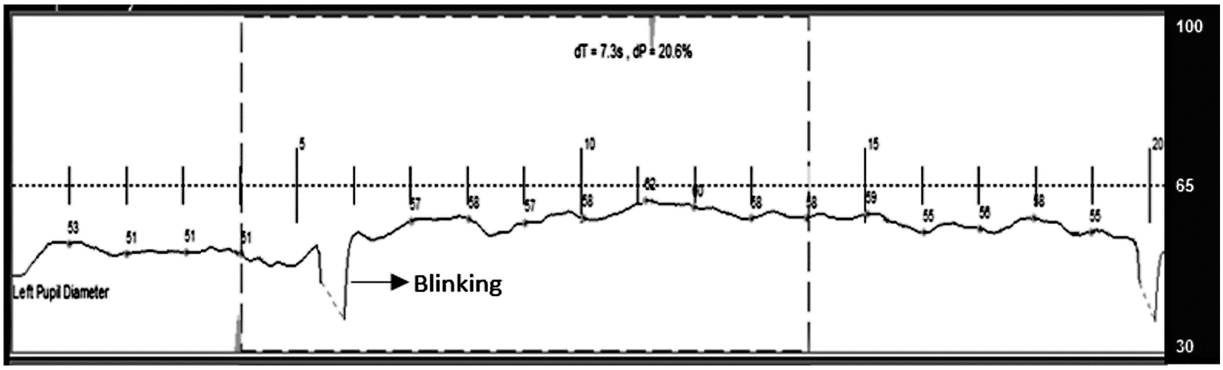

Fig. 1 Pupillometry recording of basal and listening period (dT, time difference; dP, pupil growth percentage).

videonystagmography (Chatham, IL). Participants were seated 1 m apart from the sound system in a quiet room. The Micromedical device's goggles were placed on the participants' head with a reference at eye level to prevent the participants' eyes from wandering. Participants were asked to look at this reference point without blinking as much as possible during the test. To ensure that pupil sizes are not affected by ambient conditions, during recording there was no change (such as light, noise) that would affect the test environment and the participant.

After the pupils' position and brightness were adjusted, a baseline recording was performed for each condition for the first 5 seconds. Then, while the participant was listening to the audio stimulus, a recording was taken for 15 seconds. Afterwards, during the recall process the recording continued between 10 and 20 seconds depending on the participant's performance. The more the participant blinked, the acceptance rate was lower during recordings. No mark was placed 50 ms before and 150 ms after blinking (►**Fig. 1**). PGP was calculated automatically by the device's algorithm with the marks placed on the basal pupil size and maximum size peak while listening. PGP in the device software is calculated with "dP = (Final Dimension – Initial Dimension/Initial Dimension) * 100."

Video 1

Pupillometry recording during the listening task. Online content including video sequences viewable at: https://www.thieme-connect.com/products/ejournals/html/10.1055/a-1896-5129.

Statistical Analysis

Statistical Package for the Social Sciences version 24.0 (SPSS Inc, Chicago, IL) was used for statistical analysis of the data in the study. The statistical significance level was accepted as $p < 0.05$. The distribution was evaluated with the Kolmogorov–Smirnov test. As the dependent variables did not show normal distribution, nonparametric tests were used for statistical analysis. Friedman's two-way analysis of variance was used to compare within-group situations. Bonferroni's correction was made to determine which situations differ meaningfully ($p < 0.016$ was considered significant). Mann–Whitney *U* test was used to compare the data between groups, and $p < 0.05$ was accepted as significant.

Results

The STMS while the VMPT words presented in three different conditions were compared for 20 nonmusicians and 22 musicians (►**Tables 1** and **2**). The word scores of musicians in the quiet condition were significantly higher than nonmusicians' scores ($p = 0.001$). However, the musicians' word

Table 1 Comparison of short-term memory scores of nonmusicians and musicians between groups[a]

	Nonmusicians	Musicians	
	Mean ± SD	Mean ± SD	*p*
STMS1	8.40 ± 1.63	6.63 ± 1.25	0.001[b]
STMS2	6.95 ± 1.76	8.63 ± 1.43	0.002[b]
STMS3	6.05 ± 1.27	7.13 ± 1.35	0.022[b]

Abbreviations: SD, standard deviation; SNR, signal-to-noise ratio; STMS, short-term memory score; STMS1, quiet; STMS2, +15 SNR; STMS3, +5 SNR;.

