Effects of different forms of engagement on the neuronal activity in the monkey's primary auditory cortex

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften
der Otto-von-Guericke-Universität Magdeburg

von Specialist in Physiology Stanislava Knyazeva
geb. am 27. 09. 1988 in USSR

Gutachter: Prof. Dr. Michael Brosch
Prof. Dr. Thomas Schanze

eingereicht am: 23.03.2018
verteidigt am: 17.09.2018

Acknowledgment

This work was performed in Leibniz Institute for Neurobiology in the Special Lab for Primate Neurobiology and supported by Otto von Guericke University in Magdeburg. Thanks to many people who contribute to good scientific foundation and mutual support of these organizations this work could be brought to the end.

I am especially grateful to Prof. Dr. Michael Brosch who worked actively to provide me with time, freedom for creativity and constant support. I am also very thankful for his patience and tolerance to stream of my endless ideas.

I am sincerely thankful to my mentor in past and good friend today Dr. Elena Selezneva for everything she taught me during several years of work in the institute, for all her shared knowledge that was not limited to the scientific world only.

I would like to express my happiness to meet PhD. Alexander G. Gorkin on my life's journey without whom everything would be much different today.

I am grateful to all of those with whom I have had pleasure to work all this years. Especial thanks to Cornelia Bucks and her help to me with languages that she knew better than I did, to her capability to create right atmosphere in the lab.

Thanks to Dr. Igor V. Bondar, Prof. Dr. Peter Heil, Prof. Dr. Kristine Krug, Prof. Dr. Frank Ohl and Prof. Dr. Thomas Schanze for their support and/or immediate agreement to review my thesis.

Lastly but perhaps the most important gratitude is to my family who believed in me and wisely did not ask unwanted questions about successes and failures in doctoral students' life. Separate and special thanks to Ronny Pachel for his time, for his being near me, for my peace of mind.

Table of Contents

1. Introduction

For a long period of time, it was considered that the only function of the primary auditory cortex, as follows from its name, consists in the preparation of the auditory information for further analysis in higher cortexes. However, the last century of research showed that the primary auditory cortex is much more complex than it was commonly believed because it participates in a variety of processes.

In the last three decades, more and more studies demonstrated neuronal plasticity in the primary auditory cortex of an adult during learning. The studies described changes in the responses after learning of association of acoustical stimuli and reinforcement (Ohl and Scheich 2005, Suga et al. 2002, Suga and Ma 2003, Weinberger and Diamond 1987). Later on, a bunch of studies showed that the responses to the same acoustical stimuli may change in short intervals of some seconds when the presented sounds were and were not paired with reinforcement (Fritz et al. 2007a, b).

Additionally to the neuronal plasticity in the primary auditory cortex, it was found that some neurons of the cortex are multisensory. The study of Brosch and colleagues (2005) revealed that the neurons in the cortex respond not only to acoustical stimuli but also to visual and motor/somatosensory stimuli while animals performed a task to which they were highly trained. After that, other studies confirmed the responsiveness of the neurons in the primary auditory cortex to other sensory modalities and described their sensitivity from one to three modalities (Bizley et al. 2007, Driver and Noesselt 2008, Hoefer et al. 2013, Kayser et al. 2008).

The phasic changes in the neuronal activity of the primary auditory cortex might be explained through rich connections with other cortical and subcortical areas of the brain. Indeed, the rich net connecting the structures was described by our colleagues (for review see Scheich et al. 2011). Besides the phasic changes, the slow changes were observed in the neuronal activity during some cognitive processes (for review see Brosch et al. 2011a). For instance, the slow modulations in the neuronal activity appeared during memorizing (Huang et al. 2016a, Huang and Brosch 2016), attention (Atiani et al. 2014), expectation (Shinba et al. 1995) and other processes.

Thus, the previous studies revealed that the primary auditory cortex is much more difficult than it was expected at the beginning of the 20th century. The present study continues and expands the previous discoveries. Here we aimed to find and report the influence of different forms of engagement on the neuronal activity in the primary auditory cortex.

1.1. Effects of sound meaning and unconditioned stimuli alone

One of the predecessors of the present study was a publication of Scheich and colleagues (Ohl and Scheich 2005, Scheich 1991, Scheich et al. 2011) in which the auditory cortex was considered as an area that is highly connected with the cognitive-, reinforcement-, emotional- and learning-related areas of the brain. The colleagues explained that such a rich net with other structures of the brain might be explained by the high integration of the auditory cortex in the cognitive processes such as association and learning. Indeed, many studies confirming the neuronal plasticity in the auditory cortex were made in the last three decades.

The most active research of the plasticity in the primary auditory cortex was made by Weinberger and Diamond. In one of the study, they paired acoustical stimuli with a negative reinforcement for many times and revealed three types of the neurons in the primary auditory cortex. In the first group of neurons, the responses to the acoustical stimuli increased in contrast with the control group before the pairing, whereas the activity decreased in the second group of neurons, and did not change in the third group of neurons (Diamond and Weinberger 1989, 1986, Weinberger and Diamond 1987, Weinberger et al. 1984). Similar results were described in other studies, which were conducted a little earlier (Kraus and Disterhoft 1982). A slightly different paradigm, in which one frequency of the pure tones, the target frequency, was taken as a conditioned stimulus, and all other served as the control stimuli (Ohl and Scheich 1996, 1997), revealed the same three types of the changes after comparisons of the responses to the target frequency before and after pairing with negative reinforcement.

When the earlier studies showed the bidirectional changes of the responses, the later studies reported only about one type of the neuronal changes (Blake et al. 2002, Blake et al. 2006). It was shown that the neuronal responses to target and non-target tones decreased in the primary auditory cortex during training of monkeys; in the last session, the responses to the target tones were higher than to the non-target tones (Blake et al. 2002). In the next study of this group (Blake et al. 2006), naïve and previously trained monkeys were presented with the same acoustical stimulations. The authors revealed that the neuronal responses to the target tones of the previously trained monkeys were higher than of the naïve monkeys. The authors concluded that the plasticity changes occurred only after learning of the association.

Additionally to the changes in the responses evoked by the acoustical stimuli, some of the studies described changes of the background activity during the pairing of the acoustical stimuli with the negative reinforcements (Diamond and Weinberger 1984, Weinberger et al. 1984). They also claimed that the level of arousal of the animal increased with the increase in the background activity (Weinberger and Diamond 1987). In the studies of Ohl and Scheich (1996, 1997), both, the decrease and the increase in the background activity, were found between sessions with and without pairing of the acoustical stimuli with reinforcement.

It is reasonable to assume that the differences in the neuronal activity evoked by the acoustical stimuli with different values, which were observed in the barely trained animals, will be present at the same extend or even more in the primary auditory cortex of an animal after successful training of an instrumental condition. More recent studies seek to find whether the plasticity occurred in some minutes, i.e. between the responses to the acoustical stimuli in the instrumental condition and passive presentation of the same acoustical stimuli before or after the instrumental condition. Some research groups revealed that the responses to the acoustical stimuli during the passive presentations were higher than during the instrumental conditions

(Abolafia et al. 2011, Lee and Middlebrooks 2011, Otazu et al. 2009). In contrast, many studies of another group (Fritz et al. 2005, Yin et al. 2014) described a higher response to the acoustical stimuli that predicted reinforcement in the instrumental condition compare with the response to the same acoustical stimuli that were presented to passive animals.

During instrumental conditions, additionally to the phasic changes of response, some researchers demonstrated tonic changes (Brosch et al. 2011a, 2011b, 2015, Selezneva et al. 2017). The slow changes were explained by providing neuronal mechanisms for associating, memorizing and anticipating of important events (Brosch et al. 2011a). In the study of David and colleagues (2012) neuronal activity during acoustical stimuli increased from the beginning until the moment of the reinforcement but was unchanged during acoustical stimulation during the passive listening. Another study described that 16% of neurons in the rat’s primary auditory cortex showed a ramp up or ramp down when the animal waited for the second acoustical stimulus that was necessary for the correct condition performance (Abolafia et al. 2011). In a similar experiment, where a rat was required to wait for the second acoustical stimuli to give an appropriate behavioral response, activity of some neurons in the auditory cortex ramped up during the waiting period (Shinba et al. 1995). It is interesting that the tonic changes were also observed in the sensory thalamus (Komura et al. 2001). The study described the ramp up during active licking which was required for the water delivery after acoustical stimulation.

Thus, the auditory cortex, which is more commonly known to be in charge of perception and analysis of acoustical stimuli, is also involved in the learning that would require existence of connection with the higher and lower organized structures of the brain. Such lower structures are the ventral tegmental area (VTA) and substantia nigra (SN) that were observed to be activated to novel stimuli, appetitive stimuli and to unexpected positive reinforcement (for review see Schultz 2002). The presence of the direct or indirect connection with the structures and the auditory cortex of mammals was demonstrated by Bao et al. (2001), Huang et al. (2016b) and Lou et al. (2014) and Budinger et al. (2008). One study described decrease in the spontaneous activity of the auditory cortex between response to an acoustical stimulus that was presented alone and the same acoustical stimulus paired with the VTA stimulation (Huang et al. 2016b). But the study did not reveal any changes in the responses to the acoustical events during the pairing. In contrast, another study showed that the response to the acoustical stimuli in the primary auditory cortex was lower when the acoustical stimuli were paired with the VTA stimulation (Lou et al. 2014). Interestingly, Lou and colleagues did not describe the changes in the spontaneous activity. Prolonged pairing of the acoustical and the VTA stimulations led to the changes of the tonotopical map in the auditory cortex, particularly to the expansion of the area of the frequency that was used for the pairing while the slight contraction of all other areas (Bao et al. 2001). If the connection between the VTA, as a reward-related structure, and the primary auditory cortex is indeed so strong then we can expect to find changes related to appetitive acoustical stimuli and to reinforcement in the neuronal activity of the primary auditory cortex. Moreover, other structures, such as nucleus basalis, nucleus accumbence and amigdala, responses of which are related to reinforcement and associations, also have strong connections with the primary auditory cortex (for review see Scheich et al. 2011) and, therefore, might lead to the differences in responses.

Summarizing, we divided the studies by two groups: the studies that based their knowledge on the result of the neuronal changes between the responses to acoustical stimuli that were paired with reinforcement and the same acoustical stimuli before the animals learned the association very well; and the studies in which the changes in responses were shown in the highly trained animals between the instrumental conditions and passive listening of the same acoustical stimuli before or after the session. Thus, the change of the sound value, the meaning of the sound, led to the changes in the primary auditory cortex. We wondered, which of the changes listed above we will find between the two conditions in which acoustical stimuli were and were not paired with unconditioned stimuli, in a group of monkeys with low experience. Also we questioned, will this differences be the same in the group of monkeys with higher experience.

It is important to note that the acoustical stimulation without pairing with unconditioned stimuli controls only one factor, i.e., the meaning of the sound. The second factor, the effect of the unconditioned stimuli, is not considered. Interestingly, that the effect of the unconditioned stimuli alone was also not controlled in the studies listed above. We could find only one study that partially addressed the question (Ide et al. 2012). The study revealed the effect of an aversive stimulus alone in the neurons of experienced rodents in the auditory cortex when the same aversive stimuli did not affect the activity of the naïve animals. Thus, the present study will control the effect of the positive unconditioned stimuli alone in the groups of the monkeys with low and high experience.

1.2. Sense of agency

The sense of agency, which is defined as a subjective awareness of control of its own volitional action (Jeannerod 2003), was described as a cognitive function and is more commonly associated with the posterior parietal lobe (Farrer and Frith 2002, Shimada et al. 2005) and posterior temporal cortex (Jackson and Decety 2004). Many studies have shown that neuronal activity in the auditory cortex also depended on the sense of agency. The MEG magnitudes (M100) evoked by the subject’s speech were compared with magnitudes evoked by the same speech played back to the subject. The first condition evoked lower M100 with longer latency than the second condition (Curio et al. 2000, Houde et al. 2002). Another study, in which MEG of the auditory cortex was recorded, revealed that the M100 was higher when the subject read silent than aloud (Numminen et al. 1999). Similar results were observed in the potentials of EEG recorded in the brainstem, wherein amplitude increased from aloud speech to whispering, to just articulation and had maximum to just acoustical clicks (Papanicolaou et al. 1986). Another work with artificial acoustical stimuli showed similar results (Martikainen et al. 2005). They compared magnitudes evoked by the self-initiated electronic sounds with the same sounds produced independently of the subject. The M100 in the auditory cortex was lower when the sound was initiated by the subject. Similar results were observed on potentials of EEG by Schafer and Marcus (1973).

Neuronal electrophysiological studies on human and non-human primates also revealed dependence of the firing in the auditory cortex on the changes of the sound sources. The suppression of activity was observed in superior temporal gyrus of human during speaking whereas the same phrases that were heard by this listener led to increase in activity (Creutzfeldt et al. 1989). Similar results were observed by Eliades and Wang (2003, 2005) in auditory cortex of marmosets. They also observed two groups of neurons in the auditory cortex of marmosets that exhibited inhibitory responses to their own vocalizations but excitatory responses to the play back. They also described that neurons increased their activity if played back sounds were presented during vocalizations.

It is worth noting that between most of the studies that can be found in the literature the comparison was conducted between two different situations in which one involved a subject into the process (vocalization, speech, self-initiation of a sound) when another was unimportant (play back of the vocalization or speech, externally produced sounds without any meaning). Therefore, the self-produced acoustical stimuli of the studies were a mix of two meanings in which one was the sense of agency and the second was the meaning of the sound; when the externally-initiated sounds had only one factor of sense of agency. The present study will seek to find the true effect of the sense of agency through considering the effects of the meaning of the sound.

1.3. Levels of effort

To this day, it is still poorly understood how much the engagement in an auditory tasks influences the neuronal activity in the auditory cortex. The first anecdotal study that has shown an effect of engagement on the neurons in the auditory cortex was reported by Hubel and colleagues (1959). They discovered that the sound of a toy mouse evoked much higher responses in the neuronal activity of the cats, which were previously familiar with the sound, than some other artificial sounds that they presented to the animal. Since then, many studies described the differences of the responses to the same acoustical stimulations presented in two situations, i.e., in a situations of presence and absence of engagement. One can divide the results obtained in the studies by three categories: the results that showed higher response to the acoustical stimuli in the situation of presence of engagement, the result that showed lower response to the acoustical stimuli in situations of presence of engagements and the result that found both (or none) of the changes relative to the presence of engagement. For instance, a bunch of studies was made by one research group during the last fifteen years (Atiani et al. 2014, Fritz et al. 2003, 2005, 2007c, 2010, Lu et al. 2016). The results of their studies, belonging to the first category, showed that the response to the same acoustical stimuli varied and depended on the situation in which a subject, the ferret, was. The neuronal responses to the acoustical stimuli were high during engagement of the subject in the condition when the responses were much lower when the same acoustical stimuli were presented passively before or after the session, in the same day. Similar results were also obtained by other research groups in the primary auditory cortex of monkeys (Abolafia et al. 2011, Niwa et al. 2012a, 2012b, Scott et al. 2007). However, other research groups, belonging to the second category, showed an opposite effect, where the responses to the acoustical stimuli during the passive presentation were higher than the responses to the same stimuli in instrumental conditions. Such effect was observed in the primary auditory cortex (Abolafia et al. 2011, Lee and Middlebrooks 2011, Otazu et al. 2009, Zhong et al. 2016) and also in other sensory cortexes (Fontanini and Katz 2006, Shuler and Bear 2006). Some research groups, belonging to the third category, reported absence of any differences between responses to the acoustical stimuli in condition and passive presentation (Gilat and Perlman 1984, Hocherman et al. 1976). Lastly, one study showed that some units respond higher and some lower to the same acoustical stimuli in different situations (Abolafia et al. 2011).

The question regarding the influence of engagement is intriguing and showed controversial results in different studies. Thus, a deeper research on the topic has to be done. We delved into the analysis and found it confusing that all these studies compared two conditions in which several forms of engagement were combined. Particularly the presence of unconditioned stimuli was not considered in the experiments. However, even the combination of the acoustical stimuli and reinforcement, Pavlovian conditioning, leads to the differences in the response to the acoustical stimuli in comparisons with the passive presentation (for more details see section 1.1). Thus, the results will be, probably, more accurate when two conditions, each of which leads to reinforcement, are compared. An example of such two conditions might be a design where they differ in level of effort that a subject should make in order to get a positive reinforcement.

For instance, it is well-known that the two types of the conditions, the detection and discrimination, have different level of difficulty for a subject (Kahneman 1973, Scheich et al. 2011), where the detection is easier. Thus, the two types of the conditions might be considered as different levels of effort. It was found that the detected stimuli led to lower responses than the same discriminated stimuli in the inferior temporal cortex during a visual task (Richmond and Sato 1987, Spitzer and Richmond 1991). Also some studies were conducted that varied difficulty of the detection for animals and found that the higher difficulty of the condition led to higher responses in the primary visual cortex (Chen et al. 2008). A similar effect was observed in the neuronal activity of the secondary visual area V4 during a condition with two levels of difficulty in spatial location (Boudrea et al. 2006). It is important to note that all these studies were made outside of the auditory cortex. Therefore, the present study will seek to reveal differences between the responses to the acoustical stimuli in the primary auditory cortex that are presented in the conditions and required lower and higher levels of effort.

1.4. Aim, design and hypotheses of the present study

The three previous subsections highlighted the three forms of engagements, effects of which will be reported in the present study: the effect of sound meaning and unconditioned stimuli, the effect of sense of agency and the effect of effort level. The aim of the present study was to identify how these forms of engagement influence the neuronal activity in the primary auditory cortex.

First, we aimed to find whether the different meaning of sounds influences the neuronal activity in the primary auditory cortex, and, second, whether the changes are related to the presence of the unconditioned stimuli in the monkeys with few and with much experience. In order to answer the questions, we presented an acoustical stimulations paired with water delivery (Pavlovian conditioning) to two groups of animals, to the monkeys that were previously trained to perform instrumental conditions and to the monkeys that were not trained before. Additionally, the same two groups of monkeys were presented with repeated acoustical stimulation without water delivery. Lastly, the regular water delivery was presented to the two groups of monkeys in order to control the responses to the acoustical stimulation paired with water delivery. We hypothesized that the acoustical stimuli paired with water delivery, sounds with meaning, will evoke different responses than the same stimuli without pairing in both groups of monkeys. We supposed that the level of background activity, the baseline level, of some neurons in the group with higher experience will differ between the conditions with and without presence of unconditioned stimuli. Results for the group of the low- and well-trained monkeys are presented in the sections 3.1 and 3.2. Discussion of the results is in the sections 4.2.

The third aim of the present study was to reveal the effect of sense of agency on the neuronal activity in the primary auditory cortex of monkeys. In order to find the effect, two monkeys were trained to perform the three instrumental conditions. Trials of two of the conditions were initiated by the animals and the trials of the third conditions were initiated by a training computer. The additional condition, in which the acoustical stimulation was paired with water delivery, Pavlovian conditioning, was passive and was also initiated by the training computer. Two self-initiated and two externally-initiated conditions differed in the additional requirements to the monkeys: one of conditions in each pair required an additional effort, a detection of an acoustical stimulus when other two did not. We had a hypothesis that the response to the self-initiated acoustical stimuli will be lower than to the externally-initiated. We also supposed to find suppression in the activity after the self-initiation continuing during the acoustical stimulation. Comparison of the three instrumental and one passive condition is presented in the sections 3.3, the discussion of the results is in section 4.3. The effect of sense of agency is described in the section 3.4, the discussion of the effect is in the section 4.4.

The fourth aim of the present study was to find an effect of the level of effort of the animal to get a drop of water. In order to do this, we used the same four conditions that were described in the previous paragraph. Due to the reason that two of the conditions required detection of an acoustical stimulus and another two did not, we had an opportunity to compare the two pairs of the conditions and reveal the effect level of effort. We hypothesized to find higher responses to the sounds in the conditions with high efforts compare with the low efforts. Also we expected to observe higher increase in activity during acoustical stimulation after detection before the water delivery. Comparison of the three instrumental and one passive condition are presented in the sections 3.3, the discussion of the results are in section 4.3. The effect of the effort is described in the section 3.5, the discussion of the effect is in the section 4.5.

2. Methods

2.1. Subjects

The data for the present study was obtained from four adult cynomolgus monkeys (*Macaca fascicularis*). Two of the four monkeys (one male and one female, We and Ba) were highly trained to perform three instrumental conditions. After they learned the conditions, the neuronal activity in the auditory cortex was recorded while the monkeys performed the three instrumental conditions and, additionally, while the monkeys were presented with three passive conditions (conditions that required absence of the voluntary behavior). Two other monkeys (one male and one female, Ed and Er) were not trained to the instrumental conditions but were trained to sit quietly in the primate chair. We will name the two groups the well- and low-trained monkeys. Neuronal activity in the auditory cortex of the low-trained monkeys was recorded during the three passive conditions (without any requirements to the monkeys). The monkey Ed previously participated in a study with positive reinforcement, ~2 years before the beginning of the recordings for the present study. The monkey Er did not participated in any experiments before the present study. All monkeys were water restricted for 24 hours before trainings and before recording sessions. Experiments were carried out under approval of the animal care and ethics authority of the State Saxony-Anhalt (Landesverwaltungsamt Halle).

2.2. Conditions for the well- and low- trained monkeys

2.2.1. Equipment for the training and recording sessions

The trainings and recording sessions were carried out in an electrically shielded, sound-attenuated, double-walled room (Series 1202-A, IAC Acoustics). The subjects sat in a primate chair that allowed moving upper and lower limbs but did not allow changing location of the body. The primate chair was fixed to a stable external construction. A metal hollow tube with a groove on the tip (holding capacity of about 3 ml) was attached to the primate chair in a way, so that the monkeys could reach it with their mouth but not with their hands.

The water-restricted monkeys were reinforced by drops of water. The drops were delivered through the hollow tube and flowed down into the groove. For each session, the tube was located in such way that the monkeys could reach the drops of water from the groove using only their lips or tongue. Sizes of the drops varied from session to session between 0.3 and 1 ml. The minimum size of the drops was chosen individually for each session in the way that the monkey kept attention to the process and/or had enough motivation to perform the instrumental conditions. Delivery of liquid drops was implemented through a pump (SP200IZ Syringe Pump, World Precision Instruments) equipped with a plastic syringe. The syringe was filled with a smoothie (liquid mixture consisted of fruits, vegetables and water) or pure water (independently of the type we will further always name it "water"). The type of the reinforcement depended on the monkeys' preferences at the sessions. A training computer controlled the "on/off" switch of the pump.

In order to control the behavioral performance, a metal bar (10x2x2 cm) was fixed in the front left side of the monkeys. Grasps and releases of this bar led to an on and off switching of an electrical circuit (5V). The information about the state of the circuit was transferred to the training computer. The metal bar was fixed to the chairs only for the group of the well-trained monkeys.

In order to inform the well-trained monkeys about the type of the condition, we implemented three LEDs, which had different colors, in front of the chair. One of the LEDs was on the front left side, one in the front central position and one on the front right side of the monkey. The combination of the three LEDs indicated four conditions, i.e., three combinations for the three instrumental conditions and one, the same, combination for the three passive conditions. Thus, only one type of the combination was presented to the low-trained monkeys because they were presented only with passive conditions. With the beginning of the session, one of the combinations was turned on and the LEDs illuminated constantly until the switch to another condition.

During the trainings and recording sessions, acoustical stimuli were generated using the training computer. The signal from the computer was DA converted (DA1, Tucker-Davis Technologies), amplified (A-202, Pioneer) and presented through speakers (Karat 720.2, Canton). Two speakers were located symmetrically on the same distance (100 cm) from the left and right ears (40° from both sides) of the monkeys.

The experimental chamber was equipped with a video camera (ICD, Ikegami). Through the videos, we monitored the monkeys' behavior during the training and recording sessions. Some of the recording sessions were recorded (Pinnacle Studio 10) for further analysis of the mouth movements.

2.2.2. Three instrumental and one passive conditions for the well-trained monkeys

In order to reveal the influence of the sense of agency (the source of the acoustical stimulation) and of the effort of the subject in the neuronal activity in the auditory cortex, we invented a complex behavioral model that consists of the four conditions (Figure 1). The sense of agency was investigated by the comparison of the self-initiated (Self-) conditions with externally-initiated (Ext-) conditions. Simultaneously, the effort was studied by the comparison of the pairs of conditions that required high effort (-HighE) and low effort (-LowE). Thus, each training and recording session consisted of the four conditions: self-initiated requiring high effort (Self-HighE), externally-initiated requiring high effort (Ext-HighE), self-initiated requiring low effort (Self-LowE) and externally-initiated requiring low effort (Ext-LowE) conditions.

In the three instrumental conditions, Self-HighE, Ext-HighE and Self-LowE, the monkeys were trained to interact with the metal bar. Particularly, they were required to initiated the acoustical stimulation, and the trial, with a bar grasp in the self-initiated conditions and react with a bar release on the appearance of a pure tone in the conditions with high-effort (Figure 1). In the externally-initiated instrumental condition (Ext-HighE), monkeys were required to wait for the beginning of the trial, initiated externally by the training computer, and react with a bar grasp to the onset of the pure tone. In order to get a drop of water in the passive condition of the model (Ext-LowE), the monkeys were required to inhibit the interaction with the metal bar during acoustical stimulation. In all the four conditions, the pure tone predicted water delivery. However, the water was delivered only in case of the correct performance.

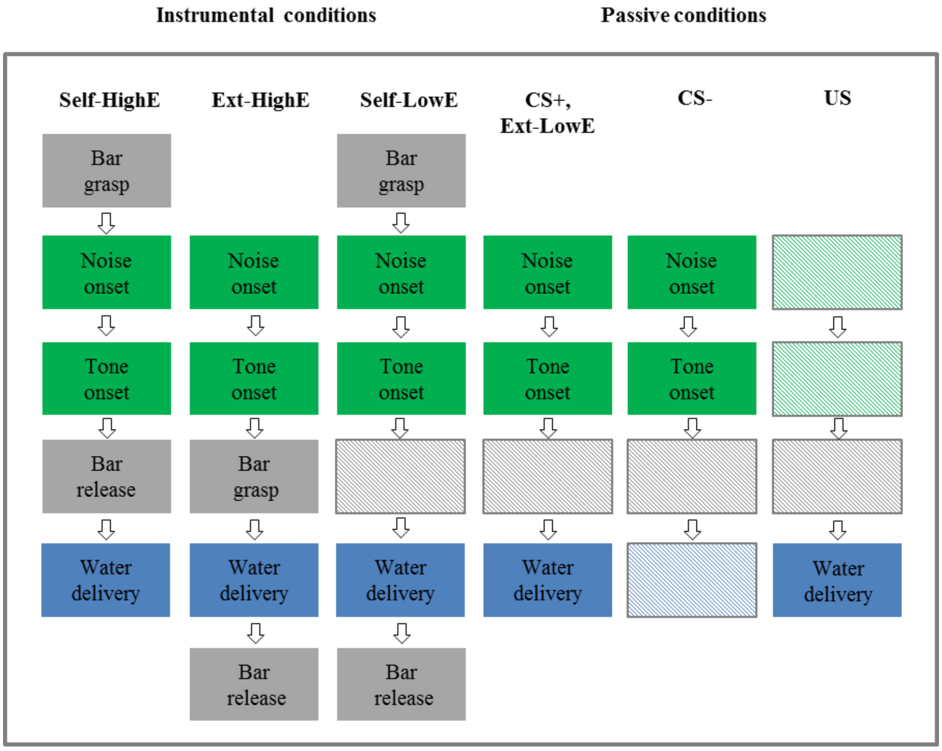

Figure 1. Scheme of the six conditions used in the present study.

The model consists of the three instrumental conditions: the self-initiated requiring high effort (Self-HighE), externally-initiated requiring high effort (Ext-HighE) and self-initiated requiring low effort (Self-LowE); and of the three passive conditions: conditioned stimuli paired with unconditioned stimuli (CS+, also named Ext-LowE), repeated conditioned stimuli (CS-) and repeated unconditioned stimuli (US). In the three instrumental conditions, the monkeys were trained to react on the tone onset with the bar grasp (Ext-HighE) or bar release (Self-HighE). In the third instrumental condition (Self-LowE) and in the passive conditions (Ext-LowE) monkeys had to actively suppress the reaction on tone onset and just wait for the water delivery. The monkeys were presented with two types of the acoustical stimuli: a noise followed by pure tone or the pure tone only. The scheme represents only one of the two types.

To exclude a possibility that the monkeys grasp and release the bar in specific intervals of the self-initiated conditions and to ensure the reaction on the tone onset in the conditions with high effort, an additional acoustical stimulus was included for each of the four conditions. The additional acoustical stimulus was a white noise. The position of the noise was always before the pure tone. Thus, two types of the acoustical stimuli could be presented to the monkeys in each of the four conditions: the pure tone only and the noise followed by the pure tone. The type of the acoustical stimuli was chosen randomly by the training computer with a probability of 0.5. The events and their durations were constructed in the way that they had the maximum correspondence between the four conditions (Figure 2).

