Krysko M.D.[1], Vartanov A.V.[1], Bronov O.Y.[2] Verbal component suppression during internal representation of songs: fMRI-study

Крысько М.Д., Вартанов А.В., Бронов О.Ю. Подавление вербального компонента внутренней репрезентации песни: фМРТ исследование

[1]Lomonosov Moscow State University

[2]Federal State Budgetary Institution National Medical and Surgical Center named after N.I. Pirogov

This paper investigates one of the processes of internal representations—specifically, the mental representation of melodies with suppressed verbalization. The objective of this study is to explore the involvement of brain structures during the mental representation of melodies. An fMRI study was conducted with 33 healthy, right-handed participants. Participants were tasked with listening to musical passages accompanied by lyrics in their original form and then mentally reproducing these excerpts with and without the lyrics. The findings suggest that the encoding of musical patterns may exhibit individual variations; however, general trends were identified. The results demonstrate that both cortical and subcortical structures contribute to the internal representation of melodies. Specifically, Wernicke's area and its right hemisphere homologue are instrumental in internal representations, while the caudate nucleus, cingulate gyrus, and superior temporal gyrus play key roles in the inhibition of the verbal component.

Keywords: music, fMRI, internal representation, inner speech, imagination, musical images, lyrics suppression.

Introduction

Currently, a large amount of research is devoted to the study of internal speech processes, as well as internal representations and images, their brain mechanisms and methods of their decoding.

The results of research aimed at studying inner speech processes opens up a large field of possibilities, including the creation and improvement of brain-computer interface (BCI) architectures, which have a wide range of functionality and can be applied to solve various tasks from therapeutic [Kopp, 2013] and compensating for various pathologies [Bocquelet, 2016] to optimising and entertaining [Minguillon et al., 2017].

Currently, a large amount of research is aimed at understanding the mechanisms of semantic features of speech, while less attention is being paid to the study of non-semantic structures. Non-semantic structures of speech can include elements that humans can use to represent music – to play or create it. In this context, the data can be used to create a BCI program capable of decoding and representing an imaginary melody as audio tracks or recognising it by comparing it with a downloaded archive. This research takes the first steps in this direction.

Inner speech operates with symbols and meanings, but a thought is not tied to a specific form of expression, which allows the same meaning to be conveyed by both verbal and gestural, spatial, metaphorical, including musical means, but at the same time music itself does not necessarily convey meaning [Sachs, 2007]. It can convey a sensation (visual, tactile and other modalities), an emotion, a state, a mood or even an image. The totality of musical images has as much range and variety as the totality of visual images [Sachs, 2007]. Recent research on musical imagery has now led to the conclusion that the brain mechanisms of musical imagination and music perception are similar, as evidenced by numerous data [Zatorre, 1996], [Herholz, Halpern et al., 2012], [Regev, Halpern et al, 2021]. The role of the auditory cortex [Zatorre, 2005], [Kraemer, 2005], including the right associative cortex [Regev, Halpern et al., 2021], motor cortex including Broca's area, frontal cortex including frontal-limbic structures, parietal sulcus, and other sensory areas has been shown. Nevertheless, in the process of musical imagination, a more extended network was recorded compared to the perceptual process, which included prefrontal cortex, cerebellum, intraparietal sulcus, as well as increased activation of the functional connectivity of the anterior right temporal cortex with frontal areas [Herholz, Halpern et al., 2012].

Musical images are complex, they appear due to the integration of various aspects such as intonation, tempo, timbre, rhythm, as well as verbal, emotional, kinetic, components, such as the image of playing a musical instrument [Zatorre, 2005; Pasely, 2012], so it is possible to observe the participation in the process of musical imagination of such structures as frontal-limbic, motor [Pascual-Leone, A. et al., 2005], as well as the parietal sulcus in the case of conscious transformation of the melodic motif [Foster et al., 2013]. Recently, experimental evidence that Broca's area and its right analogue mediate musical syntax has been obtained using MEG and fMRI in a music perception paradigm [Maess, Koelsch, Guenther, Friederici, 2001]; [Kuhnert, Willems, Casasanto, Patel, Hagoort, Berwick, 2015]; [Chaing et al., 2018].

It is important to consider the fact that there is still debate about the extent to which the processing of music, speech and other natural sounds depends on general or specific neural mechanisms [Peretz et al., 2015, Zatorre et al., 2002] and the extent to which these mechanisms are hierarchically organised [Chevillet et al., 2011; Hickok and Poppel, 2007; Staeren et al., 2009].

A musical image can be accompanied by song lyrics, and thus it is highly important to consider how pure musical images differ from complex ones that include a musical-verbal component. When linking melody and lyrics, there is evidence for activation in the inferior frontal gyrus (IFG) of the left hemisphere [Alonso, 2016; Callan et al., 2006], the middle temporal gyrus (MTG) of both hemispheres [Alonso, 2016], as well as the motor cortex of the left hemisphere [Alonso, 2016; Gabriel et al., 2016] and visual areas [Alonso, 2016]. When melody and lyrics were presented separately, a larger neural network was involved, and changes in activation were observed in the following structures: the basal ganglia and cerebellum mediating the correct order of musical phrases, the hippocampus in the right hemisphere, and the caudate nucleus and motor cortex in the left hemisphere [Alonso, 2016]. Also, different neural networks have been found to imagine verbalised and non-verbalised melodies: activation in the fusiform and inferior occipital gyrus of the left hemisphere during verbal processing, bilateral activation in the temporo-occipital cortex and in the right middle temporal sulcus during melody processing, and the left posterior inferior temporal cortex has a specific role in these processes to facilitate song recognition by integrating lyrics and melody [Saito et al., 2012].

When speaking about the verbal component of the musical image, it is also important to distinguish it from the component of direct speech, the content of which is not associated with the verbalised melody. By non-verbalised musical images, we mean the images of melody, which the subject imagines, focusing on the rhythm, tonality, melody, intonation of a musical fragment, and as verbalised ones we understand the same melodies, but supplemented with original texts (verbal component).