[a] Mann–Whitney *U* test was used.

[b] $p < 0.05$.

Table 2 Comparison of in-group short-term memory scores of nonmusicians and musicians in different noise situations

	Nonmusicians		Musicians	
	p^1	p^2	p^1	p^2
STMS1–STMS2	0.000[a]	0.002[b]	0.000[a]	
				0.000[b]
STMS1–STMS3		0.000[b]		0.600
STMS2–STMS3		0.618		0.001[b]

Abbreviations: SNR, signal-to-noise ratio; STMS, short-term memory score; STMS1, quiet; STMS2, +15 SNR; STMS3, +5 SNR.

Note: p^1: Friedman's test; p^2: Bonferroni's correction.

[a] $p < 0.05$.

[b] $p < 0.016$.

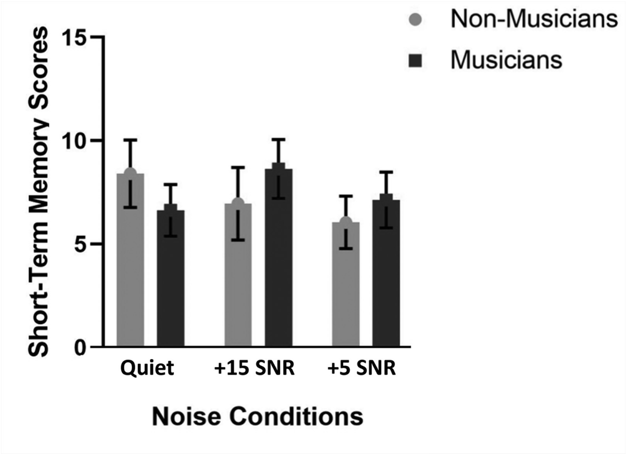

Fig. 2 Comparison of short-term memory scores of nonmusicians and musicians.

scores were significantly better at +15 SNR ($p = 0.002$) and +5 SNR ($p = 0.022$).

When the STMS obtained in various noise conditions were compared, a significant result was obtained for both groups in Friedman's two-way analysis ($p1 = 0.000$, $p1 = 0.000$) (►**Fig. 2**). In the nonmusician group, word scores were obtained significantly higher in quiet condition than +15 SNR and +5 SNR ($p2 = 0.002$, $p2 = 0.000$). In the musician group, STMS at +15 SNR were significantly higher than quiet condition ($p2 = 0.000$) and +5 SNR ($p2 = 0.022$) ($p < 0.016$).

While analyzing the pupil data, since more than 30% of the recording is distorted by blinking during listening and more than 50% during recall, three female participants and one male participant from the musician group and two female participants from the nonmusician group were excluded from the evaluation. Thus, 18 nonmusicians and 18 musicians were included in the pupil evaluation.

PGPs were compared first between all groups and then with one another for all three conditions. Comparison results between groups are presented in ►**Table 3**. PGP was observed statistically higher in the quiet condition for nonmusicians in comparison to musicians ($p = 0.027$). The comparison results of the groups are given in ►**Table 4**. There was a significant difference for both groups according to Friedman's two-way analysis ($p1 = 0.001$, $p1 = 0.001$). In paired comparisons with the Bonferroni correction, the PGP in quiet condition is significantly higher than the +15 SNR ($p2 = 0.005$) and +5 SNR ($p2 = 0.003$) for the nonmusician group. In the musician group, the PGP in the quiet condition is significantly greater than the +5 SNR ($p2 = 0.000$) ($p < 0.016$) (►**Fig. 3**).