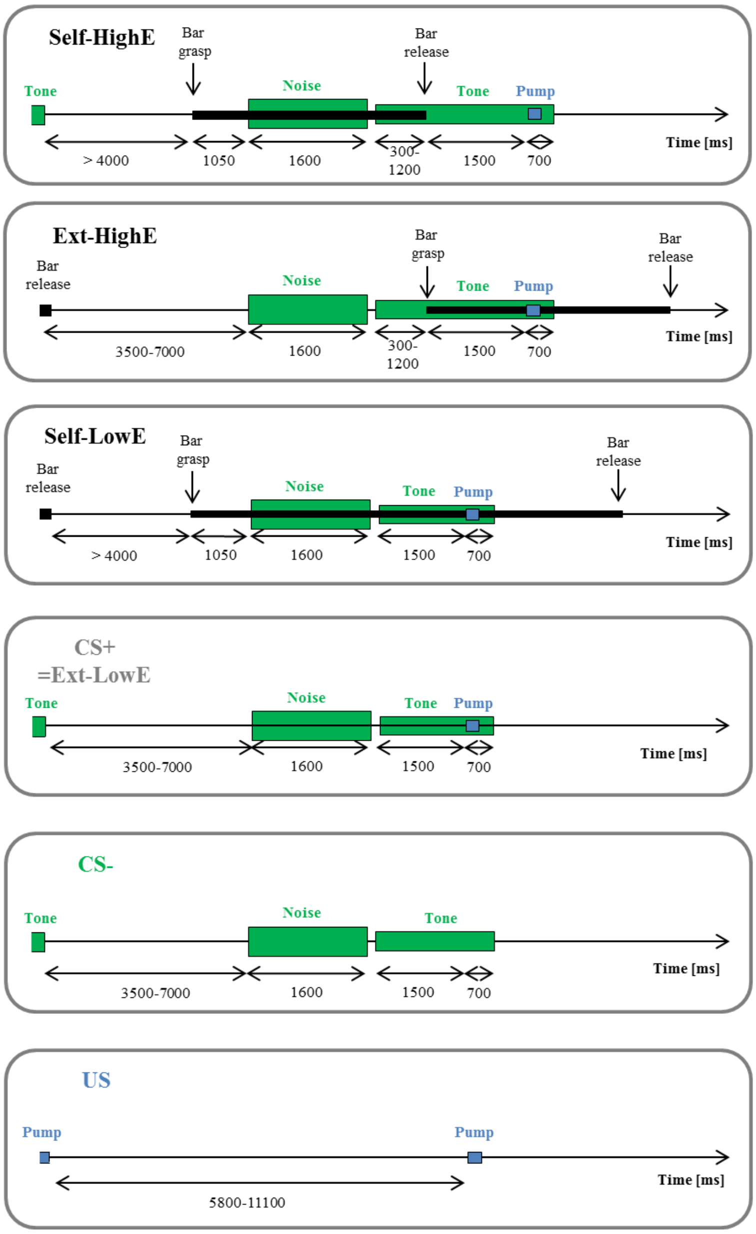

Figure 2. Time scheme of the six conditions used in the present study.

The figure shows time windows between events in the six conditions. From the two types of the acoustical stimuli presented to the monkeys, only one type, the noise followed by the pure tone, is presented in the scheme. The first brackets in the time axis of each subplot represent the last events and the end of the previous trials; the second brackets in the time axis represent the first event of the trials; the numbers between the first and the second brackets indicate durations of the pauses between trials. The time intervals between the noise offset and tone onset lasted 200 ms. The time interval between the pump onsets and pump offsets lasted 600 ms. Time intervals for the conditions with the pure tone on the first position were generally the same.

The durations of the noise stimulation and of a silent interval after were constant and lasted 1600 and 200 ms respectively (Figure 2). The duration of the pure tone stimulation varied between conditions due to their requirements and also depended on the reaction time in the conditions with high effort. In all four conditions, the water was started to be delivered with the onset of the pump that lasted 600 ms. Thus, the water was delivered during the pure tone stimulation; the time between the pump onset, the water delivery, and the tone offset was 700 ms.

In the self-initiated conditions with high effort (**Self-HighE**), the monkeys had to initiate the trial with a bar grasp and wait until the onset of the acoustical stimulation (Figure 1, Figure 2). If the first acoustical stimulus was the noise then the waiting time was 1050 ms, if it was the pure tone then the waiting time was 750 ms[1]. If the acoustical stimulation began with the noise, the monkeys were required to keep holding the bar and wait until the beginning of the pure tone. During stimulation of the pure tone, the monkeys were required to release the bar in a time window of 300-1200 ms after the onset. If the trial was performed correctly, water was delivered to the monkeys 1500 ms after the bar release. Thus, the duration of the pure tone varied between 2500-3400 ms in successful trials. In 4000 ms after the water delivery, the monkeys could initiate a new trial. In case of an earlier grasping, the onset of the acoustical stimulation was postponed and the next trial could be initiated in 4000 ms after the bar release. In the case of the absence of the bar release during the stimulation of the pure tone, the tone was terminated 1800 ms after the onset and the water was not delivered. In the case, when the monkeys released the bar before the required time window or before the onset of the pure tone, the acoustical stimulation decayed immediately and monkeys were required to wait 4000 ms before the beginning of the new trial.

In the externally-initiated conditions with high effort (**Ext-HighE**), the acoustical stimulation was initiated externally and the monkeys had to wait for it and to detect the appearance of the pure tone with a bar grasp (Figure 1, Figure 2). After the detection of the pure tone, the monkeys had to keep holding the bar 1500 ms until the water delivery. Similar to the Self-HighE condition, the reaction time of the pure tone was required to be 300-1200ms. In the case of the correct performance, the pump was turned on and a drop of water was delivered. Thus, the duration of the pure tone stimulation varied between 2500-3400 ms in successful trials. In the case of the absence of a reaction to the pure tone, the tone decayed 1500 ms after the onset. The intertrial intervals varied between 3500-7000 ms and began after the bar release. If the bar was grasped during the noise stimulation or released before the water delivery, the acoustical stimulation decayed immediately and the intertrial interval began after the bar release. In case the monkeys grasped the bar during the intertrial interval, the new intertrial interval started after the releasing.

In the self-initiated conditions with low effort (**Self-LowE**), monkeys had to initiate the trial with a bar grasp and wait until the onset of the acoustical stimulation (Figure 1, Figure 2). Similar to the Self-HighE condition, the time window before the acoustical stimulation varied and was 750 ms before the pure tone onset and 1050 ms before the noise onsets[1]. The monkeys were required to keep holding the bar until the water delivery. The water was delivered 1500 ms after the onset of the pure tone. Thus, the stimulation of the pure tone lasted 2200 ms in successful trials. The intertrial interval began after the bar release and lasted 4000 ms. After the intertrial interval, the monkeys could initiate a new trial. In the case of a

---

[1] The differences between the time windows were unavoidable due to technical reasons.

bar grasp during the intertrial interval, the beginning of the new trial was postponed 4000 ms after the bar release. If the monkeys released the bar before the water delivery, the acoustical stimulation decayed immediately and the monkeys had to wait 4000 ms before they could initiate the next trial.

In the externally-initiated condition with low effort (**Ext-LowE,** also named **CS+** when compared with passive conditions) monkeys had to wait only for the water delivery (Figure 1, Figure 2). The water was delivered 1500 ms after the onset of the pure tone. The duration of the pure tone was 2200 ms in successful trials. As in the Ext-HighE condition, the intertrial interval lasted between 3500 and 7000 ms. If the monkeys grasped the bar during the acoustical stimulation, it decayed immediately and a new intertrial interval started after the release of the bar. Unlikely to the three instrumental conditions (Self-HighE, Ext-HighE and Self-LowE), the condition Ext-LowE was not presented during the training sessions. The first presentation of this condition was in the first recording session. Despite the passive Ext-LowE conditions was a classical Pavlovian conditioning, we expected that the highly-motivated well-trained monkeys will try to interact with the metal bar. Because the monkeys learned to inhibit the reaction in each session (see Results), we considered the condition as a passive condition but not as Pavlovian conditioning.

A typical recording session of the well-trained monkeys consisted of the three instrumental and one passive condition. The frequency of the pure tone varied from session to session and was adjusted to the best frequency of the recorded units. We aimed to record 80-120 successful trials in each of the four conditions (40-60 trials for each type of the acoustical stimuli). The order of the three instrumental conditions was randomized in each session. During one session, each condition could be repeated from one to three times in order to optimize behavior of the monkeys. Usually, the passive condition was presented after the successful performance of all three instrumental conditions. Thereby, we prevented refusal of the monkeys to perform the three instrumental conditions. However, in some session we successfully presented the passive condition on the first place or in between of the three instrumental conditions. The recording sessions, in which monkeys refused to perform at least one of the three instrumental conditions, were interrupted and the neuronal activity recorded during the session was not used for the analysis.

Monkeys were informed about the type of the condition with the onset of the LEDs' combination. Despite to the successful performance of the instrumental conditions, the monkeys did not pay attention to the colors of the LEDs but when their state switched. After the LEDs blinked, the monkeys tried all their techniques that they have learned doing the instrumental conditions until the water delivery. After that, they continued to perform the rewarded condition. With this approach, the switches of behavior from one condition to another took no longer than 10 trials and were ~5 trials in average.

2.2.3. Two additional passive conditions for the well-trained monkeys

In the four conditions, the pure tones had a meaning of the water delivery for the monkey if they performed a trial correctly. We wondered whether the same acoustical stimuli that differed in meaning, with the absence of the water prediction, would evoke different activity in the auditory cortex. Therefore, monkeys were presented with the same acoustical stimuli without pairing with water delivery (**CS-,** Figure 1**,** Figure 2). Because the monkeys were well-trained and also highly motivated, they tried to perform the instrumental conditions when they heard the acoustical stimuli. Any interaction of the monkeys with the bar led to the immediate decay of the acoustical stimulus and the acoustical stimulus was presented only in 3500-7000 ms after the release of the bar.

The water delivery was an important event in each of the four conditions. We were wondering whether the unconditioned stimuli itself changed the neuronal activity in the auditory cortex. To control the effect of unconditioned stimuli, we made an additional passive condition for the well-trained monkeys. Basically, it was a regular presentation of unconditioned stimuli (**US,** Figure 1**,** Figure 2). In this condition, a drop of water was delivered with a time interval of 5800-11100 ms. All bar grasps interrupted the interval and a new one began after the release of the bar.

As well as in the four described conditions, an additional acoustical stimulus, the noise, was included in the two passive conditions, CS- and US. For the US conditions it was included as an additional time window between the water deliveries. The probability and duration of the noise stimuli, and the silent interval after were the same between all conditions.

Unlike the passive CS+ condition, the CS- and the US passive conditions were included only in some of recording sessions. The two passive conditions were presented only after successful performance of the three instrumental conditions. The order of the CS+, US and CS- conditions was randomized from session to session. The change from one to other passive conditions could be repeated from one to three times during one session. In each session, where the two additional passive conditions were presented, we tried to collect 80-120 trials (40-60 trials of the each type of acoustical stimulus) during which the monkeys did not interact with the metal bar.

2.2.4. Three conditions for the low-trained monkeys

In order to control the influence of the previous training on the neuronal activity of the well-trained monkeys and to find the real effect of Pavlovian conditioning and of the factors separately, three passive conditions CS+, CS- and US were presented to the low-trained monkeys (Figure 2). The low-trained monkeys did not have any task requirements during presentation of these three conditions and were not familiar with these conditions before the recording sessions.

Each recording session of the low-trained monkeys was composed of the three passive conditions. The order was randomized from session to session. One condition could be presented from one to two times during one session. In total, each condition was presented 80-120 times (40-60 trials of each type of acoustical stimulus).

2.3. Electrophysiology

2.3.1. Animal preparation

Each monkey was implanted with a headholder, which enforced head restraint, and with a recording chamber. The headholder was fastened to the monkeys' skulls by 4-6 bolts with sharpened points (Brosch and Scheich 2008). The recording chamber was screwed over the left or the right side of the auditory cortex and was centered on the A10 and D15 coordinates (Szabo and Cowan 1984). The two well-trained monkeys (We and Ba) received the chamber on the right side and the two low-trained monkeys (Er and Ed) had it on the left side.

Before the implantation, the animals were anesthetized with mix of Ketamine HCl and Xylazine HCl. A full course of antibiotic (Enrofloxacin) and analgesic (Carprofen) treatment followed the implantation.

2.3.2. Recording system and recording process

Two types of microdrives and head fixations were used in the present study. The first, standard, type was used for all four monkeys and required stark fixation of the head-holder based on the scull screws. In order to minimize traumatization of the animals, we established a second fixation type for one of the well-trained monkey (We).

In the standard fixation, the monkeys were fixed to the external stable construction through an additional screw in the head-holder (Figure 3). The 5-channel microdrive (Eckhorn System, Thomas Recording), with a built-in preamplifier, was attached to a manipulator. The manipulator was also fixed to the external stable construction. Quantity of the electrodes in the microdrive varied from 3 to 5 for different recording sessions. The electrodes had an impedance of 1.5-2.5 MΩ, were made from tungsten and coated in glass.

In the second fixation type (Figure 4A, B, D) the head of the monkeys was fixed by a custom-made plastic box with a rectangular opening on the front wall. The opening was needed for the protruding part of the monkey's skull, its snout. Using the opening part, the monkeys could move its mouth and reach drops of water. Four walls and the upper face of the opening touched and gently restricted the monkeys' head in three dimensions. The microdrive (MEM 10, Thomas Recoding) was fixed on the recording chamber through a custom made adaptor (Figure 4C). The adaptor was consisted of a chamber holder and chamber clamp (not shown in the figure). The adaptor was used for the x-y-z positioning of the microdrive with a multitrode. The multitrode (Multitrode for MEM 10, Thomas Recording) was fixed on the microdrive. The multitrode had 8 gold contacts from which seven were located along the length (impedance 1.5-1.9 MΩ, diameter ~40 µm) and one was on the tip (impedance 0.5 MΩ) of the core. The distance between the tip contact and the first of the seven contacts was ~700 µm; the distance between the other contacts was 125 µm. The diameter of the core was 300 µm. The microdrive allowed the multitrode to be moved in a range of 0-10000 µm. The multitrode was connected to the external custom-made preamplifier (PA-08, Thomas Recording). The pre-amplification and filtration of the signal were designed to be equal to those of the 5-channel microelectrode system.

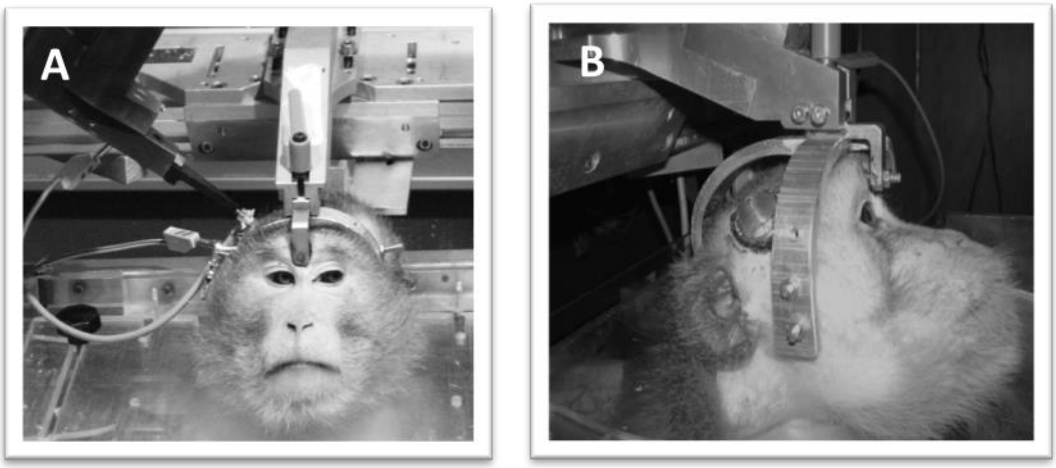

Figure 3. Design of the standard head fixation.

In these photos, the monkey sat in the primate chair. (A) and (B) show the system from the front and right sides of the monkey respectively. The chair and the implanted head-holder were fixed in the same external stable construction. For more details, see Brosch and Scheich 2008.

After the pre-amplification, the signal was amplified and band-passed filtered between 0.5- 5 kHz (PGMA, Thomas Recording). The filtered signal passed through a switcher (Distribution Box, Thomas Recording) and was collected by the recording software (Cheetah Digital Interface DT3010, Neuralynx). The signal was collected with a sampling rate at 44.1 kHz. The software recorded the forms of the spike if their amplitude exceeded a manually established threshold. The threshold for each recording channel was selected individually and varied from day to day but remained constant during one session. Voltage changes of the pump, of the metal bar and of the DA converter were used as markers for the grasps/releases of the bar, the pump onsets/offsets and of the onsets/offsets of the acoustical stimuli respectively. These signals were also connected to and collected by the recording software.

Before each recording session, we defined the best frequency of the neurons. For this, 40 pure tones in the frequency range 0.1-27 kHz were generated by the training computer and presented through the speakers. Each pure tone lasted 100 ms and was repeated 10 times. The silent interval after the pure tones lasted 900 ms. The order of the presented pure tones was randomized. The best frequency of the neurons was used for the determination of positioning. Additionally, the knowledge of the best frequency allowed us to select the frequency of the pure tone, which was used in the conditions for the well- and low-trained monkeys, in the way that it activated as many neurons as possible during a recording session.

Most of the recordings were made in the primary auditory cortex. The positioning was identified using the location of the lateral sulcus, recording depth and tonotopic gradient (Kaas and Hackett 2000).

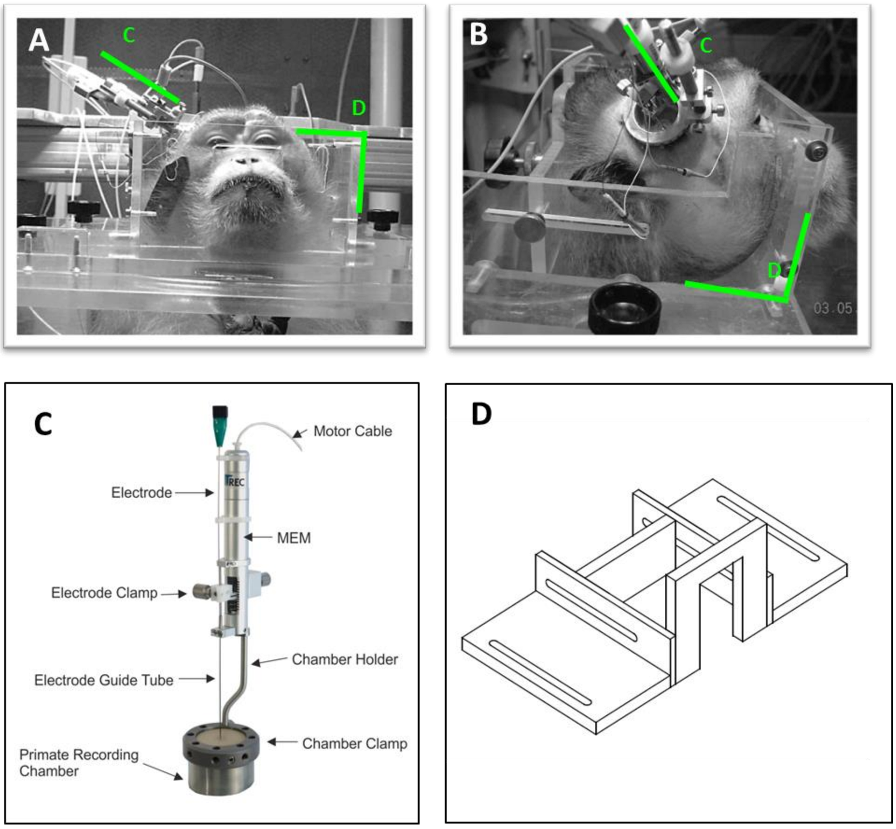

Figure 4. Design of the new head fixation (A, B and D) and of the microdrive adjustment (C).

The monkey sat in the primate chair; a plastic box (D) was placed around the head, was fixed on the chair and prevented the monkey's head from movements but allowed them to move the mouth. The photo from the front and from the right sides of the monkey are shown in (A) and (B) respectively. (C) The microdrive with multitrode was fixed on the recording chamber. The adaptor consisted of the chamber clamp, a manipulator (not shown in the scheme) and the chamber holder. The set allowed regulation of the x-y-z-positions of the multitrodes.

2.4. Data analysis

2.4.1. Preparation of data

All data analyses were implemented with custom written scripts in MatLab (2007b_1 and 2016a, MathWorks).

Before the analysis of the neuronal activity, we controlled each trial of each session on the presence of the short-termed electrical artifacts evoked by movements of the animal or other reasons. The trials with artifacts were not collected for further analysis. Only correctly performed trials were used for the analysis of the neuronal activity in the present study. After the cleaning, the quantity of the trials for each condition was on average ~49 (for more details see Supplementary table 1).

Forms of the action potentials (spike forms) recorded in all selected trials of one session were also tested and noise was removed (see also Bondar' et al. 2014, Schanze 2017). For this, they were transformed into principle components. After that, we manually saved all separated clusters, which had spike forms. After the back transformation, we revealed that some of them were single cells, single units, but most of them were pools of neighboring cells, multiunits. The present study reports about results observed on both, single and multi-units, that will be named “units”.

Each unit was analyzed for stability during a recording session. For this, we calculated peristimulus time histograms (PSTHs), i.e., the averaged neuronal activity (with bin size of 100 ms) of all trials of a certain condition. The PSTHs for each condition of each unit were compared and the units, where the baselines of the PSTHs largely between the conditions. If the levels of the baselines repeated the order of the conditions in a session then the unit was excluded from further analysis.

Having two types of acoustical stimuli in the conditions, we presented the results of the analysis of the type “noise followed by tone” only. We made this decision based on higher application of this type. The reasons were the possibilities to separate two meanings of the acoustical stimulation, i.e., (1) the beginning of a trial and prediction of the water delivery in the conditions, (2) the beginning of a trial and the cue for detection in the conditions with high effort, (3) self-initiation of a trial and prediction of the water delivery in the self-initiated condition with low effort. Thus, the stimulation type “noise followed by tone” was more informative for the analysis and allowed us to reveal the real effects of meanings and to exclude their mixes. We also analyzed neuronal activities in conditions with “tone only” informally and revealed a mixture of the differences that were found during the noise stimuli and tone stimuli in the type “noise followed by tone”. Thus, a clear view of the effects, which potentially can influence the neuronal activity, is necessary before the detailed analysis of the differences between the conditions for the type “tone only”.

Before all analysis, the firing rate of all units was normalized to be between 0 and 1. The maximal and the minimal values of firing rate of a unit during full recording session were identified in the PSTHs of all conditions. Further, each value of the PSTHs of the unit was recalculated by the next equation:

$$Xs = \frac{Xi - Xmin}{Xmax - Xmin},$$

where Xi was the initial firing rate of a bin, $Xmax$ was the maximal and the $Xmin$ were the minimal values of the all PSTHs of the unit, Xs was the new relative firing rate of a bin.

2.4.2. Common tests used for the analysis of the neuronal activity

Each unit was tested for the presence of the neuronal response to the acoustical events, i.e., to the noise and tone onsets and offsets. For this, we compared average activity during 100 ms before and after each of the acoustical events in all trials of the condition of a unit. We accepted the presence of a response when the result of the significance level of the Wilcoxon signed rank test did not exceed 5%. The same method was used for the detection of the responses to the movements, i.e., to the grasps and releases of the metal bar.

For the comparisons of the neuronal activity between the conditions in one separate unit, the average activity of a certain time window, which will be mentioned in each case, was calculated for all trials of the compared conditions using the Wilcoxon signed rank test. Due to the test required equal numbers of variables, we chose and compared the smallest number of trials of the two available. The result of the comparison was accepted significant when did not exceed 5%.

In order to exclude influence of the shift of the baseline between conditions and calculate a real change of the firing evoked by the acoustical events, we implemented the next calculation. First, the average activity during 100 ms after an acoustical event was divided by the average activity during 100 ms before the event. After the procedure in the two compared conditions, the two resulting values were subtracted. The median of the resulting values was compared with zero using the Wilcoxon signed rank test. The median was accepted as significant different from zero if it did not exceed the level of 5%. In order to conduct similar comparisons for each unit of the sample and to reduce the level of noise in the results, we averaged activity during 100 ms before an event between all trials. The responses evoked by the acoustical events (during 100 ms) of each trial were divided to the averaged value before the event. The analysis of the median was performed in the same way as for population neuronal activity.

In order to exclude influence of the shift of the baseline between conditions and calculate a real change of the firing during acoustical stimulations, we calculated the *z*-score for each 100 ms bin of the neuronal activity. The calculations were made separately for each unit and each condition using the formula:

$$z = \frac{x - \mu}{\sigma},$$

where μ and σ represent the mean and the standard deviation of the neuronal activity during 3000 ms before the noise onset. After that we found the differences between the averaged z-scored neuronal activities during the noise stimulation (200-1600 ms after onset) or during the tone stimulation (during 1400 ms before the water delivery) in the conditions. Similar analysis was made for individual comparisons between the trials and conditions.

In order to reveal units with slow modulations in the neuronal activity, we measured how the relative firing changed within 2500 ms after the water delivery (or within comparable time window for the condition without water delivery, CS-). The time window was divided into 25 segments. The row of the numbers was tested using the Spearman’s rank correlation coefficient. The units with the significant increase or decrease in activity ($r > 0$ or $r < 0$ and $p < 0.05$ controlled by t-test) were defined as units with the slow modulations.

2.4.3. Additional tests of the neuronal activity for the analysis of the sense of agency and level of effort

The second part of the study had a purpose to reveal changes of the neuronal activity that occurred with the change of the sense of agency or of the levels of effort. In order to find the effects of each of the two factors in the population level, we averaged the activities in the two relevant conditions for each unit and compared in specific time windows (will be mentioned in each case) using the Wilcoxon signed rank test. The result was accepted as significant if it did not exceed the level of 5%. After that, the neuronal activities were compared within the same time windows between the pairs of conditions with opposite factors using the Wilcoxon signed rank test. The result was accepted as significant if did not exceed the level of 2.5% (Bonferroni corrected). The effect was accepted as reliable if all three comparisons showed the same directions and were significant. In order to find the effects of the factors in the individual units, the same three steps were conducted between trials in each condition of a unit.

2.4.4. Behavioral analysis (number of errors and frequency of mouth movements)

In order to control monkey’s behavior, the numbers of errors (false alarms, missed or false initiations) were divided by the total amount of successive trials and errors in each condition in each session. The resulting values were averaged between the sessions.

All recorded videos were analyzed in order to reveal momnets of the water collection, i.e., mouth movements of the monkeys. Using the video, we could observe and detect the moments of protruding of the tongue, lip folding, smacking and other movements. In order to register the movements, we, first, found the exact time of the video, where the monkeys did not perform any movements of the mouth at least for some seconds. Using MatLab (Version 2016a), we extracted one frame of this time, “the rest frame”, and selected there two areas. The first area, the “mouth area”, had to include the snout of the monkey and the end of the metal tube where the water flowed out. The second was a “background area” fixed on a static object. The background area was used for the noise correction. Each of the two areas had size of 60 x 50 pixels. The size was optimal for including the mouth, but excluding the eyes, arms and non-movable parts of the video in the “mouth area”. The RGB colors of the pixels were transformed into the gray scale that allowed working only with intensities. Thus, each area had 3000 pixel values of intensities. Further, we calculated the differences of the sets of the intensities between the “mouth areas” of the “rest frame” and each frame of the video, taking into account positions of the pixels. The same manipulation was made with the “background areas”. The resulting values of the subtractions were converted into distributions. Thus, we had two distributions for each frame of the video. Lastly, we subtracted means of the two distributions in each frame. The obtained values for each frame were further reconstructed accordingly to their order in the video. The resulting curve was filtered using rational transfer function with window of 100 values. The initial and the filtrate of the initial curves were also subtracted. In order to get only positive values after the subtraction, we squared the obtained values. The resulting values were averaged by the known time triggers of the trials for each conditions and each session.

3. Results

3.1. Presence of unconditioned stimuli changed neuronal activity in the low-trained monkeys

The present section will describe the dependence of the neuronal activity in the primary auditory cortex of the low-trained monkeys on the sound meaning and on the unconditioned stimuli alone. In order to reveal the changes, we presented three passive conditions to the low-trained monkeys. To give the meaning to the acoustical stimulations, the last were paired with water delivery (CS+, Figure 1 and Figure 2). Animals were water deprived that insured that the drop gave the meaning to the acoustical stimulation. A drop of water was delivered shortly before the offset of the pure tone. Before or after the CS+ condition, another condition (CS-) was presented for the control of how the neuronal activity will change when the acoustical stimuli had no meaning. The acoustical stimuli in the CS- condition were the same but were not paired with the water delivery. Lastly, in order to control the effect of the unconditioned stimuli, a third condition (US) was presented before, after or between the CS- and CS+ conditions. In the US condition, one drop of water was regularly delivered to the monkeys.

Three conditions were presented to two monkeys with little experience. One of the monkeys was never trained to perform any auditory tasks. Another monkey participated in another experiment with an auditory task but the training was terminated two years before the beginning of the current experiment. Additionally to the lack of experience, the analysis of the monkeys' behavior revealed a lack of learning from the first to the last sessions where the three conditions were presented (see section 3.1.4). Therefore, we will name the two animals "low-trained" monkeys. For the present study, we used 37 (nine recording session) and 38 units (nine recording sessions) recorded in the auditory cortex of the two monkeys (Supplementary table 1). The neuronal activities were similar between the monkeys and we combined the two samples into one. We identified that the first spike latencies of these units were 16.7± 11.9 ms after the onsets of the pure tones. All 75 units of the sample responded to at least one type of the acoustical stimulus, to the onset of the pure tone or of the noise, or to the offset of one of them.