In the present study, we questioned the extent to which lyrics might be related to a melody – whether the image of a song with lyrics is encoded as a single image or as integratable but distinct images. We found additional brain activity when mentally reconstructing a song image with the verbal component excluded, which occurs when the original musical fragment is strongly associated with a particular verbal component, in our case, the lyrics of well-known children's songs popular during the subjects' childhood. In the present study, stimuli that can be called associative rather than semantic were chosen. Stimuli of this type have been used in studies of people with aphasia who cannot use speech but are able to hum melodies along with lyrics [Schlaug et al., 2008], since the pronunciation of words in this image is automatic.

The encoding of melodic and speech stimuli is determined by different mechanisms when the verbal part does not match the melodic, rhythmic and affective characteristics of the melody [Bonomo, 2022], but if they do, it remains an open question whether the same network provides encoding of a complex image [Hamilton, 2022].

In the course of this study, we abandoned the classical block paradigm of fMRI studies in favour of a newer paradigm. This paradigm consists of continuous scanning throughout the entire task and further comparing the maps, averaged from one task to another. During instruction delivery, the scanner only pauses scanning, allowing it to record all conditions with the same settings. This design organisation has several advantages: the duration of the experiment is significantly reduced, only the process of task execution is recorded (the initialisation process on each test is excluded from the design). When comparing large blocks, we have the possibility to compare several different tasks with each other, without the need to pre-record them with repetitions.

In the present study, we proceeded from the following hypotheses:

1) The automatic encoding of the song image will include a verbal component that is absent in the presented stimulation but emerges during listening and free representation of the melody, which may be captured by the presence of areas associated with lyrics during the free representation stage: left

fusiform gyrus, inferior occipital gyrus, inferior frontal gyrus, middle temporal gyrus, motor or visual cortex.

2) The additional activity in isolating the melodic component of the song will be associated with the conscious suppression of the verbal component, which can be recorded as decreased activation in structures associated with lyrics or increased activation of structures associated with pure melody: bilateral temporo-occipital cortex, right middle temporal sulcus.

3) The processes of internal representation, regardless of the presence of the text, will be associated with changes in the activation of such areas as basal ganglia, cerebellum, caudate nucleus, left lateral motor cortex

To test these hypotheses, we conducted three series of conditions to exclude activity components occurring during listening and during free melody reconstruction: with and without lyrics.

The purpose of this study was to identify specific brain activity during mental free presentation of melodies and when only the nonverbal component of a song was isolated.

Methods

Participants

The study involved 33 healthy right-handed people (10 men and 23 women ranging from 20 to 30 years old, mean age = 24 years old, σ = 2.99). All participants had no neurological or mental disorders, which was confirmed by a preliminary survey. None of them were professional musicians but they all had some experience of playing different musical. Due to the presence of a large amount of artifacts, the study took into account the data of only 30 subjects.

Stimuli

Recordings of six melodies without lyrics were used as stimuli. The duration of each stimulus was four seconds. Since all the subjects are representatives of the Russian culture, the simplest harmonic melodies of famous children's songs were taken. The selection of melodies was based on the assumption that the selected fragments would be familiar to all the subjects, since they were brought up in a Russian-speaking environment in the 2000s. Moreover, before a series of experiments, a survey regarding recognition of the selected melodies and their words was conducting. Subjects were asked whether they were familiar with the presented melodies and whether they remembered the first lines of the lyrics of these songs. The following melodies were chosen for the experiment: “В лесу родилась елочка”, “Во саду ли в огороде”, “Маленькой елочке холодно зимой”, “Катюша”, “Калинка-малинка”, “Во поле березка стояла”. The melodies (lyrics have been removed) of the first lines of the first verses of each original song were recorded using sound recording programs and an electronic musical instrument (electric guitar). After the end of the experiments, the subjects were interviewed about how they performed the task of singing the melody. Due to the post-experimental interview, it was stated that the subjects sang the specific words for each melody and were familiar with them.

Procedure

The procedure included four conditions. Before the beginning of each stage, subjects listened to pre-recorded instructions. Three conditions were conducted: simple listening to the stimulus material, free representation of the melody silently in the mind, and representation of the pure melody without a lyrics silently in the mind. All three conditions were presented one after another; during the presentation of the instruction for each subsequent condition, scanner recording was paused and continued from the start of the first stimulus of the next condition. The scanner settings were the same for all conditions.

The baseline condition was also recorded, during which the subject lay quietly in the scanner with closed eyes - two minutes before the tasks and 2 minutes after the tasks were completed. Each experimental condition lasted for two minutes, the total duration of the experiment was about 10 minutes. The duration of each stimulus was 5000 ms; then, after a pause (400 ms), a “beep”-sound was presented to indicate the onset of pronunciation (6000 ms). Presentation was performed through special headphones compatible with an MR-compatible device. The task took 6000 ms to complete. In the condition aimed at listening to musical stimuli, there was no interval allotted for internal representation, the stimuli came in succession with a pause of 400 ms and started after the “beep”-sound. All conditions were conducted in a single session. The eyes of the subjects were closed throughout the experiment.

After participant completed all the tasks, he/she underwent a post-experimental interview. Its purpose was to clarify whether the participant managed to successfully complete the tasks, whether he/she used the lyric for free representation, whether he/she managed to sing only the melody in the third condition. Most participants answered both questions in the affirmative. Their data were taken into further analyses.

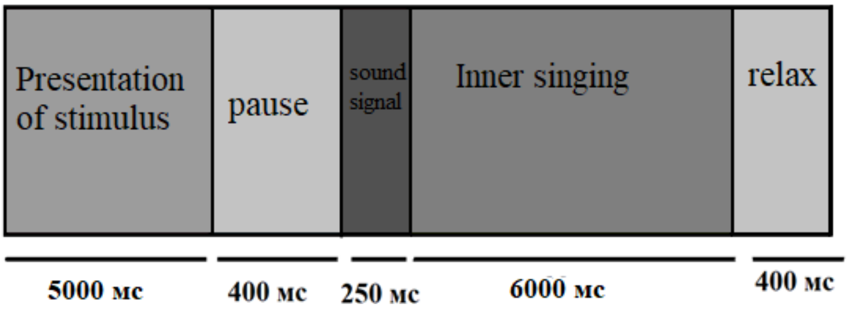

Figure 1. The timeline of experimental procedure

Image acquisition

Functional MRI was performed on a Siemens Magnetom Skyra 3T MRI Machine (Siemens Medical Systems, Erlangen, Germany) to obtain BOLD contrast (depending on blood oxygenation). A T1-weighted structural scan was performed for each subject (TR = 2200 ms, TE = 2.48 ms flip angle = 6, matrix = 256 × 256, 230 mm field of view, slices = 176, slice thickness = 1 mm). For functional scans (36 scans for a two-minute condition, three seconds each), we used a T2*-weighted gradient-echo EPI sequence (TR = 3000 ms, TE = 30 ms, flip angle = 90°, matrix = 128 × 128, 260 mm field of view, slices = 41, slice thickness = 3mm). Head movement has been minimised with a custom-fitted headrest.