Table 3 Comparison of pupil growth percentages (PGP) between groups at different noise levels[a]

	Nonmusicians	Musicians	
	Mean ± SD	Mean ± SD	p
PGP1	18.11 ± 2.43	15.79 ± 2.92	0.027[b]
PGP2	14.18 ± 3.55	12.75 ± 3.02	0.141
PGP3	13.72 ± 3.57	11.57 ± 2.75	0.066

Abbreviations: PGP, pupil growth percentages; PGP1, quiet; PGP2, +15 SNR; PGP3, +5 SNR; SD, standard deviation.
[a]Mann–Whitney *U* test was used.
[b]$p < 0.05$.

Table 4 Comparison of pupil growth percentages (PGP) within groups at different noise levels

	Nonmusicians		Musicians	
	p^1	p^2	p^1	p^2
PGP1-PGP2	0.001[a]	0.005[b]	0.001[a]	 0.091
PGP1-PGP3		0.003[b]		0.000[b]
PGP2-PGP3		1		0.287

Abbreviations: PGP, pupil growth percentages; PGP1, quiet; PGP2, +15 SNR; PGP3, +5 SNR; SNR, signal-to-noise ratio.
Note: p^1: Friedman's test; p^2: Bonferroni's correction.
[a]$p < 0.05$.
[b]$p < 0.016$.

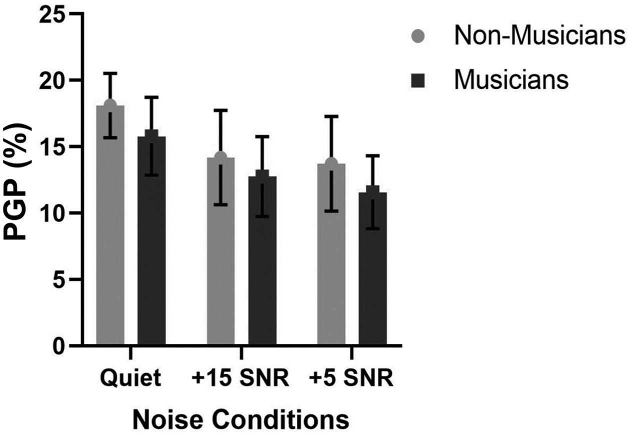

Fig. 3 Comparison of pupil growth percentage (PGP) of musicians and nonmusicians at different noise levels.

Discussion

Cognitive resources separate the target signal from background noise, which causes a listening effort—leaving fewer resources to store and process messages content in working memory.[16–19] According to Schellenberg and Peretz, musicians' enhanced speech perception in noise may depend on their general better cognitive abilities.[20] However, there are many studies which show the advantage of being a musician for short-term, long-term, and working memory.[21]

The first hypothesis was that STMS would decrease with increasing noise. Second, musicians would be less affected by noise than nonmusicians. According to the results reported by Rönnberg et al, when noise is speech-like (such as babble noise), listening effort increases due to increased demands on cognitive resources. Noise had a similar effect for our nonmusicians' results.[22] Due to the increased babble noise, nonmusicians' memory scores were decreased, and PGP showed higher listening effort.

When we compared only the musicians' word scores in our study, we found a significantly higher word score at +15 SNR than the other two conditions. Nonmusicians scored significantly better in the quiet condition, while musicians were significantly more successful in noisy situations. Rönnberg et al's study has shown that working memory capacity is essential for understanding speech in reduced SNRs.[22]

Musicians have better memory scores in noise, and this could be explained by the effect of improved working memory capacity on understanding speech in noise by looking at previous studies.[23–25] Parbery-Clark et al, evaluated the QuickSIN scores of musicians and nonmusicians and showed that working memory affected the scores, and musicians' results were significantly better in the SIN (Speech in Noise) and QuickSIN tests.[26]

On the other hand, Escobar et al evaluated speech understanding in noise and auditory memory in musicians and nonmusicians. They stated no significant difference in musicians' listening efforts and working memory capacities than those not. Differently, they have grouped participants according to their high and low working memory capacities. Results showed that musicians with high working memory capacities had better memory results in noise than nonmusicians.[1] In our study, participants' working memory capacity had not been measured before the evaluation. If we had separated the participants according to their working memory capacity, we could more reliably demonstrate the effect of higher scores of musicians in noise on working memory capacity.