3.1.1. Responses evoked by the acoustical events in the CS+ conditions were higher than in the CS- conditions but only due to the shift in the baseline

This subsection will control whether the responses evoked by the acoustical stimuli with meaning will differ from the responses evoked by the acoustical stimuli without meaning in the neuronal activity of the low-trained monkeys.

First, we analysed the average activities of the 75 units that were recorded during presentation of the CS- and CS+ conditions (Figure 5, Supplementary table 2A). In both conditions, the firing rate sharply increased after the acoustical events: the noise onset and offset, the pure tone onset and offset. Comparison of the firing during 100 ms after the events revealed that the responses to the noise onset, noise offset and to the tone offset were significantly higher in the CS+ condition compare with the CS- (Wilcoxon signed rank test, $p = 0.02$, $p = 7*10^{-5}$, $p = 0.02$). Only the responses to the tone onsets did not differ significantly between the conditions ($p = 0.15$). However, it is important to note that the baseline in the CS+ condition was also shifted upwards and could be a reason of the differences between the responses.

Additionally to the analysis of the responses in population level, we compared the responses of 75 individual units between the two conditions. For this, the responses to the acoustical events (during 100 ms after the events) in all trials of one condition were compared with the responses of other condition. The analysis revealed that ~20-22% of the units responded significantly higher to the four events in the CS+ condition (Supplementary table 3A). Responses of the most units, ~69-78%, did not differ significantly between the conditions. A smaller fraction of units of ~5-11% in the population had significantly higher responses to the acoustical events in the CS- condition. However, the differences in the levels of baselines between the conditions were again omitted.

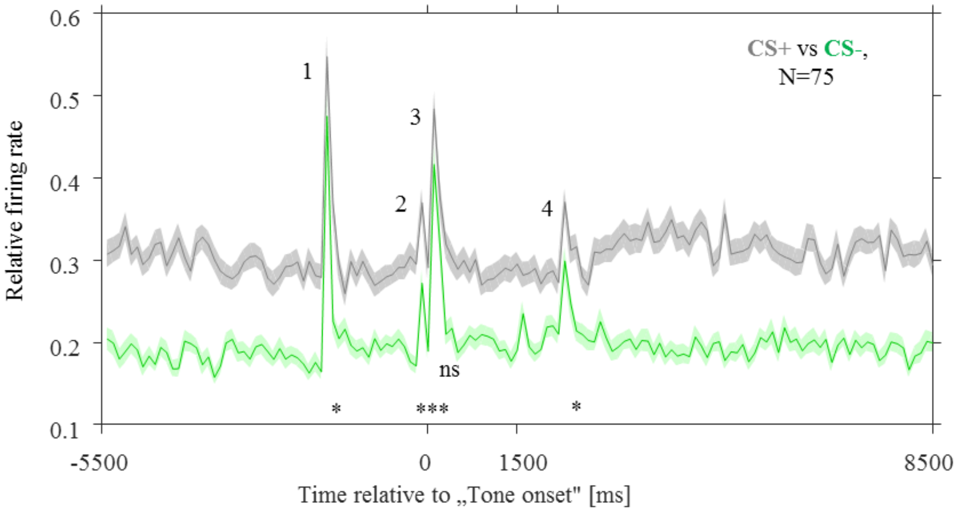

Figure 5. Neuronal activities in the CS+ and CS- condition.

The plot shows average PSTHs (bin = 100 ms) with SEM (shadow) of the population of 75 units during presentation of the CS+ (gray color) and CS- (green color) condition. In the CS+ condition, the water was delivered 1500 ms after the tone onset. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. Stars indicate significant differences between the neuronal responses to the acoustical events in the two conditions (Wilcoxon signed rank test, ns - $p > 0.05$, * - $p < 0.05$, *** - $p < 0.001$).

The average baseline levels of the population activity (during 1000 ms before the noise onset) differed significantly between the CS- and CS+ condition (Wilcoxon signed rank test, $p = 1*10^{-8}$, Supplementary table 2E). We calculated the shift between the neuronal activities in the two conditions in each unit and revealed that the baseline level in the CS+ condition rose by 47% in 75% of the units. The baseline level of the rest 25% units descended on average by 12%. Considering that the majority of the units had higher level of the baseline in the CS+

condition and that the level of the baseline was similar during the time of the trials, we assumed that the significance of the differences between the responses to the acoustical events was only a consequence of that shift.

To control this assumption, we designed a different approach where, instead of comparisons of the absolute responses, we compared the changes of the firings evoked by the acoustical events. In order to calculate the change of the neuronal activity in one condition, we found the ratio of the value of the response (100 ms after the event) and the neuronal activity before the response (100 ms before the event) in each unit. After that, we subtracted the two obtained values between the conditions and did so for each of the four events in each of the 75 units. The resulting distributions of the 75 values for each of the four events were further compared with zero (Figure 6). The analyses revealed that all four medians of the distributions did not differ significantly from zero (Wilcoxon signed rank test, p = 0.07, p = 0.38, p = 0.11 and p = 0.98, Supplementary table 2B) that indicates that the changes of the neuronal activity were similar in the two conditions. Therefore, we accepted the hypothesis that the differences between the responses to the acoustical events occurred only due to the upward shift in the baselines in the CS+ condition.

Between the 75 units, we found ~9% of units, in which the changes of the neuronal activity evoked by the noise onset in the CS+ condition were significantly larger (Supplementary table 3B). A greater amount of units, ~31%, showed that the changes were smaller in the CS+ condition. The percentages of units, in which the changes of firing evoked by the noise offset were significantly larger and smaller in the CS+ condition, were 13 and 21% respectively. Percentages for the tone onset and tone offset were 16 and 31%, and 21 and 23% respectively.

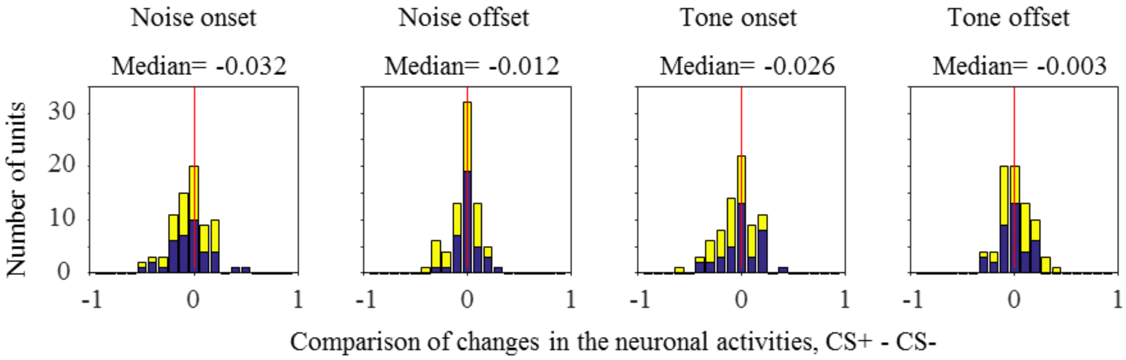

Figure 6. Comparisons of the neuronal activity evoked by the acoustical events and adjusted to the baseline levels in the CS+ and CS- condition.

Each subplot shows number of units in the population that had larger change of firing evoked by the acoustical events in the CS+ (comparison of changes > 0) or CS- condition (comparison of changes < 0). The values of medians for each distribution are inserted into the boxes. Blue and yellow parts of each bar in each histogram indicate number of units recorded in Er and Ed respectively (ratio of units between the monkeys was ~1:1).

Neuronal activity of some units did not differ in the levels of baselines between the CS+ and CS- conditions. Interestingly, that some of these units showed much higher absolute responses to the acoustical events in the CS- condition. An example that supports our assumption is presented in Figure 7.

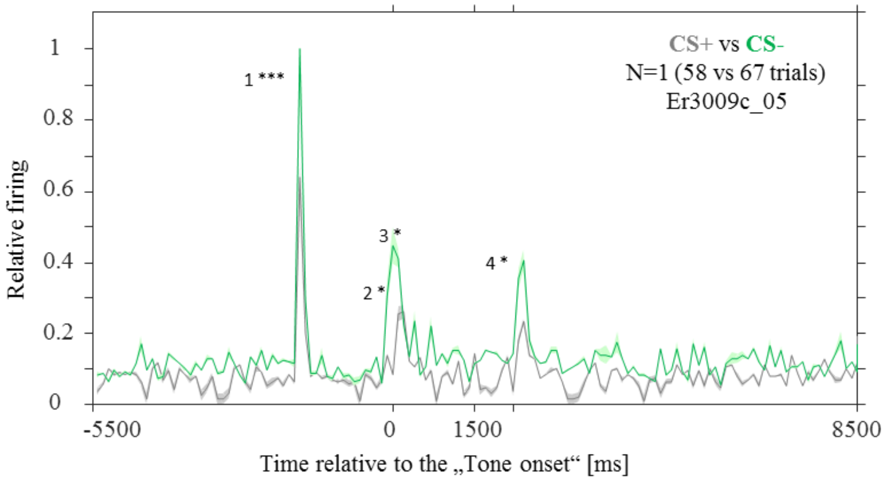

Figure 7. Example of one unit where the baseline levels were similar but the responses to the acoustical events were higher in the CS- compare with the CS+ condition.

The plot shows average PSTHs (bin = 100 ms) with SEM (shadow) of one unit during presentation of CS+ (gray color) and CS- (green color) conditions. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. The stars indicate significance of the differences between the average responses during 100 ms after the events (Wilcoxon signed rank test, * - $p < 0.05$, *** - $p < 0.001$). Note that the response to the tone onset is contaminated with the response to the noise offset in the CS- condition.

As it was expected, we found that the responses to the acoustical events with and without water delivery differed significantly in the conditions. However, the main reason of the differences was due to the shift in the baseline. Considering the baseline, we found that the changes in neuronal activity evoked by the acoustical events were similar. Thus, we concluded that the responses evoked by the acoustical stimuli with and without meaning did not affect the neuronal population activity in the low-trained monkeys.

3.1.2. Neuronal activity during the tone stimulation adjusted to the baseline was higher in the CS- condition

In this subsection, we controlled whether the neuronal activity during the acoustical stimulation depended on the meaning of the sounds in primary auditory cortex of the low-trained monkeys.

In order to reveal the changes, we compared the neuronal activities recorded during the acoustical stimulations between the two conditions, CS+ and CS-. First, we compared the average responses during the acoustical stimulation. As it can be already seen from the neuronal population activity (Figure 5), the average neuronal activities during the noise (200-1600 ms after onset) and tone (200-1500 ms after the onset) stimulations in the CS+ conditions were significantly higher than in the CS- conditions (Wilcoxon signed rank test, p

= $3*10^{-7}$ and p = $4*10^{-6}$, Supplementary table 2C). However, the effect was achieved due to the shift in the baseline.

The individual analysis of the 75 units revealed that ~54% of them had higher activity during the noise stimulation and during the tone stimulation in the CS+ condition (Supplementary table 3C). Only ~4-7% of the units had higher activity during the acoustical stimulations in the CS- conditions.

As it was reported in the previous subsection, the baselines of many units were shifted upwards in the conditions with presence of unconditioned stimuli (this will be also described in detail in the next subsection). In order to consider the baseline shift, we compared the change of the average activity during the acoustical stimulations. For this, we normalized the neuronal activity in each unit and each condition by level of the baselines during 3000 ms (*z*-scores). After that, we subtracted the average *z*-scored neuronal activity during the noise stimulation (200– 1600 ms after the noise onset) between two conditions in each unit. From the resulting values, we built a histogram and calculated its median (Figure 8, Supplementary table 2D). The same procedures were made for the average *z*-scored neuronal activity during the tone stimulations (200-1500 ms after the tone onset). The analysis revealed that the median calculated for the noise stimulation was not shifted from zero significantly (Wilcoxon signed rank test, p = 0.11). The median calculated for the tone stimulation was significantly different from zero (p = $3*10^{-3}$) and was negative. The negative value indicates that the changes in the neuronal activities of many units were larger in the CS- condition.

Between the 75 units, we found ~13% of units, in which the changes of the activities during the noise stimulation were significantly larger in the CS+ condition (Supplementary table 3D). Similar amount of units, ~15%, showed that the changes were larger in the CS- condition. The percentages of units, in which the changes of activities during the tone stimulation were larger in the CS+ and CS- conditions, were ~9 and 12%. One of the units that had higher average activity during the tone stimulation is presented in the Figure 9.

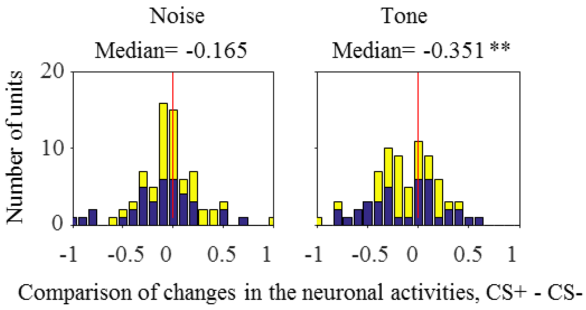

Figure 8. Comparisons of the neuronal activities evoked during the acoustical stimulations and adjusted to the baseline levels in the CS+ and CS- condition.

The plot shows number of units in the population of 75 units that had larger change of activity during the acoustical stimulation in the CS+ (comparison of changes > 0) or CS- condition (comparison of changes < 0). The value of the median of each distribution is inserted into the boxes. The stars near the median values indicate the significance level of difference from zero (Wilcoxon signed rank test, ** - p < 0.01). The blue and yellow parts of each bar indicate the number of units recorded in the monkey Er or Ed respectively (the ratio of the units between the monkeys was ~1:1).

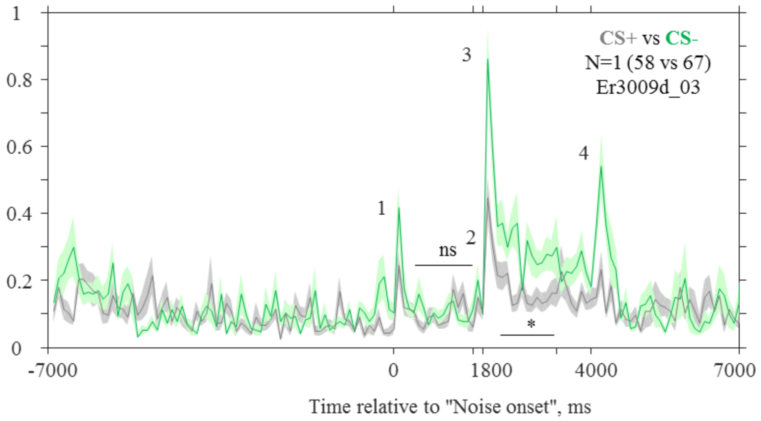

Figure 9. Example of unit with higher activity during the tone stimulation in the CS- condition compare with the CS+ condition.

The plot shows average PSTHs (bin = 100 ms) with SEM (shadow) of one unit during presentation of CS+ (gray color) and CS- (green color) conditions. Numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. The star indicates significance of the difference between the average activities during the acoustical stimulations between the two conditions (Wilcoxon signed rank test, ns - $p > 0.05$, $p < 0.05$).

3.1.3. Absence of units with slow increase (or decrease) related to the expectation of the water delivery

Previous studies showed that some of neurons in the auditory cortex had slow increase or decrease in activity when acoustical stimuli were associated with oncoming reinforcement (Abolafia et al. 2011, Fritz et al. 2010, Quirk et al. 1997, Shinba et al. 1995). The effect was observed only in the activity of animals that were previously trained to some instrumental conditions and that expected the presence of the reinforcement after the acoustical stimuli. We also controlled the presence of the slow changes in the neuronal activity during the acoustical stimulations before the water delivery in the conditions CS+. Our advantage over other studies was the possibility to control the presence of the slow changes in the neuronal activities during acoustical stimulation in the CS- condition and before the water delivery in the US on purpose to exclude the modulations induced by the acoustical stimulation itself and by the unconditioned stimuli itself respectively.

In order to reveal the slow sustained increase/decrease in the neuronal activity during the noise stimulation, we used the time window 200-1600 ms after the onset. For the tone, we considered the time window 200-1400 ms after the tone onset. In each unit, we controlled whether the activity increased (Pearson correlation coefficient, $r \geq 0.5$) during the time of the acoustical stimulation or during the relative time window in only one of the conditions but did not increase ($r < 0.5$) in the other two. Similar procedure was used for control the presence of the decrease in only one of the conditions (Pearson correlation coefficient $r \leq -0.5$) but not in other two ($r > -0.5$). After that, we compared the numbers of units, which satisfied the request, between the three conditions. The presence of the increase (or decrease) during the acoustical stimulation in the CS+ condition was accepted only if the units did not have the increase (or decrease) in the CS- and US conditions. For additional control, we calculated percentage of units that had the increase (or decrease) in CS- or US conditions only.

The analysis revealed very small amount of units, which would satisfy the requirements described above and showed slow increase in the CS+ but not in the CS- or US conditions (Table 1). The percentage of units, in which the slow increase was observed during the noise and tone stimulation in the CS+ condition, did not exceed the chance level of 5%. More units than the chance level were detected having the slow decrease during the noise and tone stimulations. However, the percentages were not significantly greater than found to have decrease in the CS- only and US only condition ($\chi^2 = 3.56$, df = 1, $p > 0.05$ for the decrease during the noise stimulation in the CS+ only and CS- only condition; $\chi^2 = 0.12$, df = 1, $p > 0.05$ for the decrease during noise stimulation in the CS+ only and US only condition). The percentages of units with decrease during the tone stimulation found only in CS+ and CS- conditions did not differ significantly ($\chi^2 = 0.08$, df = 1, $p > 0.05$). The percentages where the decrease was found during CS+ only and during US only differed significantly ($\chi^2=5.93$, df = 1, $p < 0.05$). However, based on similar percentages of units with the decrease in the CS+ only and CS- only conditions, we assumed that the decrease was related with the tone stimulation and not with expectation. From the result, we concluded that there were no slow changes in neuronal activities related to expectations in the CS+ conditions before the water delivery in the groups of the low-trained monkeys.

Change during		Noise stimulation			Tone stimulation		
Only in condition		CS+	CS-	US	CS+	CS-	US
Type of change	Slow increase	0%	2.67%	1.33%	0%	1.33%	1.33%
	Slow decrease	6.67%	1.33%	8.0%	12.0%	10.67%	2.67%

Table 1. Percentage of the units with slow increase/decrease during acoustical stimulation in only one of the three conditions.

3.1.4. Presence of unconditioned stimuli shifted the level of the baseline upward

In this subsection, we analyzed the differences between the baselines in the CS+ and CS- conditions. We supposed two reasons that could lead to this shift. (1) The acoustical stimulation, particularly the tone stimulation, had no meanings for the monkeys in the conditions CS-. Therefore, the shift might be related to the meaning of the acoustical stimulation in the CS+ condition. (2) Simultaneously, the difference between the conditions CS+ and CS- was the presence of the unconditioned stimuli, i.e., the water delivery. In order to control whether the sound meaning or the presence of the unconditioned stimuli was the main reason for the shift in the baseline, the neuronal activity in the CS+ condition was compared with the condition where the water was delivered regularly without being predicted by the acoustical stimuli (US, Figure 1, Figure 2).

The average neuronal activities of the CS+ and US conditions are presented in Figure 10A. Unlike the differences between the baselines between the CS+ and CS- conditions (Figure 5), the baseline levels in the CS+ and US conditions were similar. To control whether the baseline levels indeed did not differ significantly, we chose a 1000 ms time window before the noise onset in the CS+ condition and the corresponding time window in the US condition, i.e. from 4300 until 3300 ms before the water delivery in both conditions. The positioning of the time window ensured that the differences, if detected, did not originate from the responses to the acoustical stimulation in the CS+. Also that is the latest time window after the last reward delivery that can be measured and compared between conditions. The analysis revealed that the average activities did not differ significantly between the two conditions (Wilcoxon signed rank test, p = 0.09, Supplementary table 2F).

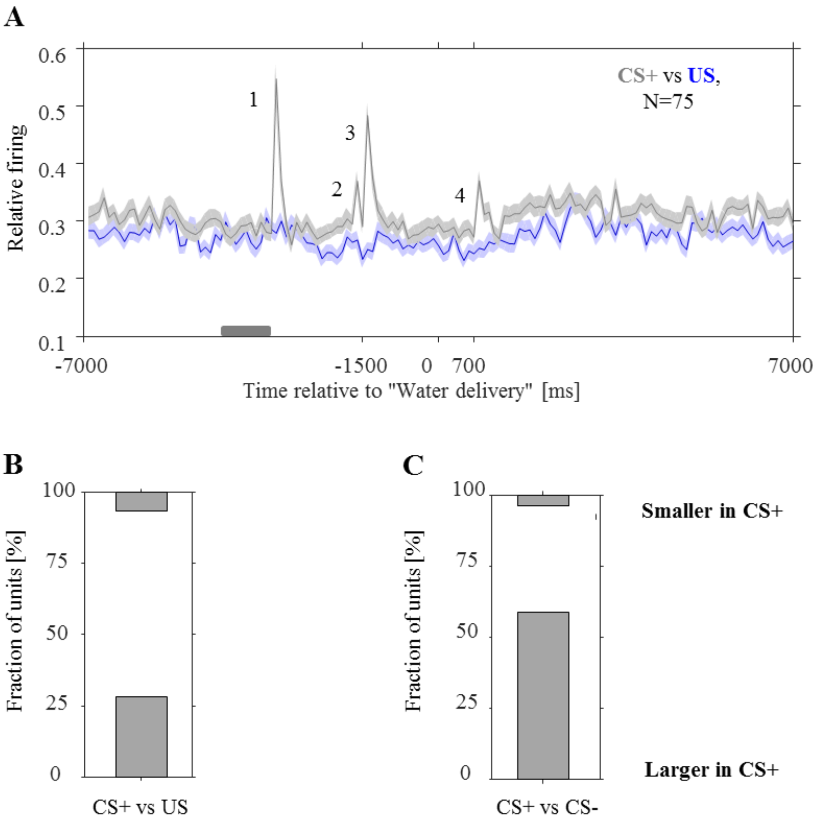

Figure 10. Neuronal population activities in the CS+ and US conditions (A). Comparison of the baselines between the CS+ and US (B) and between the CS+ and CS- conditions (C).

The subplot (A) shows average PSTHs (bin = 100 ms) with SEM (shadow) of the population, 75 units, during presentation of the CS+ (gray color) and US (blue color) condition. Numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. Subplot (B) shows results of the baseline comparisons (see the horizontal gray bar in the subplot (A)) in individual units between the conditions CS+ and US. Subplot (C) shows the result of the baseline comparisons between the CS+ and CS- conditions in the corresponding time window. (B and C) The bottom bars indicate fraction of units in which the baselines were significantly higher in the CS+ condition ($p < 0.05$, Wilcoxon signed rank test). The upper bars indicate fraction of units in which the baselines were significantly lower in the CS+ condition. The space between the bars indicate fraction of units with insignificant differences between the baselines.

Further, we compared the neuronal activities in the same time window between the CS+ and the US conditions for each of the 75 units (Figure 10B, Supplementary table 3F). More than half of the population (~65%) had similar levels of the baselines in both conditions. Around

28% of the units had higher levels in the CS+ condition. Only ~7% of the units had higher levels of the baselines in the US condition. For comparison, we also analyzed the same time window (during 1000 ms before the noise onset) between the conditions CS+ and CS-. The average population activity was significantly higher in the CS+ condition (Wilcoxon signed rank test, $p = 1*10^{-8}$, Supplementary table 2E). Large part of the population, ~59%, had higher activity in the CS+ condition (Figure 10C, Supplementary table 3E). Similar size of the group, ~37%, had no significant differences in activities between the conditions. Only ~4% of the units had higher activity in this time window in the CS- condition.

Based on the similarity of the baselines in the CS+ and US conditions and on the differences between the CS+ and CS- conditions, we concluded that the presence of the unconditioned stimuli led to the upward shift of the baselines.

3.1.5. Baseline levels in the frequency of the mouth movements

In the previous subsection (3.1.4), we reported that the baseline shift in the neuronal activity was related to the presence of the unconditioned stimuli, i.e. the baseline level in the CS+ condition was higher compared with the CS- condition. We supposed three reasons why the presence of the water delivery could lead to the shifts in the baseline. (1) The first reason is the hearing of the sounds produced during the collection of the liquid drops. The electrodes were located in the auditory cortex; therefore, the neuronal activity may respond when additional sounds are present. However, it should be noted that the self-initiated sounds led to the inhibition but not to excitation in the neuronal activity in the auditory cortex (Carcea et al. 2017, Schneider and Mooney, 2015, Buran et al. 2014, see also further in the section where the influence of sense of agency is controlled). (2) The process of water collection requires movements of lips, tongue, jaw etc. Early research of our colleagues showed already that hand movements evoked responses in the neurons in the auditory cortex (Brosch et al. 2005) but were not related to any artifacts. Thus, the second reason of the baseline shift and the higher firing may be due to the regular mouth movements, rather than hand movements, of the monkeys, (3) Most probably, the process of water collection was pleasant for the deprived monkeys and activated some areas in the middle brain such as the ventral tegmental area or substantia nigra (Ljunberg et al. 1992, Schultz 1998). Some studies showed that these structures were morphologically connected with the auditory cortex (Budinger et al. 2008); that their activation also changed neuronal activity in the auditory cortex immediately (Huang et al. 2016b, Mylius et al. 2014). Thus, the third reason of the baseline shift might be due to the signal transmitted from the non-cortical areas of the brain.

We cannot control the individual effects separately because they are all interconnected. In the current situation, the pleasure from the desired liquid could not occur without the mouth movements during licking and swallowing, which in turn are always accompanied with some sounds. Thus, we only can control the dependence of the neuronal activity on the complex of the three reasons (the triple complex) that we called the frequency of the mouth movements.

Figure 11 and Figure 12 present the relative frequency of the mouth movements during the three conditions (A) and the neuronal activities recorded during this time (B). Each figure represents the average frequency of the mouth movements during one session, which was the second session in a row for each of the monkeys. In the first example (Figure 11), the frequency of the mouth movements was on average 0.51 and 0.39 during 1000 ms before the noise onset and the relative time window in the conditions with water delivery, CS+ and US. The frequency of the mouth movements was only 0.10 during the same time window in the condition CS-. The average frequency of the mouth movements differed significantly between the CS+ and CS- conditions (Wilcoxon signed rank test, $p = 10*4^{-8}$) when was insignificantly different between the conditions CS+ and US (Wilcoxon sing rank test, $p = 0.05$). Similar changes were observed in the neuronal activities in the three conditions. The average firings during the 1000 ms before the noise onset were 0.47, 0.5 and 0.19 in conditions CS+, US and CS- respectively. The average firing during the time window was significantly higher in the CS+ condition than in the CS- (Wilcoxon signed rank test, $p = 2*10^{-5}$). Average firings of the two conditions with water delivery did not differ significantly (Wilcoxon signed rank test, $p = 0.77$). Thus, the average frequency of the mouth movements and the level of the baseline in the neuronal activity increased from the conditions without water delivery, CS-, to the conditions with water delivery, CS+ and US.

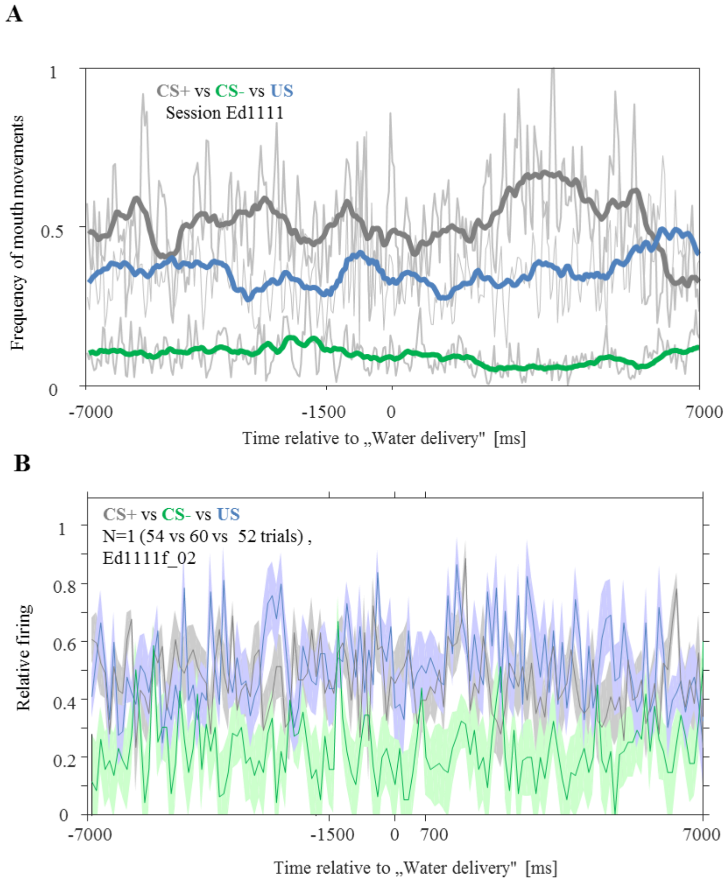

Figure 11. Frequency of the mouth movements during one session (A) and neuronal activity of one unit recorded during the same session (B).