Image analysis

Only 30 subjects were used for analysis, as images from three subjects had many motor artifacts and poor image quality (estimated rotational and translational motions were less than 2 mm for 30 subjects in this analysis). Image processing and data analysis were performed using the SPM12 statistical parametric mapping software package (The Wellcome Center for Human Neuroimaging, UCL Queen Square Institute of Neurology, London, UK, https://www.fil.ion.ucl.ac.uk /spm/software/spm12/). The first 10 volumes from each run were discarded to allow for T1 equilibrium effects. Standard preprocessing including slice timing correction, realignment and spatial normalisation. Then it was applied to the Montreal Neurological Institute (MNI) T1 high resolution template to demonstrate the results obtained. The images were further smoothed with a Gaussian kernel with a 3 voxels FWHM (FWHM – Full width at half maximum) to minimise noise and residual differences in the anatomy of the chiral structures, resulting in an effective spatial resolution of 6 × 6 × 9 mm. Each normalised image was bandpass filtered (low pass filter = canonical hemodynamic response function (HRF); high pass filter = 128 seconds to remove high and low frequency noise, respectively.

Statistical parametric maps were built using the General Linear Model (GLM) [Friston, 1994]. Next, group statistics were analysed using the one-sample Student's t-test. After that, the t-statistics were converted to normal standard deviation (z). The voxel activation threshold corresponded to pFWEuncorr. < 0.001 (pFWEuncorr – Family-Wise Error, uncorrected p-value). The choice of criterion is justified by the fact that, based on the study paradigm and the corresponding design, the compared series are identical at 50% (listening phase). Therefore, a more rigorous statistical criteria was chosen to obtain significant differences.

Statistical parametric maps were built to identify:

1) Zones of increased activation during the process of internal representation of melodies (reperesentation of a melody > listening to a melody). This will reveal the areas involved in the processing and production of melodies, as well as confirm the hypothesis about the participation of speech areas in the process of compiling melodies.

2) Zones of increased activation during the internal representation of melody (melody representation without lyrics > free internal representation of melody).

Results

Changing the activation level when listening to the melodies

Averaged SPMs were plotted for all 30 subjects. To begin with considering the zones that are more active during the process of listening to music stimuli (Fig. 2, Table 1).

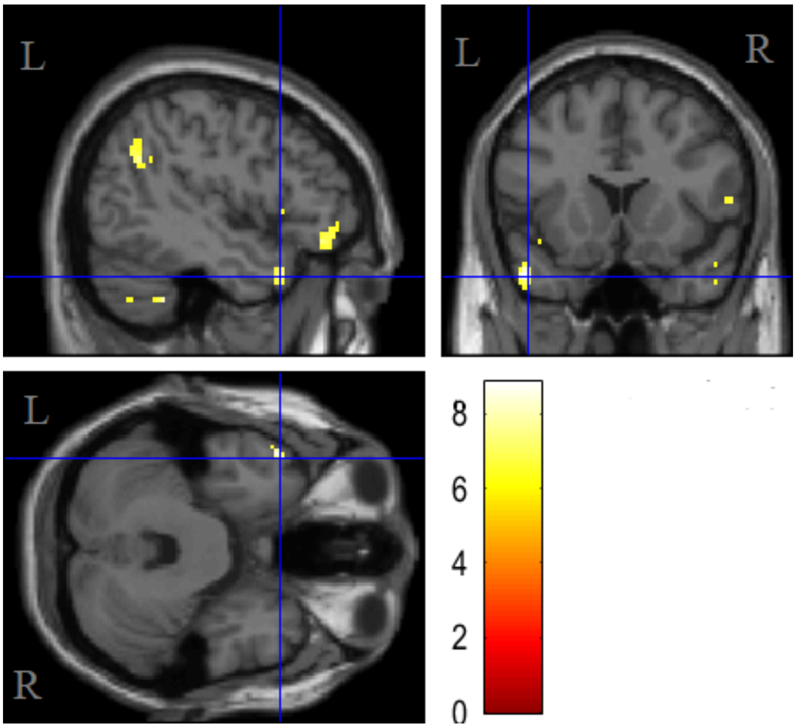

Figure 2. Group-averaged statistical parametric maps.
Note. The map demonstrates greater activation of cortical areas when listening to melodies compared to background (one sample t-test, pFWEunc < 0.001).

Table 1. Areas of increased activity when listening to music

Hemisphere	Brain structure	Cluster size	T	Z	P(unc)	x	y	z
R	Angular gyrus	47	7.913	5.731	< 0.001	–50	–56	20
L	Temporal pole	27	8.842	6.109	< 0.001	–48	16	–31
R	Lateral Orbital Gyrus	76	8.81	6.098	< 0.001	–34	48	–10

When comparing the activation of the cortex during listening to music, in comparison with the baseline series, one can see a greater activity of the angular gyrus, the lateral orbital gyrus of the right hemisphere, as well as the temporal pole of the left hemisphere. Homologous and cerebellar structures are also involved, but the activity is lower.

Changing the activation level for a given representation of a melody without lyrics

Next, areas of increased (Fig. 3, Table 2) and decreased (Fig. 4, Table 3) activity were considered during the internal representation of the melody without using lyrics.