One of our study hypotheses was to observe higher effort/greater PGP in increased noise, as PGPs are associated with listening effort. Musicians showed lower PGP/less effort than nonmusicians. According to our results, when the percentage of pupil growth from the baseline to the maximum peak is calculated, both groups' highest PGP rates were in quiet conditions. The reason for this is thought to be the difference of the starting listening efforts caused by the noise. When the participants are presented with noise before the word lists are presented, it causes a listening effort to form. Therefore, when the word list is presented, the change in PGP is relatively less compared to the PGP change that happens in quiet conditions. Thus, it is essential to consider the noise conditions when evaluating the listening effort.

Wendt et al performed speech reception threshold in 8 noise situations based on –20, –16, –12, –8, –4, 0, 4, and 8 dB SNR and recorded pupil size. Results showed pupil sizes growing with SNR decreasing down to –8 SNR.[7] However, when the task becomes more difficult than what the participant can do, the effort decreases, and the pupil dilation stops due to the decrease in the participant's motivation. In our study, it was observed that pupil sizes increased during listening tasks with decreasing SNR. However, since we reduced the SNR to +5 most in the memory evaluation, no measurement was made in the noise level that would break the motivation as in the case study. In other words, it was not observed which SNR condition pupil enlargement stopped.

When we compared the PGP between musicians and nonmusicians, the PGP of nonmusicians was higher than musicians for all conditions. While the PGP were statistically higher in the quiet condition, it could be said that the PGP are higher in the other two cases by looking at the mean values. When this situation is associated with listening effort, it has been observed that nonmusicians spend more effort on memory tasks in noise than musicians. According to the study of Escobar et al[1] mentioned before, contrary to our hypothesis, music education does not affect listening effort. On the other hand, they stated in the study that nonmusicians with low working memory capacity spent significantly more effort on listening effort measurements.

Demanding tasks, attention processes, motivation, and fatigue are related to the individual's use of cognitive resources and thus pupil size. Each of these various factors can affect different types of processing and thus cognitive resource use.[6] These factors should be considered as weaknesses of pupillometry. Considering these weaknesses, we tried to obtain consistent results using homogeneous participant groups and keeping stable environment conditions. To avoid the learning effect and fatigue, we presented the SNR conditions within intervals in random order.

Musicians' better cognitive abilities have been reported by many studies.[12,20,21,27,28] Nevertheless, there is still a contradiction about whether musical education improves cognition or whether individuals who are successful as musicians already have higher cognitive abilities.[1] We do not think that we have revealed this aspect in our study. On the other hand, our results showed that musicians performed better on memory tasks in noise, consistent with the literature. There may be a need for further study to examine the direction of this relationship with the same people, with detailed tests for pitch perception, auditory sequencing, attention, and memory.

In quiet conditions, nonmusicians have better scores than musicians in our study. Although we did not find evidence in the literature, musicians state that they cannot focus without the slight impact of background noise. However, we aimed to evaluate the difference between musicians' and nonmusicians' noise condition performance and the difference between the noise conditions for each group. Speech discrimination scores were obtained when including subjects, but speech discrimination in "noise" was not evaluated. Not evaluating the discrimination in noise's effect on memory performances is one of the study's limitations. If all participants were divided into groups by working memory assessment, more significant results would have been obtained.

Conclusion

This study aimed to compare the listening effort and memory performance in musicians and nonmusicians in three different signal noise situations. According to our results, musicians' short-term memory performance in noisy conditions was significantly more successful. When listening effort was evaluated with the pupillary growth, it is seen that nonmusicians spent higher effort than musicians. The high individual differences in pupillometer measurements and many parameters measured made it challenging to demonstrate the effect of being a musician on working memory. Further studies are needed to carry out more specific measurements for working memory with a larger musician sample.

Acknowledgment

We would like to thank musicians and nonmusicians for patience and invaluable contribution for this study. No external funding was provided for this study.

Conflict of Interest

None declared.