(A) The subplot shows filtered frequency of the mouth movements (rational transfer function, the window size of 20 points and the coefficient equal to 1) in the conditions CS+, CS- and US (gray, green and blue curves). The original curves of the frequencies during the correspondent conditions are plotted with the grey color. (B) The plot shows average PSTHs (bin =

100 ms) with SEM (shadow) of one unit during presentation of CS+ (gray color), CS- (green color) and US (blue color) condition. Water was delivered only in the CS+ and US condition; for the CS+ and CS- condition zero point is 1500 ms after the tone onset.

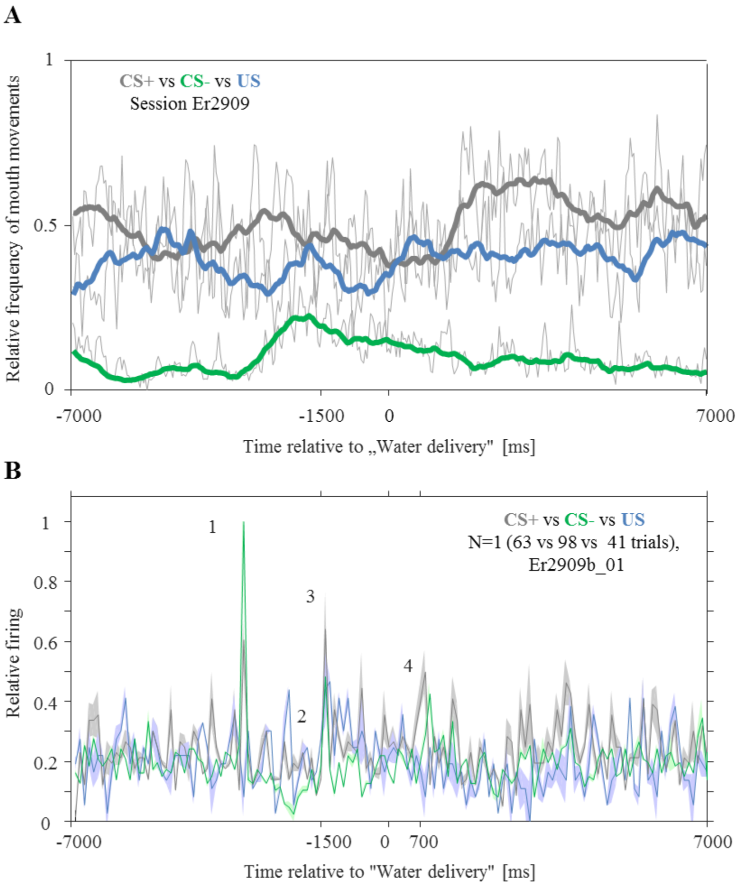

Figure 12. Frequency of the mouth movements during one session (A) and neuronal activity of one unit recorded during the same session (B).

(A) The subplot shows filtered frequency of the mouth movements (rational transfer function, the window size of 20 points and the coefficient equal to 1) in the conditions CS+, CS- and US (gray, green and blue curves). The original curves of the frequencies during the correspondent conditions are plotted with grey color. (B) The plot shows average PSTHs (bin=100 ms) with SEM (shadow) of one unit during presentation of CS+ (gray color), CS- (green color) and US (blue color) condition. Water was delivered only in the CS+ and US condition; for the CS+ and CS- condition zero point is 1500 ms after the tone onset. Numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset.

In another example (Figure 12), the frequency of the mouth movements before the noise onset and in the corresponding time window was also higher in the conditions with water delivery, CS+ and US, compared with the CS- condition. The frequencies were 0.43, 0.30 and 0.16 in the conditions CS+, US and CS- respectively. Analysis revealed that the frequency in the CS+ condition was significantly higher than in the CS- condition (Wilcoxon signed rank test, $p = 4*10^{-7}$). The frequency was significantly higher in the CS+ compare with the US condition (Wilcoxon signed rank test, $p = 0.023$). In the neuronal activity, the levels of the baselines of the same time window were also slightly higher in the conditions with water delivery, CS+ and US, than in the CS- condition and were 0.23, 0.22 and 0.19 respectively. However, the differences of the baseline levels between the CS+ and CS- conditions were insignificant (Wilcoxon signed rank test, $p = 0.79$) the same as the differences between the CS+ and US condition (Wilcoxon signed rank test, $p = 0.18$). Thus, the ratio of baseline levels in the neuronal activities was different compared with the ratio in the frequencies of the mouth movements between the three conditions in this unit. Based on the observed results we concluded that the neuronal activity of some, but not all, units correlated with the frequency of the mouth movements.

We had opportunity to analyze fourteen sessions in which the videos of the low-trained monkeys were recorded (9 for Ed and 5 for Er). We wanted to control how much the frequency of the mouth movements was higher in the conditions with water delivery. In order to test this, we calculated the average with SEM of the frequencies of the mouth movements during 1000 ms before the noise onset or in the corresponding time window in the CS+, CS- and US conditions and obtained the values 0.42 ± 0.05, 0.13 ± 0.02 and 0.33 ± 0.06 respectively. The average frequencies of the mouth movements were higher in the conditions with water delivery (Wilcoxon signed rank test, $p = 6*10^{-4}$ and $p = 0.01$ for the comparisons of the CS+ vs CS- and US vs CS- respectively). The frequencies of the mouth movements did not differ between the CS+ and US conditions ($p = 0.27$). Thus, we observed some correlations between the frequencies of the mouth movements and the baseline levels of the neuronal activities (see section 3.1.4) in the population. We supposed that the increase in the mouth movements during some sessions may influence the level of the baseline of some units.

3.1.6. Poor evidence of association of the acoustical stimuli and water delivery in the CS+ condition

We expected to find an increase in the frequency of the mouth movements immediately after the water delivery or even a little earlier. However, the analysis showed a significant increase in the frequency of the mouth movements in the CS+ condition during 2500 ms after the water delivery in the first example unit (Spearman’s rank correlation coefficient, $r=0.66$ and t-test, $p = 2.13*10^{-15}$, Figure 11) and in the second unit (Spearman’s rank correlation coefficient, $r = 0.75$ and t-test, $p=8.30*10^{-7}$, Figure 12). Unlike to the CS+ condition, no increase was observed in the frequency of the mouth movements in the US condition in the first unit (Spearman’s rank correlation coefficient, $r = 0.15$ and t-test, $p = 0.38$, Figure 11) and in the second unit (Spearman’s rank correlation coefficient, $r = -0.04$ and t-test, $p = 0.83$, Figure 12). Thus, the monkeys, probably, reacted on the tone offset and expected the water delivery after that moment; but could not predict the water delivery in the US condition.

We calculated the frequency of the mouth movements after the water delivery or during the corresponding time window in the three conditions for each of the 14 sessions. We expected that the frequency will increase after the water delivery in the conditions CS+ and US; and that the increase will be larger in the CS+ condition due to a prediction by the tone stimulation (Fanselow and Wassum 2016). Also, we expected that the frequency of the mouth movements in the relative time window will not increase in the CS- condition due to the absence of the water delivery. We supposed that these changes will be clearer from the first to the last recording session. We found that the coefficients of the changes varied largely from the first to the last session for each of the three conditions. Between all recorded sessions, the average and SEM of the coefficients for the condition CS+ was 0.37±0.09. For the US condition, where a less steep increase was expected due to the absence of predicting acoustical stimuli, the parameters were 0.28±0.09. In the CS- condition, where no increase was expected, the parameters were -0.24±0.11. In result, the average coefficients were significantly higher in the two conditions with water delivery, CS+ and US, compare with the condition CS- (Wilcoxon signed rank test, $p = 6*10^{-4}$ and $p = 1*10^{-4}$ for the comparisons between the CS+ and CS- and between the US and CS- respectively). The coefficients of the two conditions with water delivery did not differ significantly ($p = 0.5$, between the CS+ and US). Thus, we concluded that the monkeys did not associate the acoustical stimuli with water delivery in the CS+ condition. The results were probably related to the latent inhibition (Lubow and Gewirtz 1995, Lubow and Moore 1959).

We also controlled whether the increase became stronger from the first to the last session in the CS+ condition and did not reveal a positive slope. The coefficients of the US and CS- conditions also did not build any positive slopes. Based on these results, we decided that the association was not strengthened with time.

3.1.7. A subgroup of units with slow increase in the neuronal activity after water delivery

Between the 75 units recorded in the low-trained monkeys, we observed some units that, besides from the phasic responses with fast changes in the neuronal activity (less than 100 ms), had slow changes in their activity (more than 100 ms). One of the examples of such units is presented in the Figure 13. In the CS+ and US conditions, the neuronal activity increased after the water delivery and reached a maximum in ~2500 ms. After the peak, the activity decreased slowly and reached minimum at the time of the tone onset. The slow changes were observed only in the condition CS+ and US but not in the CS-, where the water was not delivered. Additionally to the neuronal activities in the three conditions, we added the spike forms of this unit recorded at the time after the water delivery (and at the relative time window in the CS- condition) that proved that the increase was not due to mechanical artifacts.

We intended to select all units in the recorded population that had the slow changes in their activities. For this, we used a common feature that was specific for the most units with the changes- the increase of the neuronal activity during 2500 ms after the water delivery. Thus, we selected all units with the positive significant coefficients of the increase after the water delivery (Spearman’s rank correlation coefficient, $r > 0$ and t-test, $p < 0.05$). The result showed that 24% and ~23% of the units had the slow increase after the water delivery in the CS+ and US conditions respectively (Figure 14A). Only ~7% of units had the increase in the CS- condition. The numbers of the units with the slow increase in activity were significantly different between the three conditions ($\chi 2 = 16.07$, df = 2, $p < 0.05$). The increase in activity was detected more often in the conditions with water delivery, CS+ and US, than in the CS- condition ($\chi 2 = 9.80$, df = 1, $p < 0.05$ for comparison between the CS+ and CS- conditions, and $\chi 2 = 8.73$, df = 1, $p < 0.05$ for comparison between the US and CS- conditions). However, the number of units with the increase did not differ significantly between the two conditions with water delivery, CS+ and US (χ2 = 0.04, df = 1, p > 0.05). Interestingly we found similar results in the frequency of the mouth movements, where the increase was larger in the conditions with water delivery and steeper in the CS- condition, where the increases were similar between the CS+ and US conditions. Probably, the frequency of the mouth movements is related to the neuronal activities of some units.

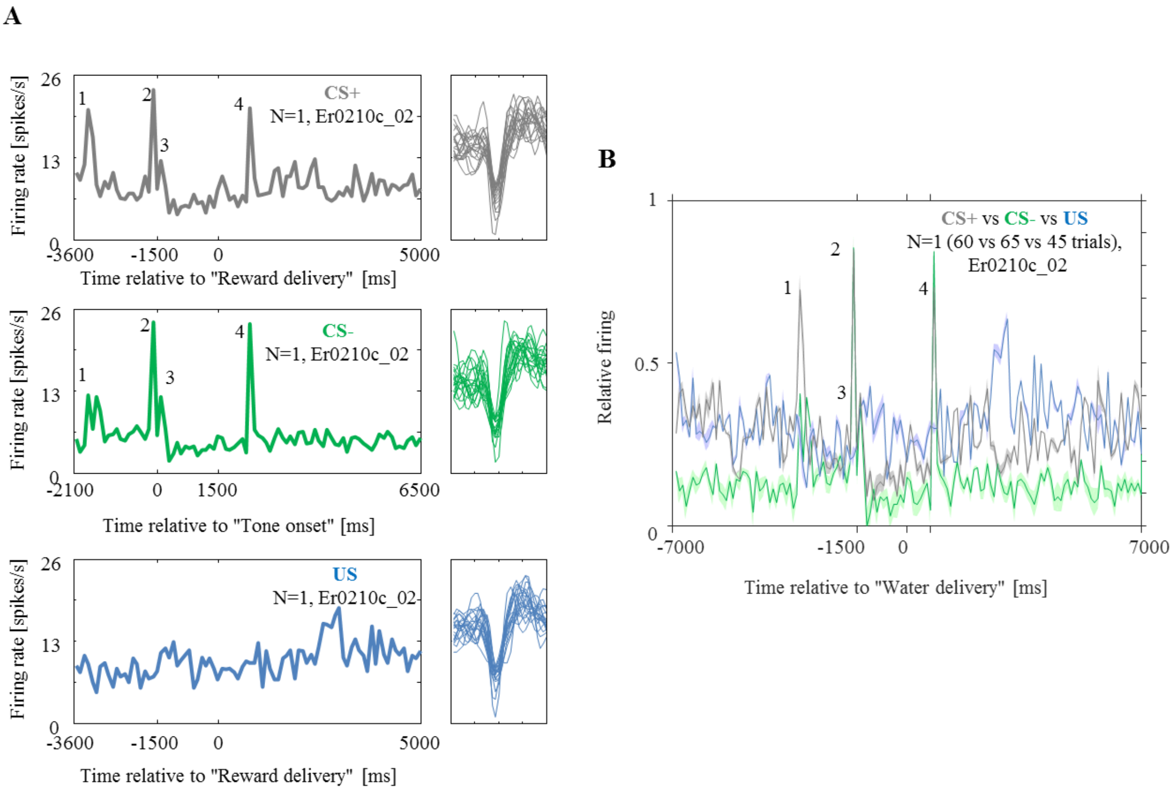

Figure 13. Example of one unit with slow increase in activity after water delivery and spike forms of this unit.
Subplots (A, left) and (B) show PSTHs (with bin size 100 ms) of the CS+ (gray color), CS- (green color) and US (blue color) conditions. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. In subplot (B), shadows indicate SEMs of the PSTHs. Subplot (A, right) also presents spike forms of the units in the three conditions.

In order to trace properties of the units with the slow increase in activity after the water delivery, we presented the averaged activity of the subgroups revealed by the tests. A subgroup of units and their activities are presented in Figure 14B. The plot shows averaged neuronal activity of 17 units in which an increase was detected in the US condition. It is worth mentioning that the average activity of these units also had increases at the same time window in the CS+ condition. Thus, we supposed that the increase after the water delivery in the US condition related to the increase in the CS+ condition. We also controlled whether the activities were different at the time of maximal activities (~2500 ms after the water delivery) between the CS+ and US conditions in the selected subgroup of the units. For this, we compared a time window of 1000 ms around the maximum value, i.e., 2000-3000 ms after the water delivery. The analysis revealed that the difference between the average activities were insignificant (Wilcoxon signed rank test, p = 0.36). Based on the results, we concluded that the units had the similar level of increase after the water delivery in the US condition and in the CS+ condition.

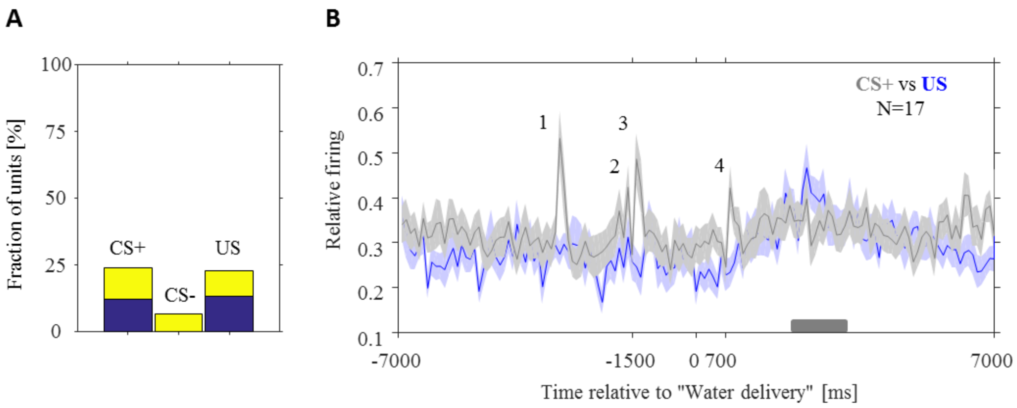

Figure 14. (A) Fraction of units that had significant increase in the activity in the conditions CS+, CS- and US. (B) PSTHs of the subgroup of units with significant increase in activity after the water delivery detected in the US condition.

(A) Blue and yellow parts of each bar in each histogram indicate the fractions of units of the monkeys Ed and Er from the population of 75 units respectively (ratio of units between the monkeys was ~1:1). (B) The subplot shows average PSTHs (bin = 100 ms) with SEM (shadow) of group of units during presentation of CS+ (gray color) and US (blue color) conditions. The subgroup showed significant increase in the neuronal activity in the US condition; the PSTH of the relative units is presented for the CS+ condition. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset.

The neuronal activities in the two conditions decreased slowly after the maximal value and reached minimum at the time of the tone stimulation before the new water delivery. We suppose that the upward shift of the baseline that was observed between the conditions with and without water delivery (section 3.1.5) was formed due to these slow changes in the neuronal activities.

Analysis of the tuning curves of the units with the slow increase in activity after the water delivery did not reveal any special properties. As was already mentioned, all of the 75 units had a significant response to at least one of the acoustical events. Thus, for instance, each of the four acoustical events led to the neuronal response in the CS+ conditions (see Figure 14B).

***

Summarizing this subsection, we revealed that the presence of the unconditioned stimuli led to the upward shift of the baseline level. The presence of unconditioned stimuli led to the slow increase in the neuronal activity in almost a third part of the population. The frequency of the mouth movements correlated with the neuronal activities of some units. All over, we concluded that the presence of the unconditioned stimuli led to the change in the neuronal activity of the monkeys with few experience.

During acoustical stimulations, the population neuronal activity adjusted to the baseline level was higher when the tone stimulation had no meaning. However, we cannot judge whether the sound meaning was the reason of the differences because we found too few evidence of the monkeys' association between the acoustical stimuli and water delivery.

3.2. Sound meaning and presence of unconditioned stimuli changed neuronal activity in the well-trained monkeys

The previous section (3.1) described changes in the neuronal activity in the auditory cortex of the low-trained monkeys. We tracked changes induced by the sound meaning and by the presence of unconditioned stimuli. It was surprisingly for us to find so many changes in the neuronal activity of monkey that were barely trained and did not associate the acoustical stimuli with water delivery. We were intrigued whether the well-trained monkeys with high experience in the performance of instrumental conditions will have similar changes in their neuronal activities. This section reports changes that were observed in the neuronal and behavioral activities of the well-trained monkeys during presentation of the three passive conditions.

For the experiment, two monkeys (other than in the low-trained group) were trained to perform three instrumental conditions (see section 2.2). The passive conditions were introduced to the monkeys for the first time only during the recording sessions, i.e. the passive conditions were new for the well-trained monkeys the same as they were new for the low-trained monkeys. The analysis of the neuronal activity was performed on the 65 units of one monkey (seventeen sessions) and 32 units of the second monkey (eight sessions). The neuronal activities were similar between the two monkeys and were combined in one sample (Supplementary table 1). We identified that the first spike latencies of these units were 18.0± 13.9 ms after the onsets of the pure tones. All (97) units of the sample responded to at least one of the four acoustical events, i.e., to the onsets and offsets of the pure tone and/or noise.

3.2.1. Well-trained monkeys actively suppressed behavioral actions in the passive conditions

Due to the long-time experience of the well-trained monkeys to interact with the metal bar in order to get the water in the instrumental conditions, we expected that the monkeys will also try to interact with the bar in the three passive conditions. Indeed, we found that the animals made unreasonable bar grasps and releases during the passive conditions (Figure 15). However, the quantity of the errors was much higher than we expected.

The first type of the error was the false alarm, i.e., the reaction to the noise or tone onset with the grasp of the bar. Due to the presence of the acoustical stimulations in only two of the three conditions, this type of the error could be observed only in the CS- and CS+ conditions (Figure 15A). The number of the bar grasps in the CS+ condition was higher than in the CS- conditions. We supposed that the monkeys tried to perform the Ext-HighE condition when they grasped the bar during the acoustical stimulations.

The second type of the errors was the grasp of the bar during the silence between the acoustical stimuli in the conditions CS+ and CS-, and during the whole time of the US condition (Figure 15B). Most probably, monkeys tried to perform one of the self-initiated conditions; therefore, we separated the error as a second type. The number of the bar grasps was lower in the US condition compared with the CS+. Also, the number of the bar grasps in the conditions with the water delivery, CS+ and US, was much higher than in the condition CS-.

The number of the bar grasps in the CS- conditions were much smaller than in the CS+ and the US conditions. We supposed that the reason of this is the presence of the unconditioned stimuli in the CS+ and the US conditions. When the monkeys made an error, a drop of water was still delivered in some time after the behavioral act. The monkeys associated the incorrect movement with the water delivery and were motivated to perform the same movement further in order to get the next drop of the water. In the CS- condition, the water was not delivered for much longer period of time that means that all additional movements of the monkeys were not rewarded. Therefore, the monkeys were not motivated to perform any new movements during that condition of the session.

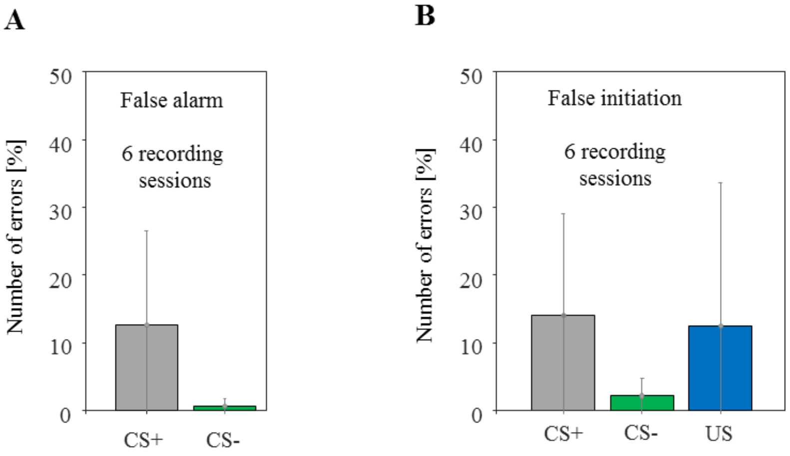

Figure 15. Frequency of errors during presentation of the three passive conditions for the well-trained monkeys. Subplot (A) shows the relative frequency of errors that were calculated as the ratio of the number of the false alarms to the number of the non-interrupted trials in the CS- or CS+. Subplot (B) shows the relative frequency of errors that were calculated as the ratio of the number of bar grasping (false initiation of the trials) to the number of the correctly performed trials in the conditions CS+, CS- or US.

In order to control the behavior of the monkeys, the metal bar was removed from the panel during the presentation of the passive conditions in some sessions (the neuronal activities were not recorded and, therefore, were not included in the sample). Even with the removed bar, the monkeys tried to perform the instrumental condition with moving their hand forth and back to the place where the bar was usually fixed. Thus, the removal of the bar would not change the behavior and the neuronal activity of the monkeys.

3.2.2. Responses evoked by the acoustical events were lower in the CS+ condition

The absolute responses to the four acoustical events in the low-trained monkeys were higher in the conditions with water delivery. However, the differences were the result of the upward shift of the baseline in the conditions with water delivery. Considering the baseline, we calculated the changes of the firing evoked by the acoustical events and found absence of the differences between the conditions in the majority of the units. We assumed that the changes in the neuronal activity of the well-trained monkeys will be larger because they associated the acoustical stimuli with water delivery.

In order to control the assumption, we first analysed the average activities of the 97 units that were recorded during the presentation of the CS- and CS+ conditions (Figure 16, Supplementary table 2A). In both conditions, the firing rate increased sharply after the onsets and offsets of the noise and pure tone. Comparison of the average neuronal responses to the four acoustical events (100 ms after the events) revealed that the responses to the three events, the noise onset, tone onset and tone offset, were significantly higher in the CS- condition (Wilcoxon signed rank test, $p = 3*10^{-8}$, $p = 6*10^{-3}$ and $p = 8*10^{-12}$) when the responses to the noise offsets did not differ significantly ($p = 0.19$). Interestingly, the baseline levels (during 1000 ms before the noise onset) differed significantly between the two conditions (Wilcoxon signed rank test, $p = 1*10^{-3}$, Supplementary table 2E). The baseline was shifted upward in the condition with water delivery. Thus, the response to the noise onset was so high in the CS- condition that even exceeded the difference of the baselines. The difference between the responses to the noise offset and tone onset, however, were not as strong as to the noise onset.

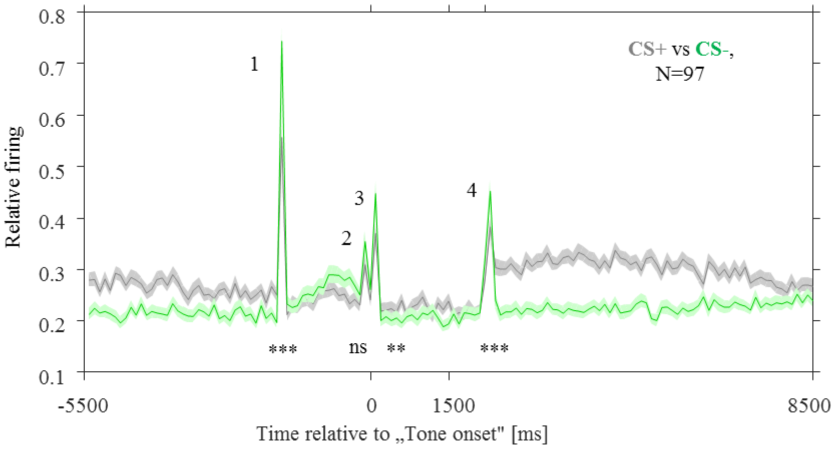

Figure 16. Neuronal activities during presentation of the CS+ and CS- conditions.

The plot shows average PSTHs (bin = 100 ms) with SEM of the population of 97 units during presentation of the CS+ (gray shadow) and CS- (green shadow) conditions. In the CS+ condition, water was delivered 1500 ms after the tone onset. The numbers indicate the acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. The stars indicate significant differences between the neuronal responses to the acoustical events between the two conditions (Wilcoxon signed rank test, ns - > 0.05, ** - $p < 0.01$, *** - $p < 0.001$).

Additionally to the analysis of the absolute responses in the population level, we analyzed the responses of each of the 97 individual units (Supplementary table 3A). We found a big fraction of units in the population, ~27%, that had significantly higher responses to the noise onsets in the CS- condition. Responses of the most units, ~69%, did not differ significantly between the conditions. The comparison revealed only ~4% of the units that responded higher to the noise onsets in the CS+ condition. The results of the comparisons of the noise offset and tone onset between the two conditions were very similar. The analysis revealed that a small group of ~16% had higher responses to the noise offset and the tone onset in the CS- condition, a group of ~7% had higher responses in the CS+ conditions. The most of the units 77% had no significant differences between the responses in the two conditions. For the tone offset, the analysis revealed a group of units, ~19%, that had higher responses in the CS- condition, a small fraction of ~5%, that had higher responses in the CS+ condition and larger group of ~76% that had no significant differences between the responses.

We also calculated the percentage of the baseline shifts between the conditions as we did for the low-trained monkey. The analysis revealed that the baseline level in the CS+ condition rose in average by 44% above the baseline level in the CS- condition in 65% of the units. The baseline of the rest of the 35% units decreased in average by 17%.

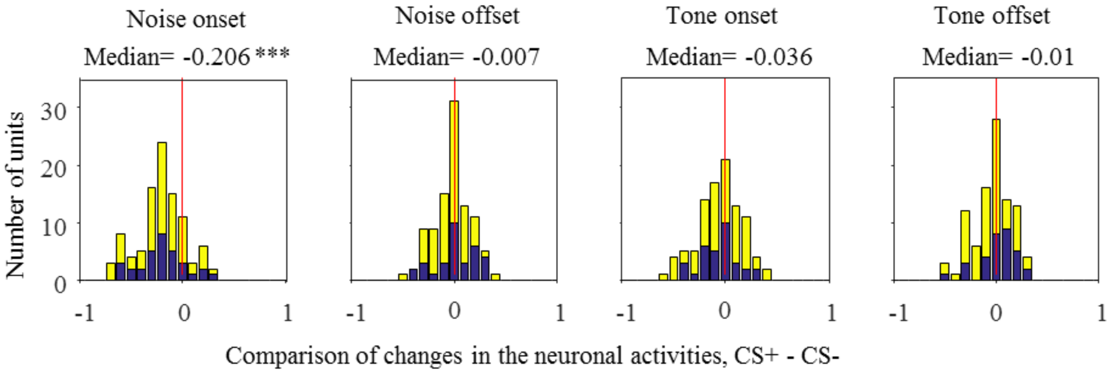

Figure 17. Comparisons of the neuronal activities evoked by the acoustical events and adjusted to the baseline levels in the CS+ and CS- condition.

Each subplot shows the number of units of the sample that had higher change of neuronal activity evoked by the acoustical events in the CS+ (comparison of changes > 0) or in the CS- condition (comparison of changes < 0). The values of the medians of each distribution are inserted into the boxes. The star near the median indicate its significant difference from zero (Wilcoxon signed rank test, *** - p < 0.001). The blue and yellow parts of each bar in each histogram indicate number of units recorded in Ba and We respectively (ratio of the units between the monkeys was ~1:2).