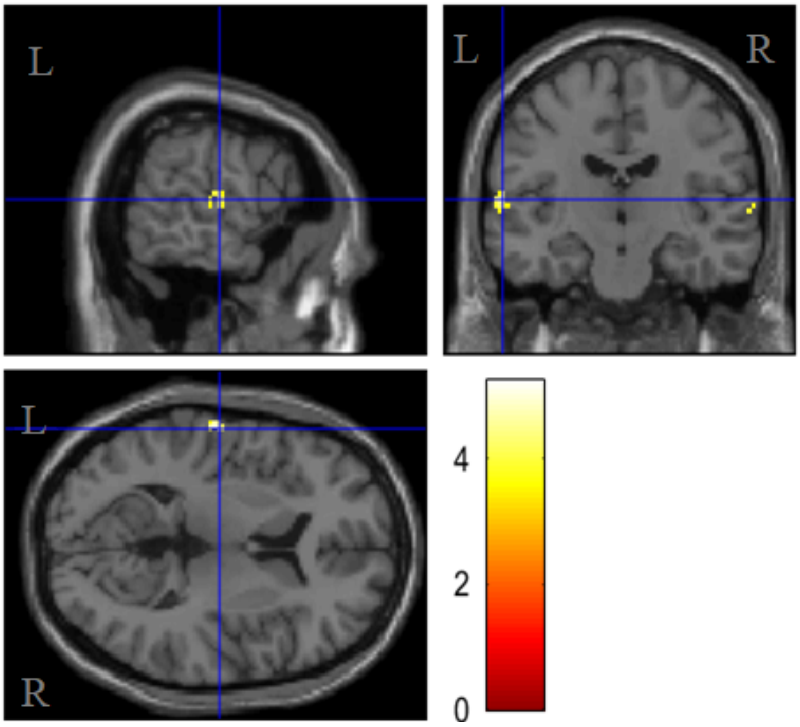

Figure 3. Group-averaged statistical parametric maps.

Note. The map demonstrates greater activation of cortical areas during internal representation of melodies without lyrics compared to listening to them (one sample t-test, pFWEunc – < 0.001).

Table 2. Areas of increased activity during the internal presentation of melodies

Hemisphere	Brain structure	Cluster size	T	Z	P(unc)	x	y	z
L	Planum Temporale	28	5.232	4.355	< 0.001	–62	–18	5

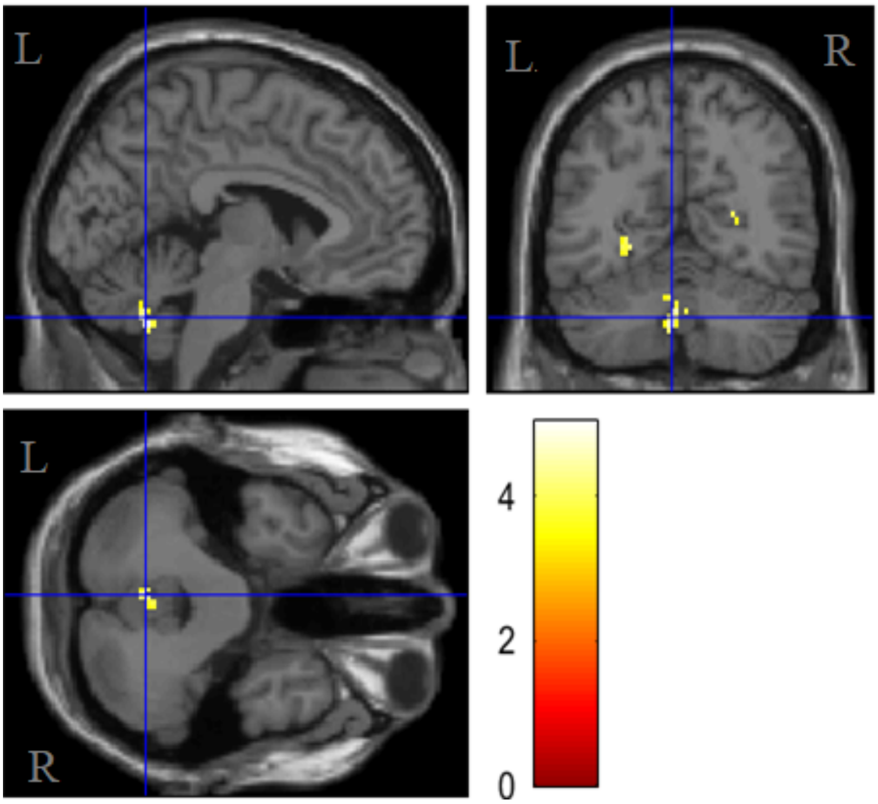

Figure 4. Group-averaged statistical parametric maps.

Note. The map demonstrates reduced activation of cortical areas during internal representation of melodies without lyrics compared to listening to them (one sample t-test, pFWEunc < 0.001).

Table 3. Areas of reduced activity in the internal presentation of the melody compared to the background series

Hemisphere	Brain structure	Cluster size	T	Z	P(unc)	x	y	z
	Cerebellum Vermal Lobules VIII-X	43	5.022	3.52	< 0.001	–4	–58	–40
L	Superior Temporal Gyrus	36	4.742	4.004	< 0.001	–26	–58	–4

Comparing the lyrics of the activation zones during the mental representation of the melody without lyrics, it can be observed that during the process of internal representation, the activity of the structures of the Planum Temporale in the right hemisphere increases, while the activity of such structures as the superior lobules of the cerebellum (7, 8, 9, 10) and the superior temporal gyrus in the left hemisphere decreases.

Changing the activation level during the free presentation of a melody

Further, the activity zones were compared during the free repetition of the melody, as well as during the repetition of the melody, excluding the lyrics. When identifying structures whose activity increases in a series with non-speech presentation, a statistically insignificant result was obtained. In reverse comparison, revealing structures that are more active with free repetition of the melody, the following results were obtained.