References

1 Escobar J, Mussoi BS, Silberer AB. The effect of musical training and working memory in adverse listening situations. Ear Hear 2020;41(02):278–288

2 Cowan N. What are the differences between long-term, short-term, and working memory? Prog Brain Res 2008;169:323–338

3 Rönnberg J, Rudner M, Foo C, Lunner T. Cognition counts: a working memory system for ease of language understanding (ELU). Int J Audiol 2008;47(Suppl 2):S99–S105

4 Baddeley A. Working memory: looking back and looking forward. Nat Rev Neurosci 2003;4(10):829–839

5 Marrone N, Alt M, DeDe G, Olson S, Shehorn J. Effects of steady-state noise on verbal working memory in young adults. J Speech Lang Hear Res 2015;58(06):1793–1804

6 Pichora-Fuller MK, Kramer SE, Eckert MA, et al. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). Ear Hear 2016;37(Suppl 1):5S–27S

7 Wendt D, Hietkamp RK, Lunner T. Impact of noise and noise reduction on processing effort: a pupillometry study. Ear Hear 2017;38(06):690–700

8 Winn MB, Wendt D, Koelewijn T, Kuchinsky SE. Best practices and advice for using pupillometry to measure listening effort: an introduction for those who want to get started. Trends Hear 2018;22:2331216518800869

9 Zekveld AA, Kramer SE, Festen JM. Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. Ear Hear 2011;32(04):498–510

10 Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music perception and production. Nat Rev Neurosci 2007;8(07):547–558

11 Pantev C, Okamoto H, Teismann H. Music-induced cortical plasticity and lateral inhibition in the human auditory cortex as foundations for tonal tinnitus treatment. Front Syst Neurosci 2012;6:50

12 Kraus N, Strait DL, Parbery-Clark A. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. Ann N Y Acad Sci 2012;1252:100–107

13 Oktem O. A verbal test of memory processes: a preliminary study. Archives of Neuropsychiatry. 1992;29(04):196–206(Turkish)

14 Moore DR, Whiston H, Lough M, et al. FreeHear: a new sound-field speech-in-babble hearing assessment tool. Trends Hear 2019; 23:2331216519872378

15 ANSI – American National Standards Institute. American National Standard Methods for the calculation of the speech intelligibility index. New York, NY: American National Standards Institute; 1997

16 Eckert MA, Cute SL, Vaden KI Jr, Kuchinsky SE, Dubno JR. Auditory cortex signs of age-related hearing loss. J Assoc Res Otolaryngol 2012;13(05):703–713

17 Piquado T, Isaacowitz D, Wingfield A. Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology 2010;47(03):560–569

18 Winn MB, Edwards JR, Litovsky RY. The impact of auditory spectral resolution on listening effort revealed by pupil dilation. Ear Hear 2015;36(04):e153–e165

19 Zekveld AA, Koelewijn T, Kramer SE. The pupil dilation response to auditory stimuli: current state of knowledge. Trends Hear 2018; 22:2331216518777174

20 Schellenberg EG, Peretz I. Music, language and cognition: unresolved issues. Trends Cogn Sci 2008;12(02):45–46

21 Talamini F, Altoè G, Carretti B, Grassi M. Musicians have better memory than nonmusicians: a meta-analysis. PLoS One 2017;12 (10):e0186773[published correction appears in PLoS One. 2018 Jan 19;13(1):e0191776]

22 Rönnberg N, Rudner M, Lunner T, Stenfelt S. Memory performance on the Auditory Inference Span Test is independent of background noise type for young adults with normal hearing at high speech intelligibility. Front Psychol 2014;5:1490

23 Ingvalson EM, Dhar S, Wong PC, Liu H. Working memory training to improve speech perception in noise across languages. J Acoust Soc Am 2015;137(06):3477–3486

24 George EM, Coch D. Music training and working memory: an ERP study. Neuropsychologia 2011;49(05):1083–1094

25 Yeend I, Beach EF, Sharma M, Dillon H. The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. Hear Res 2017;353:224–236

26 Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear Hear 2009;30(06):653–661

27 Strait DL, Kraus N, Parbery-Clark A, Ashley R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. Hear Res 2010;261(1-2):22–29

28 Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci U S A 2007;104(40):15894–15898