We controlled whether the changes in the neuronal activity evoked by the acoustical events were also larger in the CS- conditions. In order to calculate the changes in one condition, we found the ratios of the absolute value of the response to the acoustical events (100 ms after the events) and the neuronal activity before the acoustical events (during 100 ms). After that, we subtracted the two values and did so for each of the four events in each of the 97 units. The resulting distributions of the 97 values for the four acoustical events were compared with zero (Figure 17). The analyses revealed that the median value of the distribution build for the noise onset differed from zero significantly (Wilcoxon signed rank test, $p = 8*10^{-12}$, Supplementary table 2B). The median was also negative. That indicated that the change of the neuronal activity evoked by the noise onset in the CS- condition was larger than in the CS+ condition. The medians of the histograms for the noise offset, tone onset and tone offset did not differ from zero significantly (p = 0.33, p = 0.06 and p = 0.24), i.e., the changes of the neuronal activities were similar in the two conditions.

Between the 97 units, we found ~5% of the units in which the changes of the neuronal activity evoked by the noise onset in the CS+ condition were significantly larger (Supplementary table 3B). A greater amount of units, ~40%, showed that the changes were larger in the CS- condition. The percentages of the units in which the changes of the neuronal activity evoked by the noise offset were different, were ~19 and ~25%. Percentages for the tone onset and tone offset were ~22 and ~23%, and ~17 and ~24% respectively.

Baselines of some units in the population did not differ significantly between the CS+ and CS- conditions. Many of these units showed significantly higher responses to the acoustical events in the CS- condition. An example that supports our observation is presented in Figure 18. Baseline levels in the two conditions with and without water delivery were similar in the two conditions; therefore, we compared the absolute responses and accepted that the changes of the neuronal activity evoked by the events will have the same level of significance.

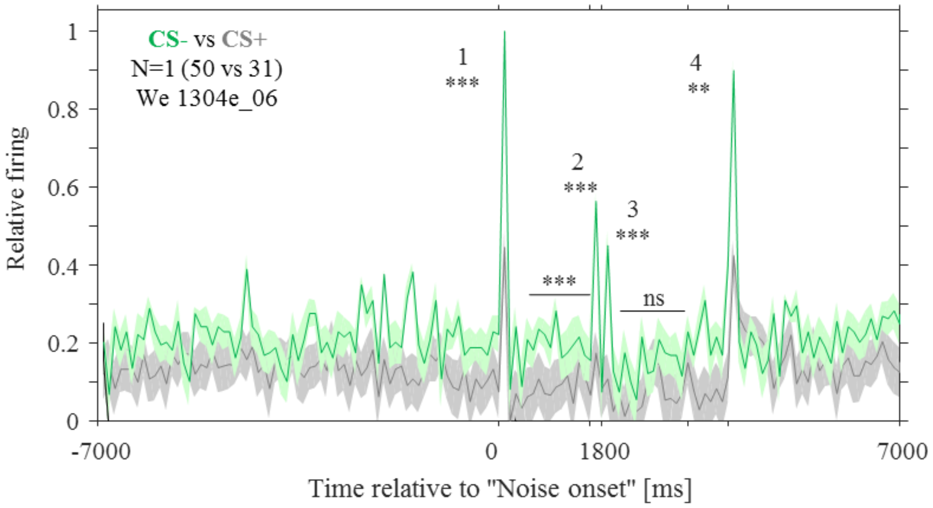

Figure 18. Example of one unit where the responses to the acoustical events were higher in the CS- condition compare with the CS+ condition.

The plot shows average PSTHs (bin = 100 ms) with SEM of one unit during presentation of the CS+ (gray shadow) and CS- (green shadow) conditions. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. The stars indicate significance of the differences between the responses evoked by the acoustical events and between the average activities during acoustical stimulations in the two conditions (Wilcoxon signed rank test, ns - $p > 0.05$, ** - $p < 0.01$, *** - $p < 0.001$).

Summarizing the results, we found that the neuronal activity of the well-trained monkeys had the upward shift of the baseline in the rewarded conditions CS+. Despite to the shift in the CS+ condition, the response to the noise onset was much higher in the CS- condition in the population level. That difference was not observed in the population of the low-trained monkeys (see also Supplementary table 2A-B). Probably, the lower response to the noise onset in the well-trained monkeys is related to the meaning of the acoustical event.

3.2.3. Neuronal activity during the noise and tone stimulations adjusted to the baseline was higher in the CS- condition

In the previous section (3.1.2), we reported that average neuronal activity recorded during the acoustical stimulations and adjusted to the baseline levels differed between the conditions with and without water delivery in the primary auditory cortex of the monkeys with few experience. We wondered whether similar changes will be observed in the neuronal activities of the well-trained monkeys.

First, we compared the average neuronal activities recorded during the acoustical stimulations (Supplementary table 2C). As it can be seen from the neuronal population activities (Figure 16), the average neuronal activities did not differ during the noise and tone stimulation. Indeed, the analysis showed that the differences were insignificant during the noise stimulation (200-1600 ms after the noise onset, Wilcoxon signed rank test, p = 0.14) and during the tone stimulation (200-1500 ms after onset, p = 0.09). However, the level of the baseline varied between the conditions and not only before the noise onset but throughout the entire trials.

The individual analysis of the 97 units revealed that ~25% of them had higher neuronal activity during the noise stimulation in the CS+ condition, and ~29% had higher neuronal activity in the CS- condition (Supplementary table 3C). For the tone stimulation, the neuronal activity was higher in ~32% of the units in the CS+ condition and was higher in ~23% in the CS- condition.

We also wanted to compare the average activities during the acoustical stimuli and consider the baseline shift between the conditions. In order to do this, we normalized the neuronal activity in each unit and each condition to the level of the baselines (*z*-scores) during 3000 ms before the noise onset. After that, we subtracted the average *z*-scored neuronal activity during the noise stimulation (200–1600 ms after the noise onset) between the two conditions in each unit. From the resulting values of each unit of the sample, we plotted a histogram and calculated the median (Figure 19). Further, we compared whether the median differed significantly from zero. The same procedure was made for the average *z*-scored neuronal activity during the tone stimulation (200-1500 ms after tone onset). The analysis revealed that both medians were shifted from zero significantly and were negative (Supplementary table 2D). The negative values mean that the changes of the neuronal activities were larger in the CS- condition compare with the CS+ conditions.

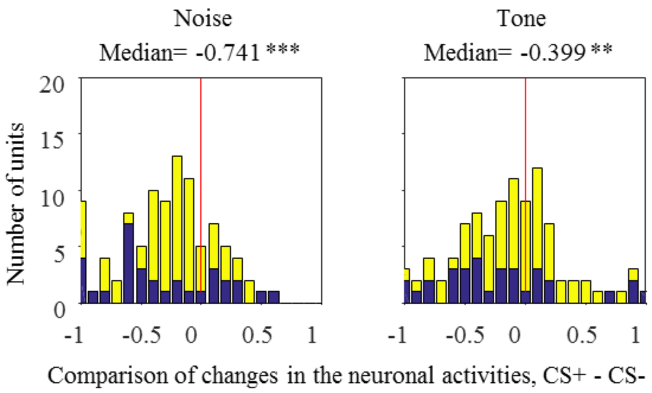

Figure 19. Comparisons of the neuronal activities evoked during the acoustical stimulations and adjusted to the baseline levels in the CS+ and CS- condition.

Each subplot shows the number of units of the sample that had larger change of activity during acoustical stimuli in the CS+ condition (comparison of changes > 0) or in the CS- condition (comparison of changes < 0). The medians of the two distributions are inserted into the boxes. The stars near the medians indicate the significant level of differences from zero (Wilcoxon signed rank test, *** - p < 0.001). The blue and yellow parts of each bar in each histogram indicate number of units recorded in monkey Ba and We respectively (ratio of units between the monkeys was ~1:2).

Between the 97 units, we found only ~7% in which the changes of the activities during the noise stimulation were significantly higher in the CS+ condition (Supplementary table 3D). Much bigger amount of units, ~38%, showed that the changes were larger in the CS- condition. The percentages of units, in which changes of activities evoked by the tone stimulation were larger in the CS+ and CS- conditions, were ~11 and 31%.

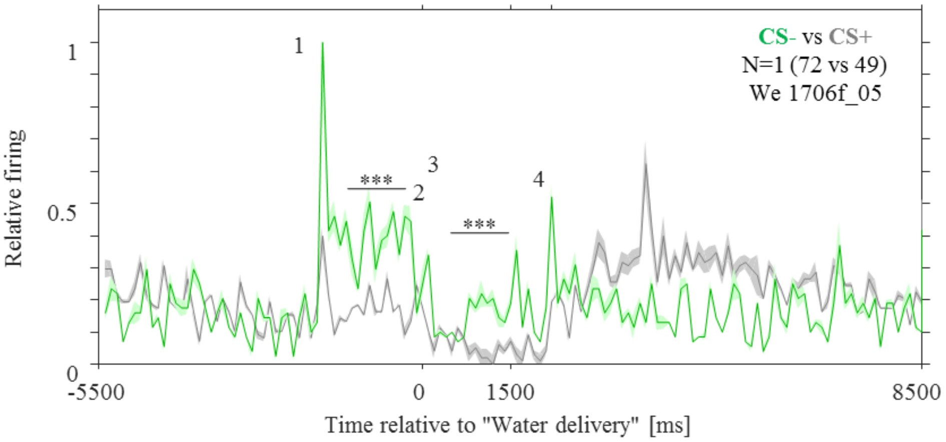

Figure 20. Example of one unit where the neuronal activity and their changes during the acoustical stimulations were higher in the CS- condition compare with the CS+ condition.

The plot shows average PSTHs (bin = 100 ms) with SEM of one unit during presentation of the CS+ (gray shadow) and CS- (green shadow) conditions. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. The stars indicate significance of the differences between the responses evoked by the acoustical events and between the average activities during acoustical stimulations in the two conditions (Wilcoxon signed rank test, *** - p < 0.001).

An example of the unit that had higher average activity during the tone stimulation is presented in the Figure 20. Due to the similarity of the baseline levels, we could compare the average activities during the noise and tone stimulation without subtraction of the baseline. The analysis revealed that, as expected, the average activity during the noise stimulation was significantly higher (Wilcoxon signed rank test, $p < 0.05$) in the CS- condition and did not differ between the conditions during the tone stimulation ($p > 0.05$).

Thus, the average neuronal activities recorded during the acoustical stimuli and adjusted to the level of baseline were higher in the condition without water delivery. Similar but less pronounced result was observed in the group of the low-trained monkeys (see also Supplementary table 2C-D). Probably, the experience led to higher differences between the neuronal activities during the acoustical stimulations, the noise, between the two conditions.

3.2.4. Absence of units with slow increase (or decrease) related to the expectation of the water delivery

Similarly as for the low-trained monkeys (section 3.1.3), we controlled the presence of the slow changes in the neuronal activity before the water delivery in the conditions CS+ in the group of the well-trained monkeys. We assumed that the experience of the animals may lead to the appearance of the slow changes in the neuronal activity.

In order to reveal the slow sustained increase/decrease during the noise stimulation, we used the time window 200-1600 ms after the noise onset. For the tone, we considered the time window 200-1400 ms after the tone onset. We controlled whether the activity increased during the time of the acoustical stimulations in only one of the conditions (Pearson correlation coefficient, $r \geq 0.5$) but did not increased in the other two ($r < 0.5$) conditions. The analysis was performed for each unit. Similar procedure was used for control the presence of the slow decrease in only one of the condition (Pearson correlation coefficient $r \leq -0.5$) but not in other two ($r > -0.5$). After that, we compared the numbers of units, which satisfied the request. The presence of the increase (or decrease) in the CS+ condition was accepted only if the units did not have the increase (or decrease) in the CS- and US conditions. For additional control, we calculated percentage of units that had the increase (or decrease) in CS- or US conditions only.

The analysis revealed that the percentages of units in which the slow decrease and increase were observed during the noise and tone stimulation respectively were lower than by chance. The percentage of units in which the slow increase during the noise stimulation was observed in the CS+ condition only, was slightly higher than in the CS- condition only, however the difference was insignificant ($\chi^2 = 0.28$, df = 1, $p > 0.05$, Table 2). The percentage of units, in which the increase during the noise stimulation was observed in the CS+ conditions, was significantly higher than in the US condition only ($\chi^2 = 8.78$, df = 1, $p < 0.01$). The result indicates that the increase during in the activity in the CS+ and CS- conditions was observed due to the noise stimulation but not due to the expectation of the water delivery. For the slow decrease during the noise stimulation in the CS+ condition only, the percentage was insignificantly different from the percentages detected for the CS- condition only ($\chi^2 = 0.69$, p

> 0.05 between the CS+ and CS-; $\chi^2 = 0.69$). The percentage was higher than in the US condition only (df = 11.65, $p < 0.05$ between the CS+ and US). We assumed that the detected increases were related to the noise stimulation rather than the expectation of the water delivery. The percentages of units in which the slow decrease of activity was observed during the tone stimulations in the condition CS+ only also did not differ significantly between the other percentages ($\chi^2 = 0.71$, df = 1, $p > 0.05$ between the CS+ and CS-; $\chi^2 = 2.32$, df = 1, $p > 0.05$ between the CS+ and US). From the result, we concluded that the slow changes in the neuronal activity in the CS+ conditions before the water delivery were not related to the expectations. Thus, experience did not lead to the increased number of units with the slow changes that were described in other studies and appeared during expectation of reinforcement.

Change during		Noise stimulation			Tone stimulation		
Only in condition		CS+	CS-	US	CS+	CS-	US
Type of change	Slow increase	14.43%	10.31%	1.03%	2.06%	2.06%	3.09%
	Slow decrease	3.09%	3.09%	3.09%	5.15%	8.25%	11.34%

Table 2. Percentage of the units with slow increase/decrease during acoustical stimulation in only one of the three conditions.

3.2.5. Presence of unconditioned stimuli shifted the level of the baselines upward

The acoustical stimulations in the conditions CS+ and CS- had different meanings for the well-trained monkeys because the first one led to the water delivery when another did not (see also section 3.2.7). We reported that the baseline levels differed between the conditions. In the present subsection, we control whether the shift in neuronal activity of the well-trained monkeys was related to the presence of unconditioned stimuli or to the meaning of the acoustical stimuli. In order to reveal the main reason, we compared the neuronal activities between the conditions where a drop of water was delivered regularly to the monkeys (US passive condition) with the CS+ condition.

The population average activities of the CS+ and US conditions are presented in Figure 21A. The baseline in the US condition had even higher level than in the CS+ condition. To control whether they differed significantly, we chose the time window during 1000 ms before the noise onset in the CS+ condition and the corresponding time window in the US condition, i.e., from 4300 to 3300 ms before the water delivery in both conditions. The positioning of the time window ensured that the differences, if detected, did not originate from the response to the acoustical stimulation in the CS+. The analysis revealed that the average activities did not differed significantly between the conditions (Wilcoxon signed rank test, $p = 0.26$, Supplementary table 2F).

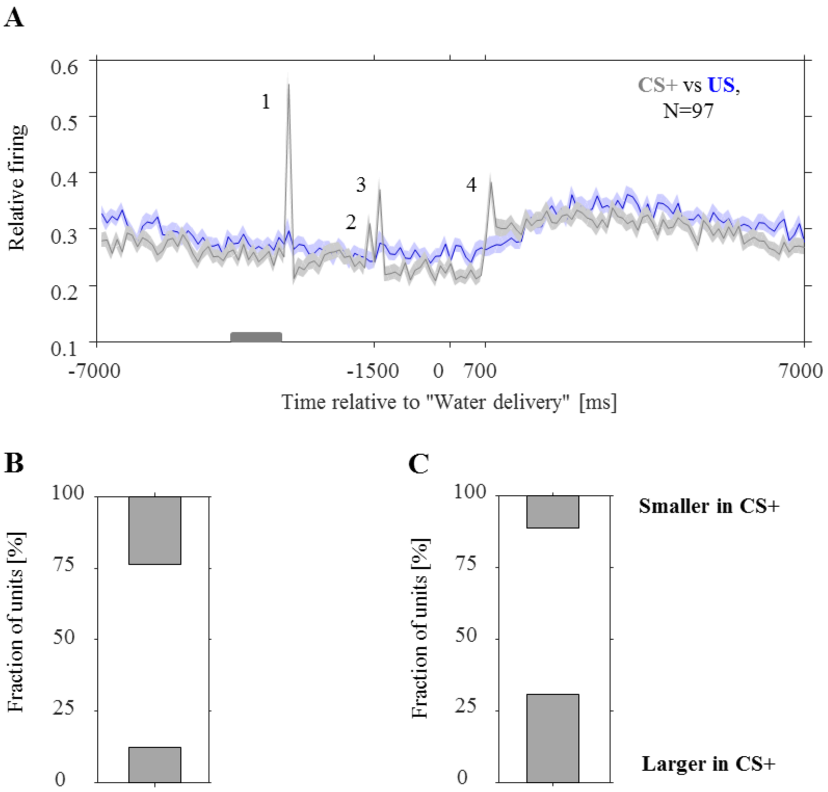

Figure 21. Neuronal population activities in the CS+ and US conditions (A). Comparison of the baselines between the CS+ and US condition (B) and between the CS+ and CS- condition (C).

The subplot (A) shows average PSTHs (bin = 100 ms) with SEM (shadow) of the population, 97 units, during presentation of CS+ (gray color) and US (blue color) condition. Numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. Subplot (B) shows results of the baseline comparisons (see the horizontal gray bar in the subplot (A)) in individual units between the conditions CS+ and US. Subplot (C) shows the result of the baseline comparisons between the CS+ and CS- conditions in the corresponding time window. (B and C) The bottom bars indicate fraction of units in which the baselines were significantly higher in the CS+ condition ($p < 0.05$, Wilcoxon signed rank test). The upper bars indicate fraction of units in which the baselines were significantly lower in the CS+ condition. The space between the bars indicate fraction of units with insignificant differences between the baselines.

Further, we compared the neuronal activities of the same time windows between the CS+ and the US conditions for each of the 97 units (Figure 21B, Supplementary table 3F). More than half of the population, ~59% of the units had similar level of the baselines between the conditions. Around 12% of the units had higher level of the baseline in the CS+ condition; and a larger group of ~24% had higher level of the baseline in the US condition. We also analysed the same time window (during 1000 ms before the noise onset) between the conditions CS+ and CS- (Figure 16). The levels of baseline was significantly higher in the CS+ condition compare with the CS- condition (Wilcoxon signed rank test, $p = 1*10^{-3}$, Supplementary table 2E). Large part of the population, ~30%, had higher activity in the CS+ condition (Figure 21C, Supplementary table 3E). A large group of ~59% of the units had no significant differences between the levels of the baselines. Only ~11% of the population had higher level of the baseline in the CS- condition. Therefore, the presence of unconditioned stimuli led to upward shift of the baseline.

We hypothesized that if the average level of the baselines were similar between the CS+ and US conditions then the reason of the baseline shift in the CS+ and US conditions consisted in the presence of unconditioned stimuli. If the average level of the baselines is the highest in the CS+ then the reason of the baseline shift is, probably, in the sound meaning. We found that the presence of unconditioned stimuli led to the upward shift of the baseline. Thus, in both groups of the monkeys, the low- and the well-trained, the effect of the presence of unconditioned stimuli was similar, i.e., the level of the baseline shifted upward (see also Supplementary table 2E-F and Supplementary table 3E-F).

3.2.6. Baseline levels in the frequency of the mouth movements

In the previous subsection, we reported that the shift of the baseline in the neuronal activities was related to the presence of the unconditioned stimuli (section 3.2.5). The same as for the low-trained monkeys, we could suppose three reasons of the shifts in the conditions with water delivery: (1) hearing of the self-produced sounds during the mouth movements, (2) mouth movements itself, and (3) pleasant component of the water collection. We cannot clearly divide one of the three possible reasons because they were all related to each other. However, we could control whether the neurons were sensitive to the frequency of the mouth movements, i.e., to the triple complex of the reasons.

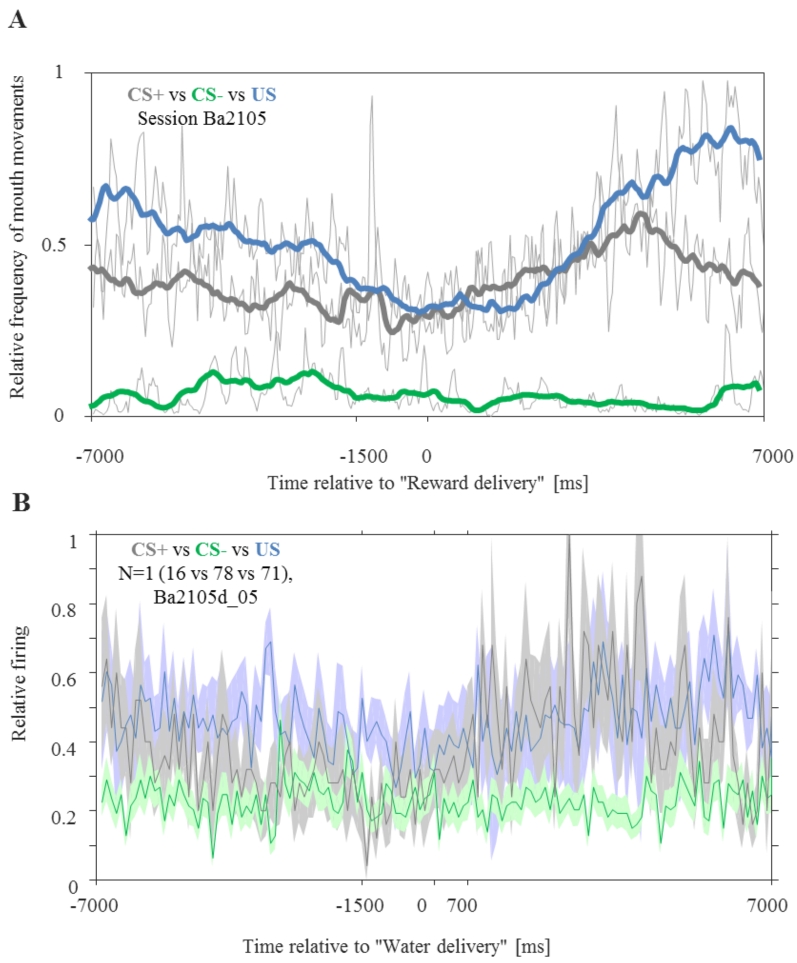

Figure 22. Example of frequency of the mouth movements during one session (A) and neuronal activity of one unit recorded during this session (B).

(A) The subplot shows filtrated frequency of the mouth movements (rational transfer function, the window size of 20 and the coefficient equal to 1) in the conditions CS+, CS- and US (gray, green and blue curves). The original curves of the frequencies in the correspondent conditions are plotted with grey color. (B) The subplot shows average PSTHs (bin = 100 ms) with SEM (shadow) of one unit during presentation of CS+ (gray color) and CS- (green color) and US (blue color) conditions. (A and B) The water was delivered only in the CS+ and US condition; the curve of the CS- condition is relative to the 1500 ms after the tone onset.

Figure 22A, B represents frequency of the mouth movements and neuronal activity in the three conditions recorded during the same session. As expected, the frequency of the mouth movements was on average higher in the conditions with water delivery, US and CS+, compare with the condition CS-. Moreover, the frequency was higher in the condition US than in the CS+. We calculated the average frequencies of the mouth movements during 1000 ms before the noise onset, or in the relative time window, in each condition and found that the frequencies in the conditions CS+ and US were 0.33 and 0.49 respectively. The values differed significantly (Wilcoxon signed rank test, $p = 6*10^{-3}$). The average frequency in the same time window in the CS- condition was 0.09 and was significantly different from the CS+ condition ($p = 4*10^{-8}$). We also controlled the levels of the baselines of the neuronal activities during the three conditions in one of the units recorded during this session (Figure 22B). The levels of the baselines seemed to be higher in the conditions CS+ and US. Indeed, the average levels of the baselines were 0.28 in the CS+, 0.49 in the US and 0.19 in the CS- condition. The difference was significant between the two conditions with water delivery (Wilcoxon signed rank, p = 0.009) and higher in the US condition. The differences were insignificant between the CS+ and CS- conditions (Wilcoxon signed rank test, p = 0.94). Thus, the baseline levels in the neuronal activity of the present unit correlated with the baseline levels of the frequency of the mouth movements only partially.

We wanted to control how much the frequency of the mouth movements was higher in the conditions with water delivery in the sessions. All together, we had seven sessions with the records of the mouth movements (all records with Ba). We compared the average frequencies of the mouth movements in the three conditions during 1000 ms before the noise onset, or in the relative time window, in all sessions. We found that the average frequency of the mouth movements and SEM were 0.42±0.03, 0.16±0.08 and 0.43±0.08 for the conditions CS+, CS- and US respectively. The average frequency of the mouth movements was significantly higher in the condition CS+ compare with the CS- (Wilcoxon signed rank test, p = 0.02). The frequency was higher in the condition US compare with the CS-, however the differences were insignificant due to the larger deviations between the sessions (p = 0.07). The frequencies in the conditions CS+ and US also did not differ significantly (p = 0.94). Thus, the frequency of the mouth movements correlated with the baseline levels in the neuronal activities only partly. Probably, the frequency of the mouth movements had some influences on the neuronal activities of some units.

3.2.7. Evidence of association of the acoustical stimuli and water delivery in the CS+ condition

We expected that the well-trained monkeys will have an increase in the frequency of the mouth movements after the water delivery. The example of the frequency of the mouth movements showed that the frequency increased after the water delivery in the two conditions, but it was much slower than we expected (Figure 22A). In the present example, the maximal value of the mouth movements was observed at the time ~4000 ms in the CS+ condition and only about ~6000 ms in the US conditions (Figure 22A). Coefficients of the slopes calculated during 2500 ms after the water delivery in the CS+ and US condition were positive and significant (Spearman's rank correlation coefficient, $r=0.59$, $p=1.8*10^{-4}$ and $r=0.66$, $p=2.05*10^{-5}$ respectively). In contrast, there was a much lower and insignificant increase in the corresponding time window in the CS- condition ($r=0.22$, $p=0.19$). From one side, it was clear that the maximal value was more distant from the moment of the water delivery in the US condition because it is less predictable for the animals due to the absence of any signal. However, we considered the time interval of 3500-6000 ms too long. The reason for this can be a large variation of the conditions that we presented to the animals. In some of them water delivery depended on the reaction time of the monkeys, in another on the time interval after the tone onset. Finally, in the US condition, they could control the time of the water delivery only by constantly looking at the water tube or estimate the time interval of 5000-11100 ms. Therefore, the longtime interval of the maximal value of the water delivery could be due to disorientation of the monkeys in the variation of conditions; or due to the generalization of behavior in order to optimize the allocation of attention and energy consumption (Sejnowski et al 2014). We also controlled the modulations of the neuronal activities recorded during this session. Similar to the frequency of the mouth movements, we found a very slow increase in the neuronal activity in the time after the water delivery in the conditions CS+ and US with a late peak at ~3500 ms (Figure 22B). The increase in the neuronal activity was significant in the CS+ condition (Spearman's rank correlation coefficient, $r = 0.40$, $p = 0.04$) but not in the US condition ($r = 0.37$, $p = 0.06$). In the CS- condition, the increase was also weak and insignificant ($r = 0.26$, $p = 0.20$). Thus, the increase after the water delivery was present in the courses of the mouth movements and of the neuronal activity in the conditions CS+ and US. However, the maximal values were different between the courses of the mouth movements and neuronal activities in the same conditions.

We calculated the frequency of the mouth movements after water delivery or during the corresponding time window in the three conditions in the seven sessions in which the video was recorded for the well-trained monkey (for monkey Ba only). We expected that the frequency of the mouth movements will increase after the water delivery in the conditions CS+ and US; and that the frequency will not change in the condition CS- due to the absence of the water delivery (Fanselow and Wassum 2016). We also assumed to find a stronger increase after the water delivery in the CS+ condition compared with the US because the water delivery of the first was predicted by the acoustical stimuli. We found that the coefficients of the increases varied largely between the sessions. Between all recorded sessions, the average and SEM of the coefficients for the condition CS+, US and CS- were 0.45±0.11, 0.06±0.19 and 0.05±0.14 respectively. The average coefficients calculated in the CS+ conditions were significantly higher compared with the CS- (Wilcoxon signed rank test,

$p = 0.04$) and compared with the US condition ($p = 0.01$). The parameters of the US and CS- conditions did not differ significantly ($p = 1$). Thus, the results confirmed our assumptions that the increase in the CS+ conditions was indeed the strongest. However, it was unexpected that the increase after the water delivery in the US condition was steep as in the CS- condition. That indicated that the monkeys collected the drops evenly during the trial time.

In addition to the slow increase in the frequency of the mouth movements in the CS+ and the US condition, we observed other interesting changes in the recorded session; an example of the changes can be seen in the Figure 22. Similar changes were observed in the neuronal activities of the monkeys (Figure 16, Figure 21A). In the CS+ condition, sometime after the increase after water delivery, the neuronal activity and the frequency of the mouth movements decreased at the moment of the tone stimulation and stayed minimal during this time. In the US condition, the neuronal activity and the frequency of the mouth movements increased after water delivery slower than in the CS+ condition. The values reached the minimal value only in the time close to the next water delivery. The frequency of the mouth movements and the neuronal activity in the condition US were rather stable during the time of the trials but were shifted upward. In the CS- condition, the firing and the frequency of the mouth movements were stable throughout the entire trial. Note that in the CS+ condition, in which the water delivery was predicted by the acoustical stimuli, neuronal activity decreased to the level of the CS- condition. No similar leveling was observed in the US condition. We supposed that the monkey suspended the tube licking in the CS+ conditions during the time of the tone stimulations because they learned the association between the tone presentation and water delivery.