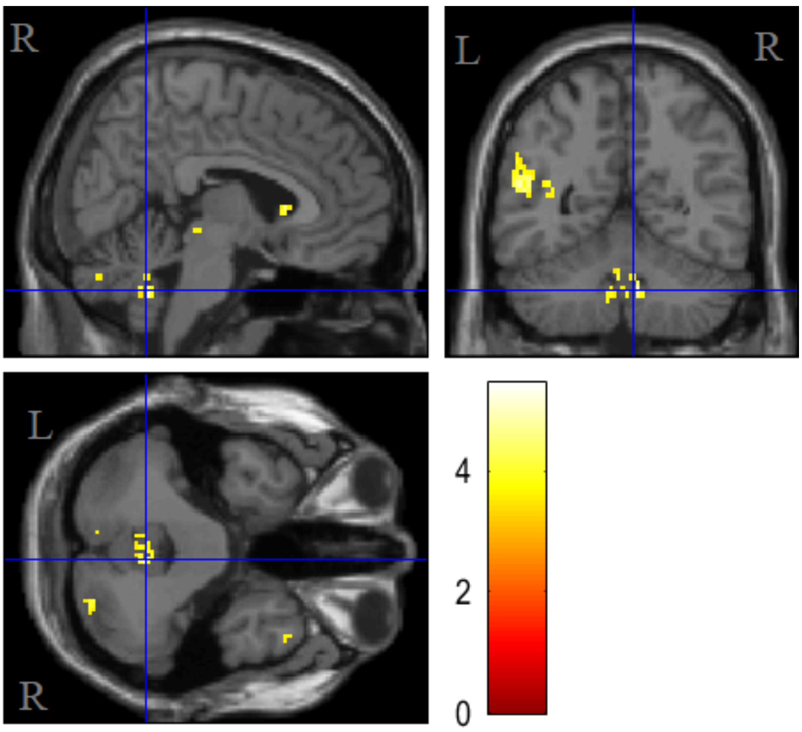

Figure 5. Group-averaged statistical parametric maps.

Note. The map demonstrates increased activation of cortical regions with free internal representation of melodies compared to their internal representation without lyrics (one sample t-test, pFWEunc < 0.001).

Table 4. Areas of increased activity in the free internal representation of melody compared to melody without lyrics

Hemisphere	Brain structure	Cluster size	T	Z	P(unc)	x	y	z
R	Cerebellum Execterior	63	5.145	4.301	< 0.001	8	–54	–37
L	Supramarginal Gyrus	176	4.921	4.152	< 0.001	–44	–48	–26
R	Caudate	10	4.51	3.89	< 0.001	6	18	5
L	Anterior Cingular Gyrus	6	3.73	3.35	< 0.001	–6	32	–20
R	MP Temporal pole	14	3.954	3.508	< 0.001	42	20	–34

It can be observed that during the presentation of a musical passage with the addition of lyrics, activation is increased in the structures of the cerebellum (Cerebellum exterior) and Caudate of the right hemisphere, the Supramarginal gyrus and Anterior Cingulate Gyrus of the left hemisphere.

Discussion

Changing the activation level when listening to the melodies

To begin with, consider the structures whose activation changes when listening to melodies. This will allow us to compare the processes of perception and mental imagery, as well as to determine the structures that are activated when listening to melodies.

Quite interesting results were obtained: in this study, there was not an increase, but a decrease in the hemodynamic response compared to the background state in the following structures: the angular gyrus, the lateral orbital gyrus, the temporal pole of the left hemisphere and their homologous zones in the right hemisphere of the brain. In addition, activity in the structures of the cerebellum of both hemispheres decreases. It was shown [Demorest, 2009] that listening to culturally unfamiliar music results in activation in the left cerebellar area, right angular gyrus, posterior precuneus, and right middle frontal area. In this experiment, subjects interacted with well-known melodies from their culture, and a decrease in activation was observed in all of the above structures compared to background, which is consistent with the findings.

Changing the activation level for a given representation of a melody without lyrics

Observing a decrease in activity in cerebellar structures, we can on the one hand assume that the process of image recreation is difficult, which does not lead to successful performance and obtaining a musical image. On the other hand, if the images were indeed successfully obtained, we can conclude that the task of melody repetition is not difficult enough to activate structures such as the cerebellum. Also, a decrease in activity is observed in the basal ganglia bilaterally. The present results correlate with the results on the role of the basal ganglia and the cerebellum in the binding of melody and text [Alonso, 2016] – here the process is reversed. Also, this phenomenon in the process of music perception may indicate the involvement of motor structures in the process of interaction with music [Martín-Fernández, 2021; Dikaya, 2015].

The observed decrease in activity level compared to listening in the angular gyrus may indicate that the representation of the melody is not compensated by the integration of visual-speech images at the stage of listening to melodies, but integration is inhibited [Bonner, 2013].

It is important to note that the parametric maps of each subject were individual: in some subjects, when listening to melodies, the activity of the visual cortical zones significantly increased, while in others, the activity of kinetic, tactile and other zones rose. Also, the average map included individual maps of subjects who experienced an increase and decrease in the activation of structures.

When internal representation of melodic constructions that exclude lyrics occurs, the strongest activation occurs in the Planum Temporale of the left hemisphere, which includes the 22nd area (Wernicke's area) and the homologous area in the right hemisphere, which is consistent with the results of other studies [Zhang, 2017]. And when listening, strong activity is observed in the cerebellum and STG. This observation may indicate that when playing a melody, excluding text, the motor components of music perception are inhibited. However, when a melody is repeated without text (the stage also included listening to a stimulus), the usual musical image is reconstructed to a greater extent, in which the inferior frontal gyrus (IFG) is involved, and, in addition, IFG is inhibited in the left hemisphere, while IFG in the right hemisphere remains active.

As has been shown, when using familiar songs with lyrics, two-way activation occurs, because a person has images of both the sung text and the musical component [Zatorre et al., 1996]. When only instrumental stimuli are used [Halpern et al., 2004], the pattern shifts towards activation in the right auditory cortex. Activation of the left auditory cortex has been demonstrated when working with non-verbal materials [Kraemer et al., 2005].

Changing the activation level during the free presentation of a melody

When comparing the activation of zones during free reproduction with the stage with the mandatory exclusion of texts, differences in activity in the structures of the cerebellum, amygdala, caudate nucleus, cingulate gyrus, superior temporal and supramarginal gyrus were revealed.