3.2.8. Almost half of the units with a slow increase in activity after water delivery

Similarly to the low-trained monkeys, we observed some of the units in the sample of the well-trained monkeys that, besides from the phasic responses with fast changes of activity (less than 100 ms), had activity with slow modulations (more than 100 ms). The slow modulations in the activity were observed only in the conditions CS+ and US (Figure 23). In the CS- condition, the baseline was stable from the beginning to the end of the trial. In the presented example unit, the activity increased slowly and reached the maximum in ~3500 ms after the water delivery in the conditions CS+ and US. Additionally to the neuronal activity of the units, we added forms of the spikes recorded during the three conditions (Figure 23A).

Similarly to the approach used previously (section 3.1.7), we intended to select all units in the recorded population that had the slow increase in their activity after water delivery. For this, we used a common feature that we observed in a large part of the population, i.e. the slow increase in the activity during ~2500 ms after water delivery. We selected only units that had significant increase in activity during this or relative time window in the three conditions (Spearman's rank correlation coefficient, $r > 0$ and t-test, $p < 0.05$). The result showed that 46.4% of the units had a slow increase after the water delivery in the CS+ condition (Figure 24A). The analysis indicated a small group of 11.3% with the slow increase in the activity in the CS- condition. In the US condition, 40.2% of units were detected to have the increase. We compared the sizes of fractions where the slow increase was observed between the three conditions and found that they were significantly different ($\chi2 = 33.6$, df = 2, $p < 0.05$). The fraction of units, where the slow increase was observed in the CS- condition, was significantly lower compare with the CS+ and US conditions ($\chi2 = 21.28$, df = 1, $p < 0.05$ for the comparison between the CS+ and CS-, and $\chi2 = 16.16$, df = 1, $p < 0.05$ for the comparison between the US and CS-). The sizes of the fractions were insignificantly different between the conditions CS+ and US ($\chi2 = 0.44$, df = 1, $p > 0.05$).

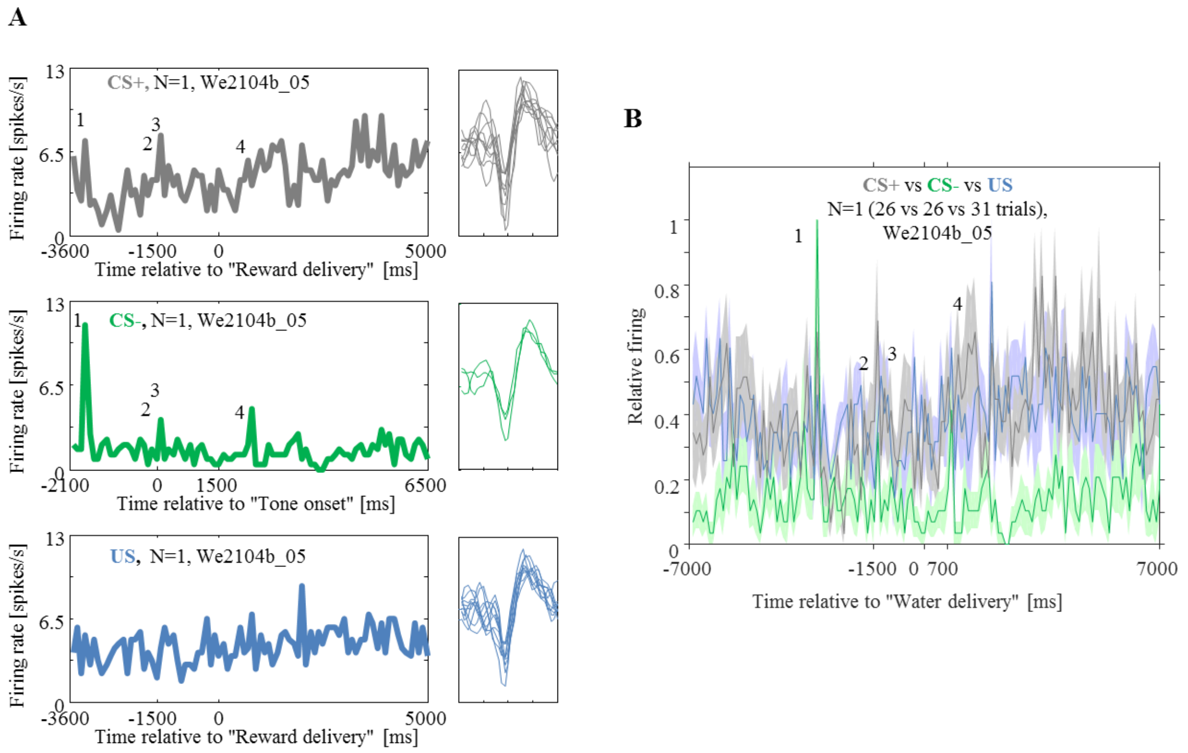

Figure 23. Example of one unit with a slow increase in the neuronal activity after the water delivery and spike forms of this unit.

Subplots (A, left) and (B) show PSTHs (with bin size 100 ms) of the CS+ (gray color), CS- (green color) and US (blue color) conditions. The numbers indicate the acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset. In the (B), shadows indicate SEMs of the PSTHs. Subplot (A, right) also presents spike forms of the units in the three conditions.

To find common properties of the units with slow modulation, we analyzed the averaged activities of the subgroup that revealed the slow modulation in the US condition (Figure 24B). The average activity of these units had similar increases after water delivery in the CS+ condition. We compared the average activities near the maximal activity, 2000-3000 ms after the water delivery, and found that the activity was slightly but significantly higher in the US condition (Wilcoxon signed rank test, $p = 0.046$). It is important to mention that the shift in the baseline activity of the US and CS+ conditions was rather constant during the time of the trials. Therefore, the differences, most probably, occurred due to this slight shift between the conditions but not due to the stronger increase in activity in the US condition.

Analysis of the tuning curves did not reveal any special properties of the units with the slow increase in activity after the water delivery. All of the 97 units had significant responses to at least one of the acoustical events and the subgroup of units had the responses to the acoustical events in the CS+ condition (see Figure 24B).

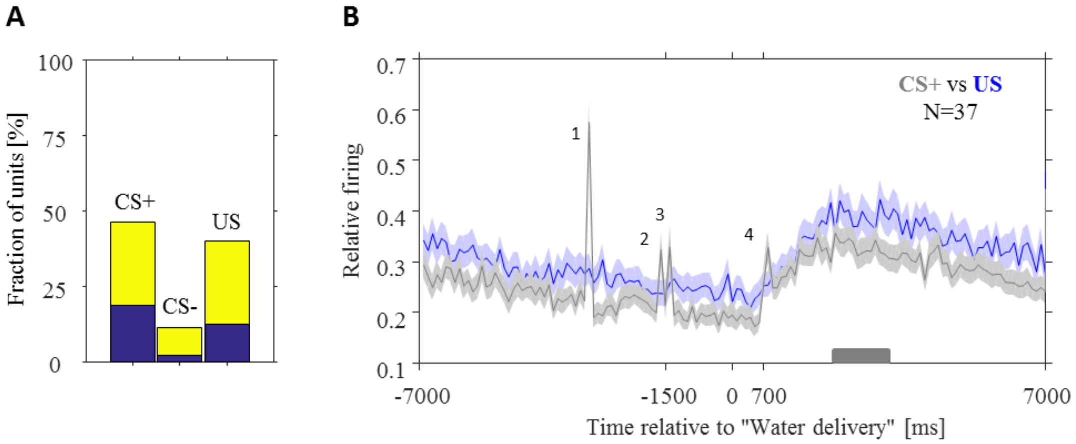

Figure 24. (A) Fraction of units that had significant increase in the activity in the conditions CS+, CS- and US. (B) PSTHs of the subgroup of units with significant increase in activity after the water delivery detected in the US condition.

(A) Blue and yellow parts of each bar in each histogram indicate the fractions of units of the monkey Ba and We from the population of 97 units respectively (ratio of units between the monkeys was ~1:2). (B) The subplot shows average PSTHs (bin = 100 ms) with SEM (shadow) of group of units during presentation of CS+ (gray color) and US (blue color) conditions. The subgroup showed significant increase in the neuronal activity in the US condition; the PSTH of the relative units is presented for the CS+ condition. The numbers indicate acoustical events: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset.

In contrast to the group of the low-trained monkeys, the group of the well-trained monkeys had more units with the slow increase after the water delivery. The slow modulations could be seen already in the population level (Figure 21) of the well-trained but not in the low-trained monkeys (Figure 10). Thus, we concluded that the experience led to larger sensitivity of the neurons in the auditory cortex to the presence of the unconditioned stimuli.

***

Summarizing this subsection, we revealed that the presence of unconditioned stimuli led to the upward shift of the baseline level. The presence of unconditioned stimuli led to the slow increase in the neuronal activity in almost half of the population. The frequency of the mouth movements correlated with the neuronal activities of some units. All over, we concluded that the neuronal activity of the experienced monkeys depended on the presence of unconditioned stimuli.

During acoustical stimulations, the neuronal population activity adjusted to the level of baseline was higher when the stimuli had no meaning. Response to the noise onset was also higher in the conditions without reward delivery. We showed that the well-trained monkeys associated the acoustical stimulation with the water delivery. Thus, we concluded that the sound meaning was the reason of the differences in the responses to the acoustical stimuli in the neuronal population.

3.3. Neuronal activity in the three instrumental and one passive condition

Many studies were made in the last two decades that seek for changes in the neuronal activities that related to complex abilities of animals such as sense of agency and level of effort. However, these studies did not exclude the presence of the unconditioned stimuli and sound meaning. In the present study, we implemented a design that allowed us to control these factors. The design consisted of four conditions, i.e., three instrumental and one passive. Before we control the effect of sense of agency and level of effort, we wanted to compare the neuronal activities and monkeys' behavior between the four conditions in order to find similarities and differences.

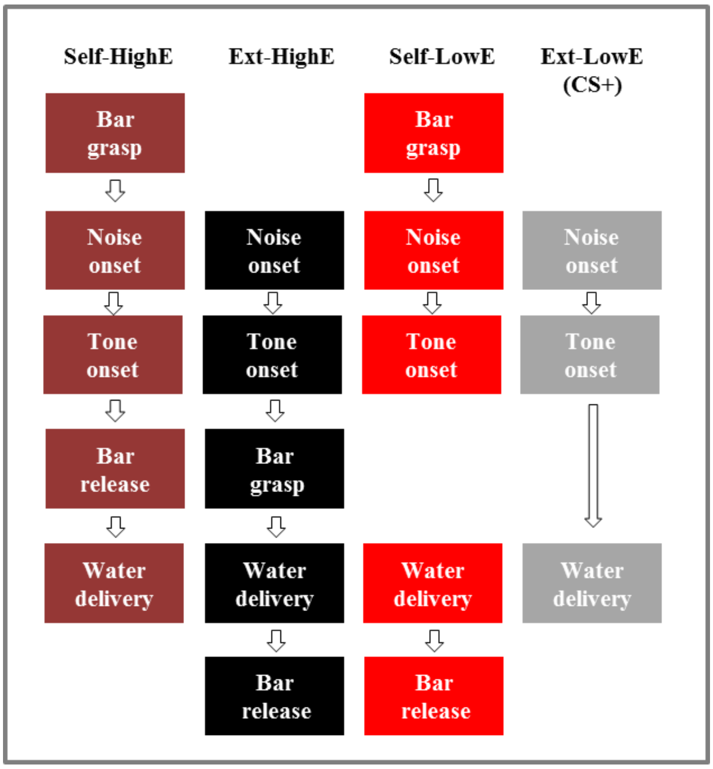

Figure 25. Scheme of the three instrumental (Self-HighE, Ext-HighE, Self-LowE) and one passive (Ext-LowE) condition[2].

Figure shows order of each of the four conditions. The scheme represents only one type of the acoustical stimuli of the two that were presented to the monkeys.

---

[2] The scheme of the four conditions was already presented in the Figure 1. The scheme is repeated and adapted on purpose to explain the comparisons of the present section.

For the experiment, two monkeys were trained to perform the three instrumental conditions (Figure 25). The passive condition was introduced to the monkeys in the first recording sessions. In the passive condition the monkeys were required to avoid interaction with the metal bar. For the analysis we used 72 units recorded in the auditory cortex of the first monkey (Ba, seventeen recording sessions) and 108 units (We, twenty one recording sessions) of another monkey. Because the neuronal activities were very similar between the two monkeys, the sample was combined into one with 180 units (see also Supplementary table 1). We identified that the first spike latencies of these units were 12.1± 8.3 ms after the onsets of the pure tones. All units of the sample responded to at least one acoustical event between the four, to the onsets and offsets of the pure tone and/or noise.

3.3.1. Number of errors did not relate to the difficulty of the conditions

Between the four conditions that the well-trained monkeys were presented with (Figure 25), three were instrumental and one was the passive condition that required an absence of any interaction with the metal bar. Two of the three instrumental conditions, the Self-HighE and Ext-HighE, required detection, the high effort, and were positively reinforced only after reaction in the specific time window. In contrast, the Self-LowE and Ext-LowE[3], the conditions with the low effort, did not require detection. Two of the conditions, the Self-HighE and Self-LowE, required self-initiation of the trials. One instrumental and the passive condition, the Ext-HighE and Ext-LowE, were initiated by the training computer. Because of the complexity of the conditions, we expected to find many errors made by monkeys during the sessions.

In the two self-initiated conditions, pauses between the end of the previous trial and the initiation of the new trial had to be longer than 4000 ms. As expected, it was difficult for the monkeys to withstand the pauses and sometimes they tried to initiate trials earlier. We revealed that the average relative frequency of the early initiation between the sessions was 6 and 14% in the Self-HighE and Self-LowE conditions respectively (Figure 26A). That was also predictable that during the pauses, which were determinated externally by the training computer in the Ext-HighE and Ext-LowE conditions, monkeys also tried to initiate the trials with a grasp of the metal bar. We found that they tried the initiation with similar frequency of 11 and 9% in the Ext-HighE and Ext-LowE conditions (Figure 26A, was also reported earlier, see Figure 15B). It is worthwhile to mention that some of the trials in the Ext-HighE conditions where the bar grasps were detected during the pause could in fact be not the initiation but very late detection of the pure tone. There was no way to find the real reason and type of the errors and we combined both of them.

Conditions with the high effort required detection of the pure tone in 300-1200 ms after the onset. The reaction time of the monkeys varied between the conditions (Figure 26D, E). The analyses revealed that the median of the reaction time in the Self-HighE condition was insignificantly different from the Ext-HighE condition in the types “tone followed by noise”

---

[3] The condition Ext-LowE was named CS+ in the previous sections. The differences of the names were made for emphasizing of the relationships between the conditions in the present section.

(Wilcoxon signed rank test, p = 0.049). In some of the trials, the pure tone was detected earlier than 300 ms (Figure 26B). Some of the pure tones were detected too early (Figure 26B).

Obviously, some of the trials were performed too slowly and the reaction times exceed the limit of 1200 ms (Figure 26C). However, we cannot be sure of whether such long reaction time was related to the late reaction or to the animals' confusion between the conditions. Thus, for example, the late reaction time in the condition Ext-HighE can be interpreted as the monkeys' performance of the condition Ext-LowE. The late reaction in the Self-HighE condition was, probably, a performance of the Self-LowE condition. The Self-LowE condition can be confused by the monkeys with the Self-HighE condition and be performed with the bar release. As well, the Ext-LowE condition can be confused with the Ext-LowE condition (was already described earlier, Figure 15A, Figure 26C).

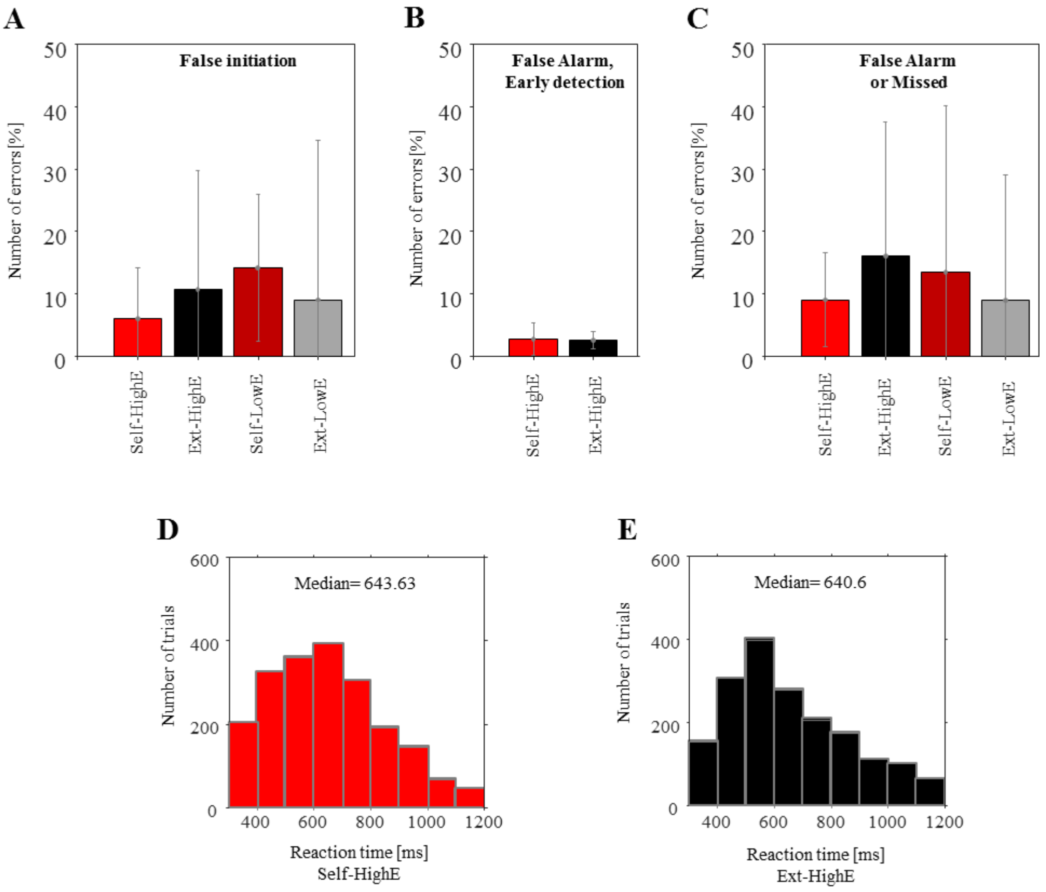

Figure 26. Relative frequency of errors (A, B, C) in the three instrumental and one passive condition and reaction time of detection in the two instrumental conditions (D, E).

The subplot (A) represents frequency of false initiations of the trials, (B) represents relative frequency of the early and late detection of the pure tones in the conditions with high effort, (C) represents relative frequency of the false alarms and/or missed trials. Two types of errors presented together because we do not know whether it was one or another error in each case. Whiskers indicate standard deviations of the mean between the sessions. The subplots (D) and (E) represent reaction times of detections in the Self-HighE and Ext-HighE conditions respectively. The reaction times of the type "tone followed by noise" are illustrated only.

The condition Ext-HighE was the easiest condition to train the animals that also took the smallest amount of the time between the three instrumental conditions. However, after all the three conditions and switches between them were successfully learned by the monkeys, the Ext-HighE condition was the last preferred by them. The performance of the condition can be characterized as inconsistent or fragmented. Usually, the monkeys made the successive trials of the Ext-HighE conditions in short sequences with long intervals between them. Thus, for instance, it happened that we were forced to stop a recording session because the monkeys denied performing the Ext-HighE condition. In contrast, it was always possible to motivate them to perform the Self-HighE or Self-LowE conditions. It is important to note that both monkeys preferred to begin the sessions and the search of the condition after a switch with the Self-LowE condition. Therefore, we supposed that the difficulty of the instrumental conditions increased from the Self-LowE to Self-HighE to Ext-HighE.

We hypothesized that number of errors and reaction time correlated with the monkeys' preferences of the conditions. The number of errors increased from the Self-HighE to Self-LowE to Ext-HighE condition (Figure 26A-C). Reaction time of the two conditions with high effort was also similar and insignificantly different for the type “tone followed by noise”. Thus, we did not find the correlations between the number of errors and the reaction times and the preferences of the monkeys.

3.3.2. Slow modulations in the neuronal activities in the four conditions were generally similar

In this subsection, we address the question whether there were some dramatic differences between the courses of the neuronal activities in the four conditions. For this, we first compared the PSTHs of the four conditions (Figure 27).

Figure 27 shows average neuronal activities of the 180 units recorded in the four conditions. As it can be seen from the figure, the time courses of the conditions, the slow modulations, were similar. The units responded to the acoustical stimulations (i.e., to the noise onset and offset, to the tone onset and offset) in all of the four conditions. The amplitudes of the responses were also similar between the conditions (see numbers 1-4 above the curves). The activities during the acoustical stimulation also did not differ drastically between the four conditions. The three instrumental conditions had higher baseline levels, the same as the Ext-LowE condition. We also found the slow increase in activity after the water delivery in all four conditions. Note, that the upward shift in the baseline levels and slow increase in the time corresponding to the water deliveries were not observed in the condition without presence of the unconditioned stimuli (see section 3.2).

However, we found two important differences between the neuronal activities. These differences were related to the requirements during the instrumental conditions, with the self-initiation of the trial and with the detection of the pure tone. Note that both were related to the monkeys' movements. Thus, the activities of the units in the two self-initiated conditions increased sharply after the grasps of the bar (see number 5 above the curves in the Figure 27), when no similar increase was observed in the two externally-initiated conditions. In the two conditions with high effort, in which the monkeys were required to detect the pure tone, the activities also increase sharply after the grasps/releases of the bar (see numbers 6 above the curves in the Figure 27) but no similar increase was observed in the conditions with low effort.

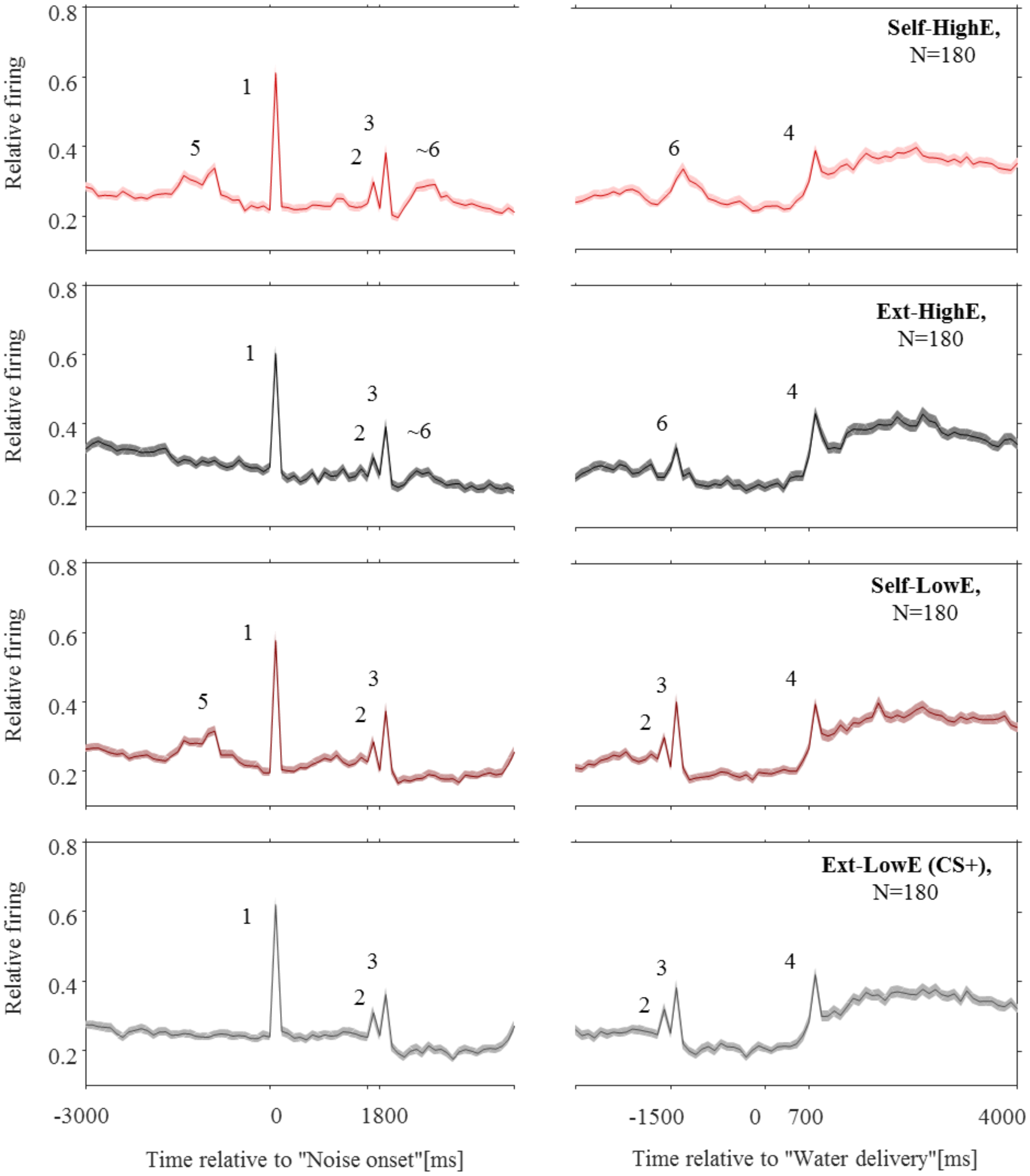

Figure 27. Average neuronal activity during presentation of the Self-HighE, Ext-HighE, Self-LowE and Ext-LowE condition.

The plot shows average PSTHs (bin = 100 ms) with SEM of the population, 180 units, in the Self-HighE (red shadow), Ext-HighE (black shadow), Self-LowE (bordo shadow) and Ext-LowE (gray shadow) condition. The left and the right parts of the subplots have time related to different events. Such consideration of the PSTHs was needed due to the different reaction times in the conditions where detection was required; the conditions with the low effort were presented in the same way for consistency. Numbers above the curved indicate acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset, 5. Grasp of the bar for the self-initiation of the trials, 6. Grasp/Release of the bar for detection of the tone onset.

The main focus of the present study was to find the specific effects of one of the two factors, i.e., the sense of agency and the level of effort. Due to this reason, we skipped the direct comparisons between each of the four conditions but will provide detailed comparisons of the pairs of the conditions in the next two sections (section 3.4 and 3.5).

3.3.3. Frequency of the mouth movements decreased after the self-initiation and increased after the detection

In one of the previous section (Section 3.2), we reported that the presence of the unconditioned stimuli led to the slow modulations in the neuronal activities of the well-trained monkeys. We described large fluctuations of the frequency of the mouth movements during the trials in the Ext-LowE condition. To the moment of the trial beginning, the frequency was minimal during the acoustical stimulation, and then increased slowly during ~ 2500 ms after the water delivery. Due to the presence of the unconditioned stimuli in the instrumental conditions and due to the prediction in the three instrumental conditions, we also expected to find such fluctuations in the frequency of the mouth movements in the Self-HighE, Self-LowE and Ext-HighE condition.

We had the opportunity to analyze seven video recordings (all with monkey Ba). Using the recordings, we calculated the frequency of the mouth movements. Figure 28 shows frequencies of the three sessions, each presenting results of the mouth movements in the four conditions. The examples show all variations that we could find between the four conditions in the seven records. The course of the mouth movements in the Ext-LowE condition repeated the course that we described previously. The analysis revealed that the increase in the all curves of the seven sessions was significant (Pearson correlation coefficient, $r > 0$, $p < 0.05$). In the seven sessions, the increases in the frequency of the mouth movements in the instrumental conditions repeated the increase observed in the Ext-LowE condition. The slow increases after the water delivery were also significant in the seven sessions in the three instrumental conditions ($r > 0$, $p < 0.05$). Moreover, other parts of the courses, the slow decrease in the frequency of the mouth movements to the beginning of the trials and minimal frequencies during the acoustical stimulations, were also similar between the four conditions.

However, we also found some differences between the four conditions. Thus, the frequency of the mouth movements decreased in the Self-HighE and Self-LowE conditions after the bar grasp (Figure 28A, B). This decrease developed shortly after the self-initiation and reached the minimum shortly before the noise onset. In ~700 ms after the noise onset, the frequency leveled to the frequency in the Ext-LowE and Ext-HighE condition. Visual analysis of the videos confirmed that after the initiation of the trials, the monkeys froze for some hundred milliseconds expecting the acoustical stimulation and gazed to the water tube. Similar decrease was observed in the self-initiated conditions after the grasp of the bar in the four of the seven sessions. It is important to note that the presence and the level of decrease were not related to the number of the recording session, i.e., the decrease did not intensify or diminish with experience. Thus, the frequency of the mouth movements plotted in Figure 26A was recorded in one of the first recording sessions, when other three were recorded in the three last recording sessions of the animal.

Another interesting detail that was observed in the frequency of the mouth movements was the increase in the Self-HighE and Ext-HighE conditions at the time of detection of the tone onsets (Figure 28C). The increase developed after the detection, reached the maximal value in ~1000 ms and leveled to the frequencies of the Self-LowE and Ext-LowE conditions shortly after the time of the water delivery, close to the moment of the tone offset. Visual analysis of the videos confirmed that the monkey, after the relatively low frequency of licking during the noise stimulation, briskly grabbed the water tube with lips shortly after the detection of the pure tone. It is worth pointing out that such changes were observed in both conditions with high effort, where detection of the tone was required, but was not present in the conditions with low effort. We found similar increase in the four of the seven recording sessions. The decrease was not related to number of the session, i.e., the frequency of the mouth movements plotted in the Figure 26C was recorded in one of the first sessions, when other three were recorded in the middle recording sessions of the animal.