The level of activation of these structures is much higher with free representation. This can be explained by memorised motives [Stracheus, 2003], which form the basis of the image with free instruction. This is an integral indivisible structure, which, when memorised, becomes automated, that is, a holistic image of a sequence of sounds is created in memory, and not a set of images of each note that makes up this motif. Transformations of motives are provided by functional activation within the parietal sulcus (IPS), which is similar in mechanism to visuospatial and numerical transformations provided by the activity of the same structure [Foster et al., 2013].

That is, with a free representation, a person operates precisely those forms of well-known motives that are associated with the image of a particular melody. Thus, this image includes motor and emotional (including the fact that these images are familiar) components, which explains the greater activation of the caudate nucleus and cingulate gyrus. This can be explained by the fact that the usual images of various melodies differ in affective colouring, in contrast to the constructed image of melodic sequences. This is due to the fact that the melodies selected as stimuli are known primarily for their verbal content, so the memorised motive includes the text, and its inhibition leads to the fact that emotional stimuli do not differ from each other.

An increase in the activation of the superior temporal gyrus (STG) can be associated with the fact that the habitual motives of these stimuli may include a verbal component. Since due to STG there is a perception of the human voice, as well as the arbitrariness of the functioning of attention between different aspects in a complex stimulus, it can be assumed that musical images consist of several components, such as rhythmic structure, main motive, background motive, verbal component, while attention can focus on each of the elements. Further studies are required to examine this phenomenon in more detail.

An increase in activation in the supramarginal gyrus in the left hemisphere may be associated with the activation of mirror neurons [Wang and Agius, 2018]. In music production, the mirror neuron system in the left fronto-parietal network plays an important role, since this process is based on a person's experience (including the selection of a way to form a habitual way of representing a melody) [Singer, 2015].

In the task of internal representation of the pure melodic component, no structures were identified whose activation increased compared to the free representation task. This may be because the free representation process was more complex and included an additional component (e.g., verbal, emotional ones) that was absent in the series with text exclusion. The presented stimuli contained only a melodic line and the subjects were required to recall and recreate the melody they had just heard. During the free representation task the image became more complex, as shown by the activation of such structures as the cerebellum and temporal pole (cognitively complex verbal information) [Herholz, Halpern et al., 2012], as well as the cingulate gyrus and caudate (emotional and motor component in the perception and imagination of familiar melodies and songs) [Sikka, 2015].

The absence of an increase in STG activation may indicate that the components included in the song are more likely to be divided into verbal and musical. The latter includes rhythm, main motive, background motive etc. The specifics and mechanisms of each of these components will be studied in further research.

Conclusions

The activation of brain structures that changes when imagining melodies that include and exclude components related to the memorisation of the motif (verbal, emotional) has been revealed. There are different patterns of activation of brain areas during different tasks, the map of which includes cortical and subcortical structures. The encoding of musical patterns can have individual characteristics and be variable, which can be related to individual strategies for representing the melodic image, but we can speak of general tendencies.

It was found that listening to these stimuli decreases the activity of such structures as the angular gyrus, lateral orbital gyrus, temporal pole in the left hemisphere, their homologous areas in the right hemisphere, and also decreases the activity in the cerebellar structures of both hemispheres.

During internal representation of pure melodic constructions, greater activation is observed in the temporal plane of the left hemisphere, including area 22 (Wernicke's area), as well as its homologous area in the right hemisphere, and during listening, greater activity is observed in the cerebellum and STG. This may indicate that the singing of instrumental melodies, excluding additional components absent in the original stimulation differs from the encoded automatic musical image. Further studies are required to investigate in more detail the neural mechanisms of attention switching to a less dominant image.

During free representation, greater activation of the caudate nucleus, cingulate gyrus is detected, which may indicate a more complex image of familiar melodies (including a verbal component) that differs from and overlaps with the original stimulation.

In addition, there is increased activation of the superior temporal gyrus (STG), which may be responsible for a number of complex coding and integrating functions. The changes in free presentation involve several subcortical structures: the basal ganglia, cerebellum are associated with the motor component, which is very different from the series that examined changes in activation level during a given presentation of a melody without lyrics. The cerebellum and left temporal areas are one indicator of the difficulty of the melody representation task.

Limitations

The current study contains several limiting factors. First of all, researchers rely on the subjects' self-report about inner speech, however, the process of inner speech itself cannot be controlled by researchers. This leads to the fact that we only assume that the subjects correctly followed the instructions.

It is also necessary to take into account the temporal resolution of fMRI, which is 3000 ms in this experiment, does not allow us to draw unambiguous conclusions about the dynamic interaction between the brain regions that are activated during perception and the process of internal representation. To solve this problem, variants of additional EEG and MEG studies were proposed, which would allow measuring the activity of brain structures on a millisecond scale.

There is also a difficulty in separating the process of internal representation of a melody, taking into account the components of rhythm and intonation, from the processes of memory, perception, and problem solving, given that all of these processes occur at the same moment. In addition to separating the process of inner speech from other processes, there is a problem of separating this process itself, as well as separating the process of verbalising a melody from the process of presenting a non-verbalized melody.

Accordingly, when using the results of the study of internal speech processes, these limitations should be taken into account.

Ethics statement

All participants signed informed consent after the experimental protocol was explained to them. «Application #4 approved during the Ethic Commission meeting #6 held on 18.06.2021.

Funding

The research was financially supported by the Russian Science Foundation, project No 20-18-00067-П.