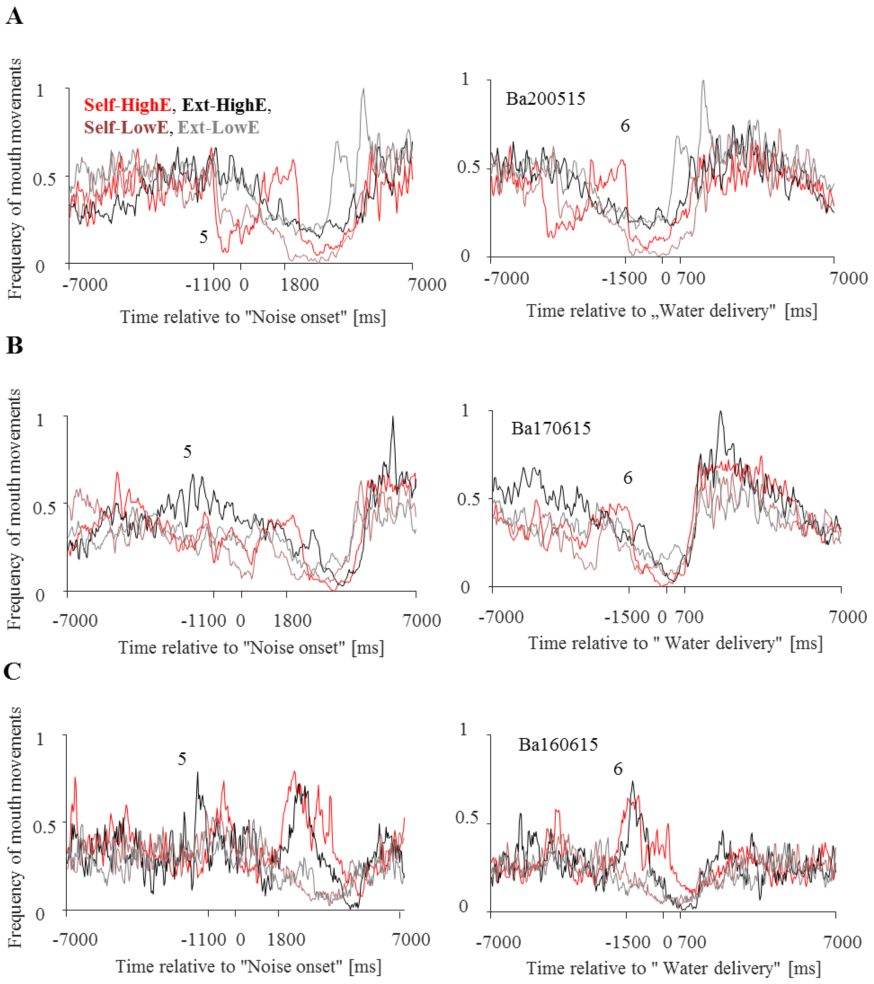

Figure 28. Frequencies of the mouth movements recorded in three sessions.

Subplots (A, B, C) present frequency of the mouth movements of three sessions in the four conditions. The left and right subplots have time related to different events. Numbers indicate time of the hand movements: 5. Grasp of the bar for the self-initiation, 6. Grasp/Release of the bar for the detection of the tone onset.

It is worthwhile to note that the sessions where the decrease after the self-initiation and increase after the detection were observed did not belong to the same sessions. Both, the decrease and increase, were found only in one of the seven recorded sessions. That indicates that the increase and decrease were not interrelated, i.e., the presence of one did not obligated the appearance of the second. Because the method of the detection of the mouth movements was used for the first time in the present study, we did not trace the recording picture, therefore in was not ideal for the analysis. Thus, the absence of both, the increases and decreases, does not guarantee that they indeed were not occurred. However, we suppose that when one of both was detected then the other would be detected too if it was present in the video.

More detailed analysis, which regards the increase and decrease in the frequency of the mouth movements after the self-initiation and detection respectively, and comparison of the frequency of the mouth movements with the neuronal activity recorded during these sessions, will be provided in the next two sections (section 3.4 and 3.5).

3.3.4. Half of units in the population had slow increase in activity after water delivery in the four conditions

In the previous sections (section 3.2.), we analyzed slow modulations that occurred in the neuronal activities when unconditioned stimuli were present. In this section, we questioned whether the three instrumental conditions also had the slow increases in their activities.

Figure 29 shows the fractions of units that had a significant increase in activity during 2500 ms after the water delivery in each condition (Spearman correlation coefficient, $r > 0$, $p < 0.05$). As it was already reported in the previous section, around half of the units had the slow increase in activity in the Ext-LowE condition. Thus, we validated the result in a greater amount of units. The sizes of fractions were very similar between the four conditions and the test did not revealed differences ($\chi 2 = 0.47$, df = 3, $p > 0.05$). Thus, the increase observed in the units was independent of the condition types.

In order to reveal whether the same units had the increase in the activity in different conditions, we chose the 108 units that showed the increase in the condition Ext-LowE and analyzed changes of activities in the same units in the other three conditions (Figure 30). As can be seen from the figure, neuronal activities in the four conditions had the slow increase after the water delivery. Thus, the presence of the slow increase after the water delivery in one condition interrelated with the presence of the increase in other three conditions. More detailed comparisons of the increases, regarding the factors of the sense of agency and level of effort, will be provided in the next two sections (section 3.4 and 3.5).

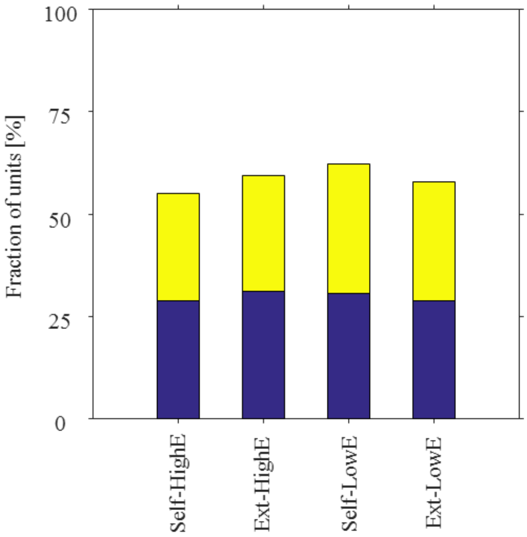

Figure 29. The fractions of units that had the slow increase in activity after the water delivery in the conditions Self-HighE, Ext-HigheE, Self-LowE and Ext-LowE.

The blue and yellow color parts of each bar indicate the fractions of units recorded in the monkeys Ba and We (the ratio of units between the monkeys was ~1:1.5).

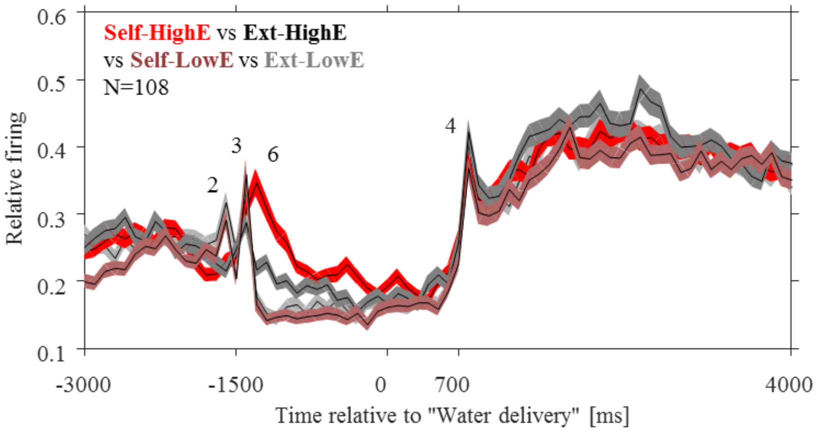

Figure 30. Average neuronal activity of the subgroup of units with the slow increase after the water delivery in the Ext-LowE condition.

The plot shows the average PSTHs (bin = 100 ms) with SEM of 108 units in the Self-HighE (red shadow), Ext-HighE (black shadow), Self-LowE (bordo shadow) and Ext-LowE (gray shadow) condition. The group showed significant increase in activity after the water delivery in the conditions Ext-LowE. The figure shows PSTHs of the same units in other three instrumental conditions. Note that the curve of the Ext-High condition is almost completely overlapped with the curve of the Self-LowE condition. The numbers above the curves indicate the acoustical events and movements: 2. Noise offset for the externally-initiated conditions, 3. Tone onset for the externally-initiated conditions, 4. Tone offset for all four conditions, 6. Grasp/Release of the bar for the conditions with high effort.

3.3.5. Number of units that responded to the movements depended on the importance of the movements

Some previous studies (Brosch et al. 2005, Brosch et al. 2015) revealed that the units in the primary auditory cortex responded to the movements that were related to the auditory tasks, i.e., to the bar grasp that led to the beginning of the tone sequence and to the bar release that predicted the offset of LEDs and stopped the tone sequence. In our study, the movements that were required from the monkeys could be divided in three categories: movements that led to the initiation of the trial (bar grasp in the conditions Sel-HighE and Self-LowE), detection of the pure tone (bar grasp and release in the conditions Ext-HighE and Self-HighE respectively) and continuation of a condition (bar release after the water delivery in the conditions Ext-HighE and Self-LowE). We were interested in whether the probability or amplitude of the responses to the movements differed between (1) bar grasps and releases, (2) initiation and detection, and between (3) more important (initiation and detection, before the water delivery) and less important (after the water delivery) movements.

To address these questions we first controlled whether the present neuronal sample had units that responded to the bar grasp and/or release. For this, we compared the firing rates of 100ms before and after the movements in each unit. The analysis revealed 15-22% of the units that had significant responses after the bar grasp in each of the three instrumental conditions (Wilcoxon signed rank test, $p < 0.05$, Figure 31A). Similar percentage of units, 19%, had responses to the release in the Self-HighE condition. Only 3-7% of units showed significant responses to the bar releases in the conditions Ext-HighE and Self-LowE (Figure 31B). Then we controlled the change of firing evoked by the movements. The changes of firing varied between 0.27-0.33 for the bar grasps (Figure 31C) and between 0.22-0.26 for the bar releases (Figure 31D).

Due to the absence of the bar release related to the initiation in the present design, we compared only the bar grasps and bar releases related to the detections (1). The number of units that responded to the bar grasps and releases related to detection did not differ significantly ($\chi^2 = 0.56$, df = 1, $p > 0.05$ between the conditions Self-HighE and Ext-HighE). The change of the firing was also insignificantly different (Student's t-test, $p = 0.12$). Therefore, we concluded that the number of units that responded and the amplitudes of the responses were independent of the direction of the movement, i.e., whether it was a bar grasp or a bar release.

We analyzed whether the number of units that responded to the movements which related to the initiation and detection and sizes of the responses differed significantly (2). In order to reveal the differences, we compared the responses to the bar grasps related to the initiation and detection. The analysis showed that the number of units that responded were similar between the movements for the initiation and detection ($\chi^2 = 0.28$, df = 1, $p > 0.05$ between the conditions Self-HighE and Ext-HighE and $\chi^2 = 1.0$, df = 1, $p > 0.05$ between the conditions Self-LowE and Ext-HighE). We also compared the changes of firing evoked by the movements and found that the responses evoked by the bar grasps in initiations and detections did not differ significantly (Student's t-test, $p = 0.35$ between the conditions Self-HighE and Ext-HighE and $p = 0.33$ between the conditions Self-LowE and Ext-HighE). We concluded that neither the number of the units nor the change of the firing was related to the type of the actions.

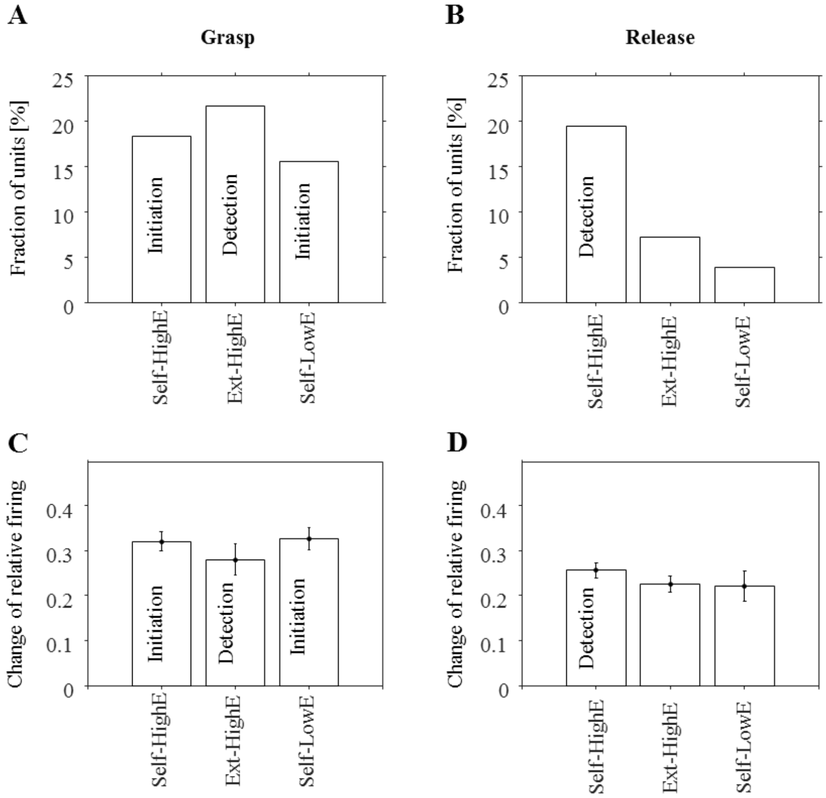

Figure 31. The number of units with responses to the bar grasps (A) and bar releases (B) in the population. Average changes of firing rate evoked by the bar grasps (C) and releases (D) in the units.

Note that the subplots (A) and (C) illustrates results of the bar grasps, the subplots (B) and (D) illustrate results of the bar releases. Error bars in the subplots (C) and (D) indicate SEM of the means.

The movements could be also differentiated by the level of importance. Thus, there were more important movements that determined the water delivery and less important movements that determined only the end of the trials. We controlled whether the importance of the movement was a factor that changed the number of the units with responses or sizes of the responses (3). Due to the absence of the less important movements with the bar grasps, we compared only the responses to the bar releases. Between the three types of the bar releases, we had the opportunity to compare only between the detection in the condition Self-HighE, the important movement, with the bar releases after the water delivery in the conditions Ext-HighE and Self-LowE, the unimportant movements. The analysis revealed that the number of the units that responded to the important movements were significantly higher than the number of units that responded to unimportant movements ($\chi^2 = 5.6$, df = 1, $p < 0.05$ between the conditions Self-HighE and Ext-HighE and $\chi^2 = 10.4$, df = 1, $p < 0.05$ between the conditions Self-HighE and Self-LowE). We also compared the changes of the firing evoked by the movements and found that they did not differ significantly (Student's t-test, $p = 0.31$ between the conditions Self-HighE and Ext-HighE, $p = 0.40$ between the conditions Self-

HighE and Self-LowE). Therefore, the number of units with responses depended on the importance of the movements.

3.3.6. Additional activity before the bar grasps in the three instrumental conditions

Analyzing the neuronal population activities, we found their increase not only after the self-initiation but also some time before the bar grasps (Figure 27, Figure 32A). We wondered whether this increase was related to the responses evoked by the movements. To address this question, we found all units in the population that had significant responses to the bar grasps in the self-initiated conditions and controlled whether the increase in activity was present in these units. Additionally, we averaged the activity of the units without significant responses to the bar grasps and controlled whether the increase in activity was absent in such units.

After comparisons of the average firings 100 ms before and after the bar grasps, we detected 51 units that had significant responses to the self-initiation (Wilcoxon signed rank test, $p < 0.05$, Figure 32B). As we assumed, the increase in the neuronal activity before the grasps accompanied the responses to the movements. The increase began, peaked and ended ~700-600, ~300 and ~100 ms before the movement respectively. The average firing of the rest of the 129 units, in which the response to the bar grasp was not detected, had no similar increase in activity before the movement (Figure 32C). Thereby, we concluded that the units that had the increase in the neuronal activity before the bar grasp in the self-initiated conditions also had the responses to this movement and the activities were related to each other.

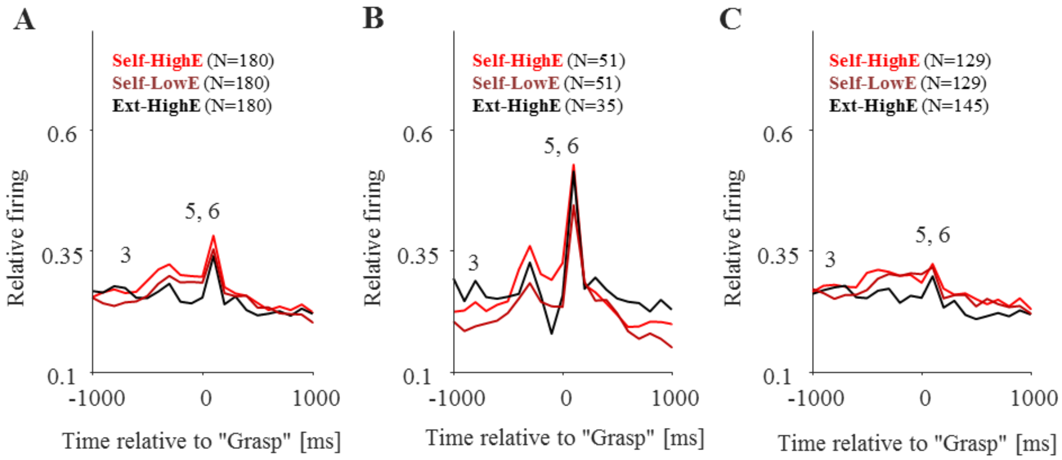

Figure 32. PSTHs of the population (A) and two subgroups of the units with (B) and without (C) responses to the bar grasps.

Each subplot shows PSTHs (with bin size 100 ms) in the Self-HighE (red), Self-LowE (bordo) and Ext-HighE (black) conditions. Numbers in parentheses near condition types indicate the number of units showed for a certain condition. The numbers above the curves indicate the acoustical events and movements: 3. Tone onset in the condition Ext-HighE, 5. Grasp of the bar in the self-initiated conditions, 6. Grasp of the bar in the condition Ext-HighE.

We assumed also to find a similar increase before the bar grasps in the Ext-HighE condition (Figure 27, Figure 32A). In order to address the question, we conducted the same procedures and found all units that responded to the bar grasp in the condition Ext-HighE. The analysis revealed 35 units that had responses to the bar grasp and 145 units without responses. Similarly to the units with responses to the bar grasps in the self-initiated conditions, the activity with responses to the bar grasps in the Ext-HighE conditions had increased before the movements. The increase began, peaked and ended ~700-600, ~300, ~100 ms before the movement respectively. In the units without responses to bar grasps, such increase in activity was not observed. Thereby, we concluded that the units that had the responses to the bar grasp also had the increase in activities before the movements.

3.3.7. Baseline levels differed between the four conditions

In the previous sections we revealed that the levels of baselines in different conditions may lead to differences in the absolute responses to the acoustical stimuli (section 3.1 and 3.2). Due to our intention to find dependencies of the sense of agency and the level of effort on the responses to the acoustical stimuli, we wanted to control whether the adjustment to the baseline levels was necessary.

To address this question, we compared average neuronal activities before the trial beginning in all four conditions. Previously, we also found that the neuronal activities of some units were modulated ~800-700 ms before the bar grasps in the conditions with the self-initiation (section 3.3.6). In order to exclude the changes, we compared an earlier time window, namely 1700-700 ms before the beginning of the trials. The analysis revealed that the averaged neuronal activities of the time window were significantly different between the four conditions ($\chi^2 = 35.89$, df = 537, $p < 0.001$). A more detailed analysis of the baselines is provided further and includes the dependencies on the sense of agency and the level of effort. To the present moment, we note that the baselines of the four conditions were different that indicates the need of their adjustments in further comparisons of the response to the acoustical stimuli.

3.4. Influence of sense of agency on neuronal activity

Some studies, in which neuronal activities in the auditory cortex of human and non-human primates were recorded, demonstrated that the activity showed inhibitory responses when the subject pronounced some words or vocalized respectively (Creutzfeldt et al. 1989, Müller-Preuss and Ploog 1981, Eliades and Wang 2003). The same neurons showed excitatory responses when the same sounds were played back. It is important to note that the studies did not consider the differences between the meaning of the sounds that were self-initiated and an absence of the meaning of the sounds when they were initiated externally. In our study, we intended to reveal whether the sense of agency of the artificial sounds will change the activity in the neuronal activity of the monkeys' primary auditory cortex. In contrast to the previous studies, we considered the meaning of the externally-initiated sounds.

In order to reveal such differences, we compared neuronal activities in the primary auditory cortex recorded during monkeys performed three instrumental conditions (Self-HighE, Ext-HighE and Self-LowE) and were presented with one passive condition (Ext-LowE, Figure 33). Between the four conditions, two required self-initiation of the trials, i.e. a grasp of the metal bar (Self-HighE and Self-LowE), and another two were initiated by the training computer. The same sample of the 180 units was used for the present comparisons as was introduced in the previous section (Section 3.3, see also Supplementary table 1).

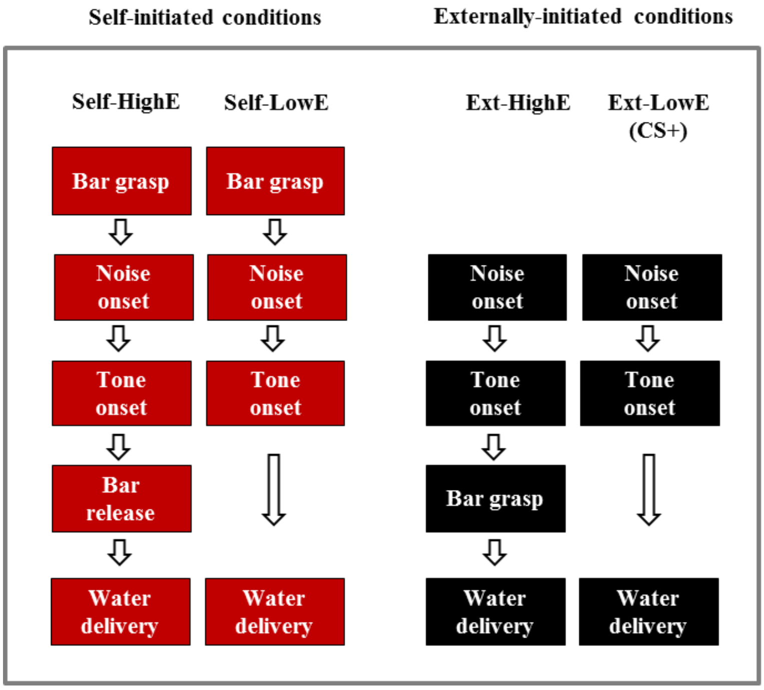

Figure 33. Scheme of the experimental design[4] adapted to show the contrast between the self- and externally-initiated conditions.

The figure shows the sequence of events in the four conditions. Two of the four conditions were self-initiated by the monkeys (red blocks) and the other two were externally-initiated by the training computer (black blocks). The scheme represents only one type of the acoustical stimuli from the two presented to the monkeys.

To address the question whether the sense of agency had an influence on the neuronal activity, we first compared the average activity in a time window of interest in the two self- and two-externally-initiated conditions (Wilcoxon signed rank test, $p < 0.05$). After these comparisons, we also controlled whether the pairs of the conditions had the same direction and significance of the results (Self-HighE vs Ext-HighE and Self-LowE vs Ext-LowE, Wilcoxon signed rank test, Bonferroni corrected $p < 0.05/2$). The result was accepted as

---

[4] The scheme of the four conditions was already presented in Figure 1. The scheme is presented on purpose to explain the comparisons of the present section. Note also that the order of the four conditions is changed to emphasize the differences between the pairs of conditions under the consideration of the present section.

reliable only if the three comparisons had the same directions of differences and were significant.

3.4.1. Responses to the tone onset adjusted to the baseline were higher in the self-initiated conditions

Because previous studies reported inhibitory responses to the pronounced words and vocalization compared with the play back of these sounds (Creutzfeldt et al. 1989, Müller-Preuss and Ploog 1981, Eliades and Wang 2003), we hypothesized that the response to the self-initiated acoustical stimuli will be lower than the responses to the externally-initiated stimuli. In order to control the hypothesis, we compared the absolute responses evoked by the acoustical events between the self- and externally initiated conditions. Figure 34 shows the average population activity of the sample of 180 units where peristimulus histograms of the two self-initiated (Self-HighE and Self-LowE, see also Figure 33) and two externally-initiated conditions (Ext-HighE and Ext-LowE) were averaged.

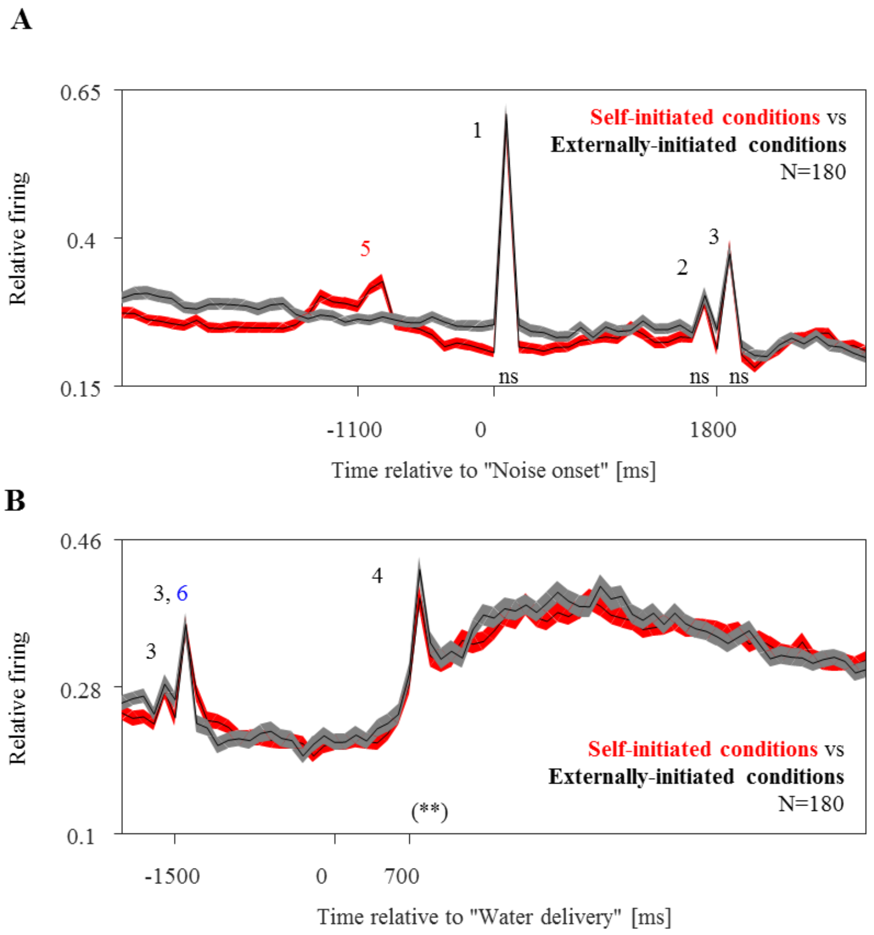

Figure 34. Average neuronal population activities in the self- and externally-initiated conditions.

The subplots show PSTHs (with bin size 100 ms) with SEM in the self-initiated (red shadow) and in the externally-initiated (black shadow) conditions. Subplot (A) represents PSTHs with time relative to the "Noise onset"; subplot (B) represents PSTHs with time relative to the "Water delivery". Note that the responses to the bar grasps/releases in the conditions with high effort were averaged with the response to the tone onset in the conditions with the low effort in subplot (B). Symbols at the bottom of the subplots indicate significance levels of comparisons between the responses to the acoustical events between the self- and externally-initiated conditions (Wilcoxon signed rank test, ns - $p > 0.05$, ** - $p < 0.01$). The symbols in parentheses indicate that the significance level was reached only for comparisons of the averages of the two self- and two externally-initiated conditions but not for the comparisons of the pairs of the conditions. The numbers above the curves indicate acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset, 5. Grasp of the bar for the self-initiation (red), 6. Grasp/Release of the bar for detection of the tone onset (blue).

For the comparisons of the responses evoked by the acoustical stimulations, we used the time windows 100 ms after the events. The analysis revealed that the responses to the noise onsets was smaller in the two averaged self-initiated conditions compared with the two averaged externally-initiated conditions (Table 3A, a copy of the table is in the Supplementary table 4A). The response to the noise onset in the Self-LowE condition was significantly lower than in the Ext-LowE condition. However, the response to the noise onset in the Self-HighE condition was similar compare with the Ext-HighE condition. We concluded that the responses to the noise onsets did not depend on the sense of agency. The responses to the noise offset did not differ significantly between the two averaged self- and two externally-initiated conditions. The pairs of the comparisons also did not have significant differences. Thus, the responses to the noise offsets also did not depend on the sense of agency. The responses to the tone onsets also did not differ significantly between the two averaged self- and two externally-initiated conditions and between the pairs of the comparisons. Thereby, the responses to the noise offsets did not depend on the sense of agency. Responses to the tone offsets were significantly lower in the two averaged self-initiated conditions compared with the externally-initiated conditions. In the conditions with high effort, the responses to the tone offsets were significantly lower in the self-initiated condition. However, in the conditions with low effort, the responses to the tone offsets did not differ significantly. Thus, we did not accept the effect of the sense of agency on the differences between the responses to the tone offset.