Definition of abbreviations

FWHM – Full width at half maximum
pFWEuncorr/pFWEunc – Family-Wise Error, uncorrected p-value

References

Dikaya L.A., Naumova M.I., Naumov I.V. Psikhofiziologicheskiye korrelyaty psikhicheskogo ispolneniya improvizirovannogo tantsa. Rossiyskiy psikhologicheskiy zhurnal, 2015, Vol. 12, No 4, 110–126 (In Russian)

Saks O. Muzykofiliya: rasskazy o muzyke i mozge. N'yu-York: Al'fred A. Knopf, 2007. (In Russian)

Starcheus M.S. Slukh muzykanta. M.: Mosk, Gosudarstvennaya konservatoriya im. PI. Chaykovskiy, 2003 (In Russian)

Alonso I., Davachi L., Valabrègue R., Lambrecq V., Dupont S., Samson S., Neural correlates of binding lyrics and melodies for the encoding of new songs, NeuroImage. 2016, No 127, 333–334. DOI:10.1016/j.neuroimage.2015.12.018

Bocquelet F., Hueber T., Girin L., Chabardès S., Yvert B. Key considerations in designing a speech brain-computer interface. Journal of Physiology-Paris. 2016, No 110(4A), 392–401. DOI:10.1016/j.jphysparis.2017.07.002

Bonner M., Peelle J., Cook P., Grossman M. Heteromodal conceptual processing in the angular gyrus. NeuroImage. 2013, No 71, 175–186. DOI:10.1016/j.neuroimage.2013.01.006

Boulenger V., Hauk O., Pulvermueller, F. Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. Cerebral Cortex. 2009, No 19, 1905–1914. DOI:10.1093/cercor/bhn217

Chiang J., Rosenberg M., Bufford C, Stephens D, Lysy A, Monti M. The language of music: Common neural codes for structured sequences in music and natural language. Brain and Language. 2018, No 185, 30–37, DOI:10.1016/j.bandl.2018.07.003

Demorest S., Morrison S., Stambaugh L., Beken M., Richards L., Johnson C. An fMRI investigation of the cultural specificity of music memory. Social Cognitive and Affective Neuroscience. June/September 2010, No 5(2-3), 282–291. DOI:10.1093/scan/nsp048

Ferstl E. C., Neumann J., Bogler C., von Cramon D. Y. The extended language network: A meta-analysis of neuroimaging studies on text comprehension. Human Brain Mapping. 2008, No 29, 581–593. DOI:10.1002/hbm.20422

Foster N., Halpern A., Zatorre R., Common parietal activation in musical mental transformations across pitch and time. Neurolmage. 2013, No 75, 27–35. DOI:10.1016/j.neuroimage.2013.02.044

Gabriel D.; Wong T.C.; Nicolier M.; Giustiniani J.; Mignot C.; Noiret N.; Monnin J.; Magnin E.; Pazart L.; Moulin T.; Haffen E.; Vandel P. Don’t forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs. Neurobiology of Learning and Memory. 2016, No 132, 18–28. DOI:10.1016/j.nlm.2016.04.011

Halpern A. Memory for the absolute pitch of familiar songs. Memory and Cognition. 1989, No 17(5), 572–581.DOI:10.3758/BF03197080

Halpern A. Mental scanning in auditory imagery for songs. Journal of Experimental Psychology. Learning, Memory, and Cognition. 1988, No 14(3), 434–443. DOI:10.1037/0278-7393.14.3.434

Halpern A., Zatorre R. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. Cerebral Cortex. 1999, No 9(7), 697–704. DOI:10.1093/cercor/9.7.697

Halpern A., Zatorre R., Bouffard M., Johnson J. Behavioral and neural correlates of perceived and imagined musical timbre. Neuropsychologia. 2004, No 42(9), 1281–1292. DOI:10.1016/j.neuropsychologia.2003.12.017

Hamilton L. Human song: Separate neural pathways for melody and speech. Current Biology. 2022, No 32(7), R311–R313. DOI:10.1016/j.cub.2022.02.037

Herholz Sibylle C., Halpern Andrea R., Zatorre Robert J.; Neuronal Correlates of Perception, Imagery, and Memory for Familiar Tunes.Cognitivr Neuroscience, 2012; No 24(6): 1382–1397. DOI: 10.1162/jocn_a_00216

Kopp R. Metaphor Therapy: Using Client Generated Metaphors In Psychotherapy. Routledge. 2013, 216

Kraemer D., Macrae, C., Green, A. et al. Sound of silence activates auditory cortex. Nature. 2005, No 434, 158. DOI:10.1038/434158a

Kunert R., Willems, R., Casasanto D., Patel A., Hagoort P., Berwick R. Music and Language Syntax Interact in Broca's Area: An fMRI Study. PLOS ONE. 2015, No 10(11). DOI:10.1371/journal.pone.0141069

Lacey S., Stilla R., Sathian K. Metaphorically feeling: Comprehending textural metaphors anctivates somatosensory cortex. Brain and Language. 2012, No 120, 416–421, DOI:10.1016/j.bandl.2011.12.016

Maess B., Koelsch S., Gunter T., Friederici A. Musical syntax is processed in Broca's area: An MEG study. Nature neuroscience. 2001, No 4, 540–545. DOI:10.1038/87502.

Martín-Fernández J., Burunat I., Modroño C., González-Mora JL, Plata-Bello J. Music Style Not Only Modulates the Auditory Cortex, but Also Motor Related Areas. Neuroscience. 2021, No 457, 88–102. DOI:10.1016/j.neuroscience.2021.01.012

Minguillon J., Lopez-Gordo M. Angel P. Trends in EEG-BCI for daily-life: Requirements for artifact removal. Biomedical Signal Processing and Control. 2017, No 31, 407–418. DOI:10.1016/j.bspc.2016.09.005

Norman-Haignere S., Kanwisher N., McDermott J. Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. Neuron. 2015, No 88(6), 1281–1296. DOI:10.1016/j.neuron.2015.11.035

Pinker S. How the Mind Works. Norton. 1997, 50.

Mor Regev, Halpern Andrea R, Owen Adrian M, Patel Aniruddh D, Zatorre Robert J, Mapping Specific Mental Content during Musical Imagery, Cerebral Cortex, 2021, Volume 31, Issue 8, 3622–3640, DOI:10.1093/cercor/bhab036

Saito Y, Ishii K, Sakuma N, Kawasaki K, Oda K, Mizusawa H Neural Substrates for Semantic Memory of Familiar Songs: Is There an Interface between Lyrics and Melodies? PLoS ONE 2012, No 7(9): e46354. DOI:10.1371/journal.pone.0046354

Schmithorst V. Separate cortical networks involved in music perception: preliminary functional MRI evidence for modularity of music processing. Neuroimage. April 2005, No 25(2), 444–451. DOI:10.1016/j.neuroimage.2004.12.006

Sikka R. et al. An fMRI comparison of neural activity associated with recognition of familiar melodies in younger and older adults //Frontiers in Neuroscience., 2015, No 9, 356. DOI:10.3389/fnins.2015.00356

Singer N, Jacoby N, Lin T, Raz G, Shpigelman L, Gilam G, Granot RY, Hendler T. Common modulation of limbic network activation underlies musical emotions as they unfold. Neuroimage. 2016, No 141, 517–529. DOI:10.1016/j.neuroimage.2016.07.002

Whitehead, Jocelyne C.; Armony, Jorge L. Singing in the brain: Neural representation of music and voice as revealed by fMRI. Human Brain Mapping, 2018, No 39, 4913–4924. DOI:10.1002/hbm.24333

Zatorre R., Halpern A. Music. Brain Mapping, Academic Press. 2015, 551-557. DOI:10.1016/B978-0-12-397025-1.00283-9

Zatorre R., Halpern A. Mental Concerts: Musical Imagery and Auditory Cortex. Neuron. 2005, No 47(1), 9–12. DOI:10.1016/j.neuron.2005.06.013

Zatorre R., Halpern A., Perry D., Meyer E., Evans A. Hearing in the mind's ear A PET investigation of musical imagery and perception. Journal of Cognitive Neuroscience. 1996, No 8(1), 29–46. DOI:10.1162/jocn.1996.8.1.29

Zatorre R., Halpern A., Mental Concerts: Musical Imagery and Auditory Cortex. Neuron. 2005, No 47(1), 9-12. DOI:10.1016/j.neuron.2005.06.013

Zatorre R., Halpern A., Perry D., et al., Hearing in the mind's ear: a pet investigation of musical imagery and perception. Journal of Cognitive Neuroscience. 1996, No 8, 29. DOI:10.1162/jocn.1996.8.1.29

Zhang Y., Chen G., Wen H. et al. Musical Imagery Involves Wernicke's Area in Bilateral and Anti-Correlated Network Interactions in Musicians. Scientific Reports. 2017, No 7(1), 1–13. DOI:10.1038/s41598-017-17178-4

Поступила в редакцию: 20 ноября 2023 г. Дата публикации: 24 июля 2024 г.

Information about authors

Maria Dmitrievna Krysko, Department of Psychology, Lomonosov Moscow State University (Moscow, Russia), ORCID 0000-0001-9263-5203 – Corresponding author.

Email: mariya.krysko@mail.ru

Alexander Valentinovich Vartanov, PhD in Psychology, Associate Professor, Senior Researcher, Department of Psychology, Lomonosov Moscow State University (Moscow, Russia). ORCID 0000-0001-8844-9643. SPIN: 9511-2918, AuthorID: 72326. ResearcherID: D-9907-2012. Scopus Author ID: 6603018783.
Email: a_v_vartanov@mail.ru

Oleg Yurevich Bronov, PhD in Medical Sciences, Radiologist, Associate Professor of the Department of Radiation Diagnostics with a course in Clinical Radiology, Head of Department of the Federal State Budgetary Institution N.I. Pirogov National Medical Clinical Center of the Ministry of Health of Russia (Moscow, Russia). ORCID 0000-0002-2784-302X

For citation: Krysko M.D., Vartanov A.V., Bronov O.Y. Verbal component suppression during internal representation of songs: fMRI-study. Psikhologicheskie Issledovaniya, 2024, Vol. 17, No. 94, p. 2. https://psystudy.ru

Адрес статьи: https://doi.org/10.54359/ps.v17i94.1578

Крысько М.Д.[1], Вартанов А.В.[1], Бронов О.Ю[2]. Подавление вербального компонента внутренней репрезентации песни: фМРТ исследование

[1] Московский государственный Университет им. М.В.Ломоносова

[2] Федеральное государственное бюджетное учреждение «Национальный медико-хирургический центр имени Н.И. Пирогова

В данной работе исследуется один из процессов внутренних репрезентаций - в частности, мысленная репрезентация мелодий с подавленной вербализацией. Цель исследования - изучить вовлеченность структур мозга в процесс мысленного представления мелодий. В фМРТ-исследовании приняли участие 33 здоровых праворуких человека. Перед участниками ставилась задача прослушать музыкальные отрывки, сопровождаемые текстами песен в их оригинальной форме, а затем мысленно воспроизвести эти отрывки с текстом и без него. Полученные результаты свидетельствуют о том, что кодирование музыкальных паттернов может иметь индивидуальные вариации, однако были выявлены общие тенденции. Результаты показывают, что как корковые, так и подкорковые структуры вносят вклад во внутреннее представление мелодий. В частности, область Вернике и гомологичная ей зона в правом полушарии играют важную роль во внутренней репрезентации, а хвостатое ядро, поясная извилина и верхняя височная извилина - в торможении вербального компонента.

Ключевые слова: музыка, фМРТ, внутренняя репрезентация, внутренняя речь, воображение, музыкальные образы, подавление текста.

Финансирование

Исследование выполнено при поддержке Российского научного фонда, проект № 20-18-00067-П

Сведения об авторах

Крысько Мария Дмитриевна, факультет психологии, Московский государственный университет имени М.В. Ломоносова (Москва, Россия), ORCID 0000-0001-9263-5203 -- Корреспондирующий автор.

Email: mariya.krysko@mail.ru

Вартанов Александр Валентинович, кандидат психологических наук, доцент, старший научный сотрудник, факультет психологии, Московский государственный университет имени М.В. Ломоносова (Москва, Россия). ORCID 0000-0001-8844-9643. SPIN: 9511-2918, AuthorID: 72326. ResearcherID: D-9907-2012. Scopus Author ID: 6603018783.

Email: a_v_vartanov@mail.ru

Бронов Олег Юрьевич, кандидат медицинских наук, врач-рентгенолог, доцент кафедры лучевой диагностики с курсом клинической радиологии, заведующий отделом ФГБУ «Национальный медицинский клинический центр имени Н.И. Пирогова» Минздрава России (Москва, Россия). ORCID 0000-0002-2784-302X

Ссылка для цитирования

Крысько М.Д., Вартанов А.В., Бронов О.Ю. Подавление вербального компонента при внутренней репрезентации песен: фМРТ-исследование. Психологические исследования. 2024. Т. 17, № 94. С. 2. URL: https://psystudy.ru