Comparison		Time window		Average of the Self- vs Ext- initiated conditions		Pairs of the Self- vs Ext-initiated conditions				Conclusion
		Main	Subtracted from		P-value	HighE	P-value	LowE	P-value	Significance
Absolute responses evoked by the acoustical events										
A	Noise onset	0-100 ms after event	-	<	0.03	=	0.67	<	$6*10^{-3}$	ns
	Noise offset	0-100 ms after event	-	=	0.86	=	0.75	=	0.31	ns
	Tone onset	0-100 ms after event	-	=	0.54	=	0.57	=	0.23	ns
	Tone offset	0-100 ms after event	-	<	$7*10^{-3}$	<	$4*10^{-3}$	=	0.12	ns
Change of the neuronal activity evoked by the acoustical events										
B	Noise onset	0-100 ms after event	-100-0 ms before event	>	$2*10^{-4}$	>	$4*10^{-5}$	=	0.46	ns
	Noise offset	0-100 ms after event	-100-0 ms before event	=	0.57	=	0.96	=	0.32	ns
	Tone onset	0-100 ms after event	-100-0 ms before event	>	$1*10^{-4}$	=	0.07	>	$2*10^{-4}$	ns
	Tone offset	0-100 ms after event	-100-0 ms before event	<	0.03	=	0.07	=	0.50	ns

Table 3. Effect of the sense of agency on the responses in the population activity evoked by the acoustical events.

Symbols in the conclusion indicate significance level of differences after the three comparisons: ns - $p > 0.05/2$.

We also controlled the influence of the sense of agency in each of the 180 units (Table 4A, a copy of the table is in the Supplementary table 5A). The analysis detected 2-8% of the units that had higher responses to the noise onset, noise offset, tone onset and tone offset in the two averaged self-initiated conditions compared with the two externally-initiated conditions. None of the units in the population had higher responses in the self-initiated conditions in all three comparisons. We found that 4-11% of the units had lower responses to the noise onset, noise offset, tone onset and tone offset in the two averaged self-initiated conditions compared with the two externally-initiated conditions. A very low fraction of units in the population, 0-2%, had significantly lower responses in the self-initiated conditions in all three comparisons.

Comparison		Time window		Self > Ext		Self < Ext	
		Main	Subtracted from	Averaged	Averaged and pairs	Averaged	Averaged and pairs
Absolute responses evoked by the acoustical events							
A	Noise onset	0-100 ms after event	-	7.22%	0%	10.56%	1.11%
	Noise offset	0-100 ms after event	-	6.11%	0%	5.56%	0.56%
	Tone onset	0-100 ms after event	-	6.11%	0%	4.44%	0%
	Tone offset	0-100 ms after event	-	2.78%	0%	10.56%	0%
Change of the neuronal activity evoked by the acoustical events							
B	Noise onset	0-100 ms after event	-100-0 ms before event	17.78%	3.89%	7.22%	1.67%
	Noise offset	0-100 ms after event	-100-0 ms before event	13.33%	3.33%	14.44%	0.57%
	Tone onset	0-100 ms after event	-100-0 ms before event	23.33%	5.56%	10%	1.67%
	Tone offset	0-100 ms after event	-100-0 ms before event	8.33%	5%	17.78%	1.67%

Table 4. Effect of the sense of agency on the responses of individual units evoked by the acoustical events.

The tables (A) and (B) show the fractions of units evoked responses of which depended on sense of agency.

To consider the differences between the baselines in the conditions (see section 3.3.7), we compared the changes of the neuronal activity evoked by the acoustical events (during 100 ms before and after the events). After calculating the ratio of the values in each unit for each of the four conditions, we subtracted the resulting values of the externally-initiated conditions from the resulting values of the self-initiated conditions. The obtained 180 values were compared with zero. The analysis revealed that the change of the neuronal activity evoked by the noise onset was significantly higher in the self- initiated conditions compare with the externally-initiated conditions (median = 0.03, Table 3B, a copy of the table is in Supplementary table 4B). However, the tests of the pairs revealed significant difference only between the conditions with the high effort but not between the conditions with the low effort. Thus, the response to the noise onset adjusted to the baseline level did not depend on the sense of agency. For the noise offset, the changes of the neuronal activity were similar between the self- and externally-initiated conditions (median = -0.002). The pairs of the conditions with high and low effort were also similar. Thereby, the response to the noise offset adjusted to the baseline level also did not depend on the sense of agency. For the tone onset, the changes of the neuronal activity were significantly larger in the self-initiated conditions (median = 0.02). Changes of the neuronal activities in Self-LowE conditions were higher than in the Ext-LowE conditions. However, the changes in the conditions Self-HighE and Ext-HighE did not differ significantly. We concluded that the response to the tone onset adjusted to the baseline level did not depend on the sense of agency. The changes of the neuronal activity evoked by the tone offset were significantly lower in the self-initiated conditions (median = -0.02). However, the pairs of comparisons revealed no significant differences between the changes of the neuronal activities. Thereby, the response to the tone offset adjusted to the baseline level did not depend on the sense of agency.

We made a similar analysis for each of the 180 units (Table 4B, a copy of the table is in Supplementary table 5B). The analysis revealed that the changes of the neuronal activity evoked by the noise onset, noise offset, tone onset and tone offset were larger in the self-initiated conditions in 8-23% of units. In the population only 3-6% of units also had significantly larger changes of the neuronal activities evoked by the acoustical events in the three comparisons. The results revealed that 7-18% of units had significantly smaller changes of neuronal activities evoked by the acoustical events in the self-initiated conditions. Only 0-2% of units showed significantly lower changes of the neuronal activities in the self-initiated conditions in all three comparisons.

3.4.2. The average activity during the noise stimulation was lower in the self-initiated conditions

Previous studies of the sense of agency revealed that the neuronal activity was inhibited during the time of speaking or vocalization (Creutzfeldt et al. 1989, Müller-Preuss and Ploog 1981, Eliades and Wang 2003). In order to control whether the self-initiated acoustical stimuli led to a similar effect in the neuronal activity, we compared the average activities during the noise and tone stimulations.

To address the question, we averaged the activities in the time window from 200-1600 ms after the noise onset. For the tone, we included 1400 ms before the water delivery. Thus, we excluded the responses evoked by the onset of the acoustical stimuli and the response evoked by the movements in the conditions with high effort.

The average neuronal population activity during the noise stimulation was significantly lower in the two self-initiated conditions compared with the two externally-initiated conditions (Figure 35A, Table 5, a copy of the table is in Supplementary table 4C). Moreover, the comparisons between the pairs also revealed significant differences. The activities were significantly lower in both self-initiated conditions with high and low effort. Thus, the average neuronal activity during the noise stimulation depended on the sense of agency. The neuronal activities during the tone stimulation were similar in the self- and externally-initiated conditions (Figure 35B). The average activity during noise stimulation in the Self-HighE condition was significantly higher than in the Ext-HighE condition. The average activity in the Self-LowE condition did not reveal significant differences from the Ext-LowE condition after a Bonferroni correction. Thereby, the average activity during the tone stimulation did not depend on the sense of agency. The difference between the neuronal activities during the noise stimulation could be also seen in the figure that represent the average population activity (Figure 34A). Unlike to the noise stimulation, the average neuronal activities during the tone stimulation were similar in the self- and externally-initiated conditions in the population level (Figure 34B). Some examples of the individual units in which the differences during the noise stimulation were observed are presented in the Figure 36A, B.

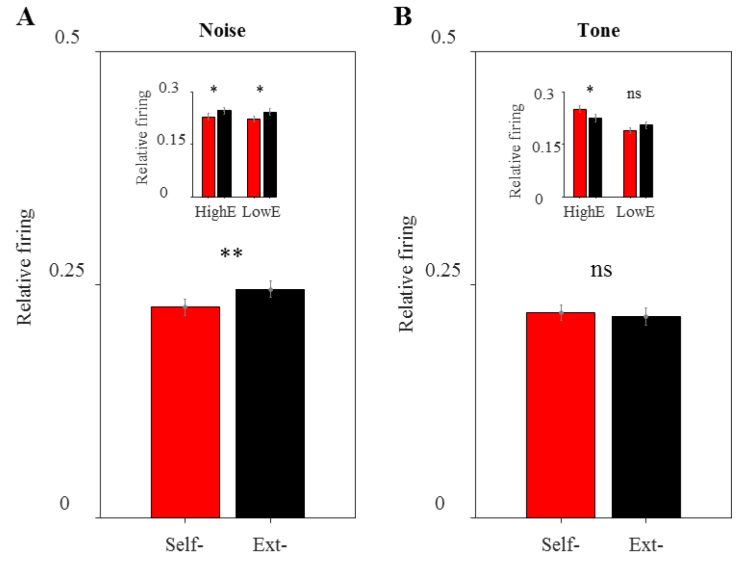

Figure 35. Average neuronal activities during the noise (A) and tone (B) stimulation in the self- and externally-initiated conditions.

The bars in (A) indicate the average neuronal activities during the noise stimulation (200-1600 ms after the onset). The bars in (B) indicate the average neuronal activities during the tone stimulation (during 1400 ms before the water delivery). The stars indicate the significance levels of differences between the activities in the self- and externally-initiated conditions (Wilcoxon signed rank test, ns - $p > 0.05$, ** - $p < 0.01$). Inserts in (A) and (B) show the average neuronal activity during the noise and tone stimulations respectively in each of the four conditions. Average neuronal activities of the self-initiated and externally-initiated conditions colored with red and black respectively. The stars indicate significance levels of differences between the neuronal activates in the pairs of conditions (Wilcoxon signed rank test, ns - $p > 0.05/2$, * - $p < 0.05/2$). Whiskers in the general plots and in the inserts indicate SEM of the means.

Table index	Comparison	Time window		Average of Self- vs Ext-initiated conditions		Pairs of Self- vs Ext-initiated conditions				Conclusion
		Main	Subtracted from		P-value	HighE	P-value	LowE	P-value	Significance
Absolute responses evoked during acoustical stimulations										
A	Noise	200-1600 ms after event	-	<	$1*10^{-3}$	<	0.02	<	0.01	*
	Tone	-1400-0 ms before water delivery	-	=	0.6	>	0.02	=	0.03	ns
Change of the neuronal activity evoked during the acoustical stimulation										
B	Noise	200-1600 ms after event	-1700 – -700 before trial beginning	=	0.07	=	0.03	<	$4*10^{-8}$	ns
	Tone	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	>	$7*10^{-10}$	>	$7*10^{-12}$	<	$1*10^{-3}$	ns

Table 5. Effect of sense of agency on the neuronal activity in the population during the acoustical stimulations.

Symbols in the conclusion indicate significance level of differences after the three comparisons: ns - $p > 0.05/2$, * - $p < 0.05/2$.

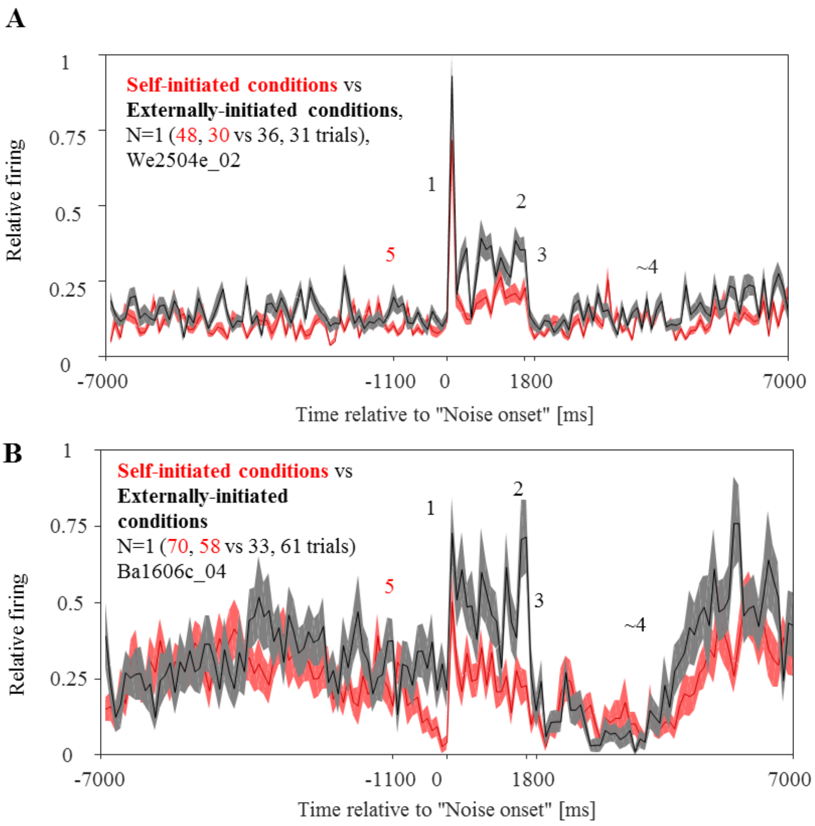

Figure 36. Examples of units (A and B) where the neuronal activity during the noise stimulation was lower in the self-initiated conditions.

The figure presents PSTHs (with bin size 100 ms) with SEM of individual units in the self-initiated (red shadow) and externally-initiated (black shadow) conditions. The numbers above the curves indicate the acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar for the self-initiation (red).

Individual analyses of each of the 180 units showed that ~27% of the units had lower and only ~12% of units had higher activity during the noise stimulation in the two self-initiated conditions (Table 6A, a copy of the table is in Supplementary table 5C). In the population, only 1-3% of the units revealed significant differences in the neuronal activity during the noise stimulation between the self- and externally-initiated conditions in the three comparisons. During the tone stimulation, ~24 and ~17% of units showed significantly higher or lower activity in the two self-initiated conditions respectively. Less than 3% of units in the population had differences in the activities during the tone stimulation between the self- and externally-initiated conditions in the three comparisons.

Table index	Comparison	Time window		Self > Ext		Self < Ext	
		Main	Subtracted from	Averaged	Averaged and pairs	Averaged	Averaged and pairs
Absolute responses evoked during acoustical stimulations							
A	Noise	200-1600 ms after event	-	11.67 %	2.22 %	26.67 %	6.11 %
	Tone	-1400-0 ms before water delivery	-	24.4 %	1.67 %	16.67 %	2.22 %
Change of the neuronal activity evoked during the acoustical stimulation							
B	Noise	200-1600 ms after event	-1700 – -700 before trial beginning	13.33 %	1.67 %	26.11 %	3.33 %
	Tone	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	25.56 %	6.67 %	20 %	2.22 %

Table 6. Effect of the sense of agency on the neuronal activities of the individual units during the acoustical stimulations.

The tables (A) and (B) show the fractions of units evoked responses of which depended on the sense of agency.

It is worth taking into account the differences of the baselines that we observed between the four conditions (section 3.3.7). In order to consider possible influences of the baselines, we also compared the changes of the neuronal activities using the average *z*-scored neuronal activity during the noise stimulation in the self- and externally-initiated conditions (200 -1600 ms after the noise onset). Similar procedures were made for the averaged *z*-scored neuronal activities during the tone stimulation (during 1500 ms before the water delivery). For each unit, the resulting values for the externally-initiated conditions were also subtracted from values of the self-initiated conditions. Further, the medians of the 180 values were compared with zero. For the noise stimulation, the analysis revealed that the median of the values was not significantly different from zero (median = 0.08, Table 5B, copy of the table is in Supplementary table 4D) that indicated that the changes of neuronal activities were larger in the self-initiated conditions. The analysis of the pair of the conditions with high effort also revealed a positive median that was not significantly different from zero after Bonferroni correction (median = 0.16). However, the comparisons of the conditions with the low effort showed that the median was negative. That indicated that the change of the neuronal activity in the Self-LowE conditions was smaller than in the Ext-LowE condition (median = -0.33). The median was also significantly different from zero. Thus, the differences had opposite directions. That indicated that the changes of the neuronal activities during the noise stimulation adjusted to the baseline level did not depend on the sense of agency. Similar results were observed for the tone stimulation, where the median was positive and significantly different from zero for comparison of the average self- and externally-initiated conditions (median = 0.45). The comparison of the pairs showed that the median was also positive and significantly different from zero for the conditions with the high effort (median = 0.59) but negative and significantly different from zero for the conditions with the low effort (median = -0.33). Thus, the change of the neuronal activity evoked during the tone stimulation did not depend on the sense of agency.

We also made an analysis of the neuronal activities during the acoustical stimuli adjusted to the baseline level in 180 individual units (Table 6D, a copy of the table is in Supplementary table 5D). We found that ~13% of the units had significantly larger and ~26% of units had significantly smaller change of the neuronal activities for the noise stimulation in the self-initiated conditions compare with the externally-initiated conditions. Less than 3% of units had significant differences between the self- and externally-initiated conditions in the three comparisons. The number of the units, where the changes of the neuronal activities for the tone stimulation was significantly larger and smaller in the self-initiated conditions, were ~26 and 20% respectively. Less than ~7 and 2% of units had significant differences between the changes in the self- and externally-initiated conditions in the three comparisons.

3.4.3. Absence of units with a slow increase (or decrease) in the neuronal activity that depended on the sense of agency

Previous studies showed that some of units in the auditory cortex slowly increased or decreased in activity when the sound was associated with reinforcement (Abolafia et al. 2011, Fritz et al. 2010, Quirk et al. 1997, Shinba et al. 1995). In the paradigm presented in our experiment, we had a possibility to find out whether the neuronal activity increased (or decreased) slowly during the acoustical stimulations and to control the dependence of the changes on the sense of agency.

In order to reveal a slow sustained increase or decrease during the noise stimulations, we used the time window of the noise (tone) stimulation. For the noise stimulation, we considered the time 200-1600 ms after the onset. For the tone, we considered 1400 ms before the water delivery in order to exclude influence responses to detection in the conditions with the high effort. We controlled whether the activity increased or decreased during the time of the acoustical stimulation in the two averaged self-initiated conditions (Pearson correlation coefficient, $r \geq 0.5$ or $r \leq -0.5$ respectively) but did not decrease or increase in the two externally-initiated conditions ($r < 0.5$ or $r > -0.5$ respectively) in each unit. We calculated the number of such units in the population. After that, we compared the numbers of units, which satisfied the requirement. The presence of an increase (or decrease) in the self-initiated condition was accepted only if the units did not have an increase (or decrease) in the externally-initiated conditions. For additional control, we calculated the percentage of units that had the increase (or decrease) in the externally-initiated but not in the self-initiated conditions. The dependence of the increase on the sense of agency was accepted only if the percentages differed significantly.

The analysis revealed a very small amount of units (less than 5%, Table 7), which was lower than by chance, that would satisfy the requirements reported above and showed slow changes in the self- but not in the externally-initiated conditions. Also very few units were detected having the slow changes in activities in the externally-initiated but not in the self-initiated conditions. From the result, we concluded that the sense of agency did not influence the slow changes in activities during acoustical stimulations.

Change during		Noise stimulation		Tone stimulation	
Only in conditions with		Self-initiation	Ext-initiation	Self-initiation	Ext-initiation
Type of change	Slow increase	1.11%	0%	0.56%	1.11%
	Slow decrease	0%	0 %	2.22%	0 %

Table 7. Percentage of units where the slow increase or decrease in activity was detected during the noise or tone stimulation.

3.4.4. Baseline levels did not depend on the sense of agency

Our findings described in the present study revealed that the baseline activities varied between the conditions where the unconditioned stimuli were and were not presented (section 3.2.5). Particularly, we found that the presence of unconditioned stimuli led to an upward shift in the baseline level. Moreover, we reported that the baseline levels differed between the four presented conditions (section 3.3.7). We wondered whether the levels of the baselines depended on the sense of agency.

The durations of the pauses preceding the trial beginning were in average similar between the four conditions. The first event in the externally-initiated conditions was the onset of the noise, the first event in the self-initiated conditions was the grasp of the bar. Previously, we described that the neuronal activities before the bar grasp had some increase during ~700 ms (section 3.3.6). For the analysis, we wanted to exclude all changes in the neuronal activities related to the bar grasp. Thus, we compared the neuronal activities between the pairs of the self- and externally-initiated conditions during 1700-700 ms before the beginning of the trials (Figure 37).

The analysis revealed that the two self-initiated conditions had significantly lower baselines than the two externally-initiated conditions (Wilcoxon signed rank test, p=0.01, Supplementary table 4E). However, the pairwise comparison of the conditions with different levels of effort revealed that the activity in the Self-HighE condition was significantly lower than in the Ext-HighE conditions ($p=2*10^{-6}$) but the activities were similar between the conditions with low efforts (p=0.94). From the results, we concluded that level of baselines did not depend on the sense of agency.

Comparison of the neuronal activities during the time windows between the conditions in individual units revealed that ~17% of the units had lower activity in the two self-initiated conditions (Wilcoxon signed rank test, $p < 0.05$, Supplementary table 5E) but only ~3% of the 180 units had lower activity in the three comparisons (Wilcoxon signed rank test, $p < 0.05$, Wilcoxon signed rank test, $p < 0.05/2$). Only ~8% of the units had higher activity in the two self-initiated conditions and no units had higher activity in all three comparisons.

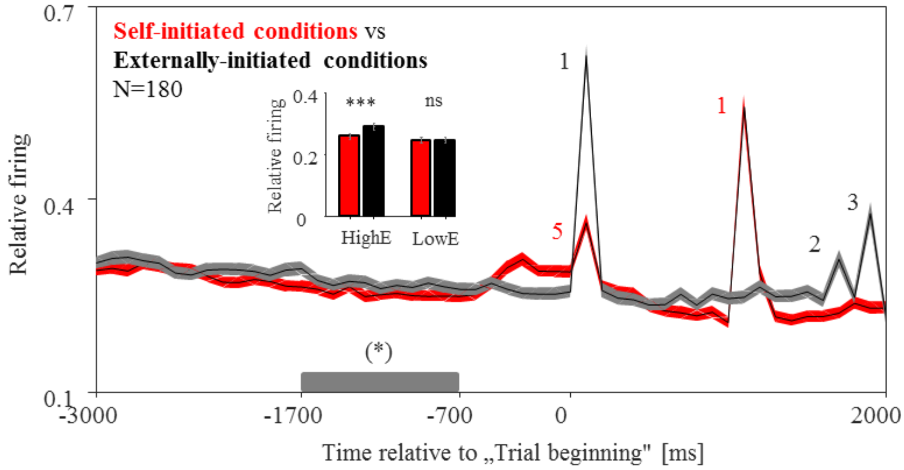

Figure 37. Average neuronal population activity in the self- and externally-initiated conditions.

The plot shows PSTHs (with bin size 100 ms) with SEM in the self-initiated (red shadow) and in the externally-initiated (black shadow) conditions. The time of the PSTHs is related to the beginning of the trial that was the bar grasp in the self-initiated conditions and the noise onset in the externally-initiated conditions. The horizontal gray bar indicates the time interval where the baseline levels were compared. The stars in the parentheses indicate the significance levels of the differences between the self- and externally-initiated conditions (Wilcoxon signed rank test, * - $p < 0.05$). The stars are in parentheses because only one pair of conditions showed significance. The numbers above the curves indicated the acoustical events and movements (red and black colors for the self- and externally-initiated conditions respectively): 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar. The insert shows average neuronal activities during the marked time window in the four conditions. The red bars indicate the self-initiated conditions; the black bars indicate the externally-initiated conditions. The stars indicate the significance levels (Wilcoxon signed rank test, ns - $p > 0.05/2$, *** - $p < 0.001/2$). Whiskers show SEM of the means.

3.4.5. The neuronal activity before the noise onset depended on the sense of agency

One of the studies (Eliades and Wang 2003) reported that the inhibition in the neuronal activity started shortly before the beginning of the animal’s vocalization. The researchers assumed that the inhibition was related to the preparation to the movement. A similar effect was observed in other studies in the auditory cortex of rodents, where the inhibition of the neuronal activity started to increase after the animals initiated a trial before the acoustical stimulation (Buran et al. 2014, Carcea et al. 2017). In order to control whether we also find the inhibition in the neuronal activity after the initiation, we compared the activities between the self- and externally- initiated conditions before the noise onset.

Figure 34A showed that the neuronal activities in the self-initiated conditions decreased slowly after the grasp of the bar. Quantitative analysis of the population revealed that, indeed, the average activity during the 500 ms before the noise onset was significantly lower in the self-initiated conditions compared with the externally-initiated conditions (Wilcoxon signed rank test, $p=4*10^{-7}$, Figure 38, Supplementary table 4F). The comparisons of the conditions with high and low effort also showed that the activities were significantly lower in the self-initiated conditions ($p=1*10^{-8}$ and $p=4*10^{-4}$ respectively). Thus, the effect was significant and very strong in the population level. Some examples of such units are presented in Figure 36B, Figure 39, Figure 41B. In the examples, the activity decreased after the self-initiation of the trials, after the grasp of the bar.

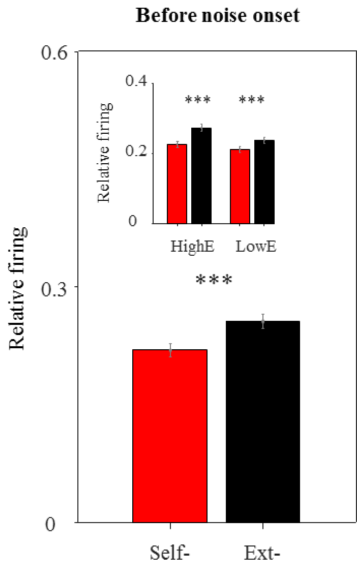

Figure 38. Average neuronal activity before the noise onset in the self- and externally-initiated conditions.

The bars indicate the average firing during the last 500 ms before the noise onset in the self-initiated (red) and externally-initiated (black) conditions. The stars indicate the significance level of differences between the firing in the self- and externally-initiated conditions (Wilcoxon signed rank test, *** - $p < 0.001$). The insert indicates average firing of the same time window in each of the four conditions. The whiskers show the standard deviations of the means. The stars indicate significance level of differences between the firing in the self- and externally-initiated conditions (Wilcoxon signed rank test, *** - $p < 0.001/2$).

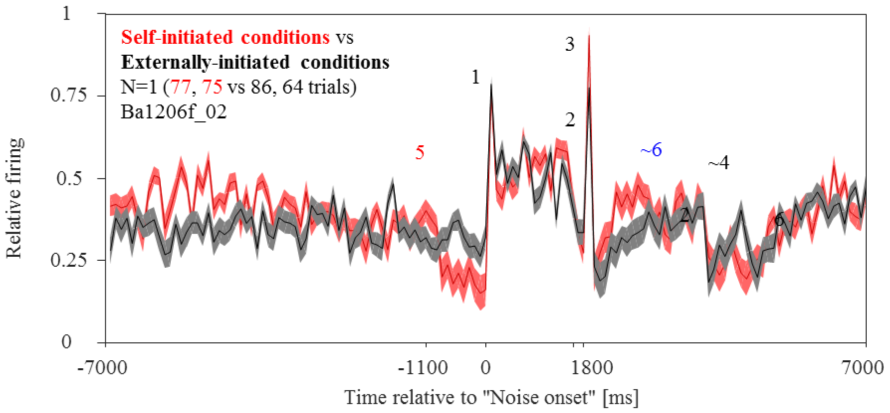

Figure 39. Examples of unit with a decrease in activity after the self-initiation.

The figure shows PSTHs (with bin size 100 ms) with SEM of individual unit in the self-initiated (red shadow) and externally-initiated (black shadow) conditions. Note that the unit had no response to the bar grasps. The numbers above the curves indicate the acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset, 5. Grasp of the bar for the self-initiation (red) 6. Grasp/Release of the bar for detection (blue).

We also compared individual units separately for the differences between the activities during the 500 ms before the noise onset (Supplementary table 5F). The analysis revealed that ~31% of the units had a lower activity during the time window in the self-initiated conditions. However, between the 180 units, ~7% of the units showed that the activity of the both pairs differed in the same directions. Between the 180 units, only 8% of them had a higher activity in the self-initiated conditions before the noise onset. None of the 180 units had significant effects in the two pairs of comparisons.

Because the differences between the conditions increased after the grasp of the bar, the self-initiation, we wondered whether the presence of the inhibition and the responses to the bar grasps were related, whether the same units had the changes in activities. In order to address the question, we selected all units in the population where the units responded to the bar grasp, i.e., the activities during 100 ms before and after the bar grasp were significantly different (Wilcoxon signed rank test, $p < 0.05$). Using the test, we found 51 units that had the response to the bar grasp at least in one of the self-initiated conditions (Figure 40A). In the other 129 units, the activity did not show responses to the bar grasp (Figure 40B). Further, we compared the activities during 500 ms before the noise onset between the self- and externally-initiated conditions in each of the two groups. We supposed that if the response to the bar grasp and the decrease in activity were related, then we will see the differences in the group with the response to the bar grasp and will not see the differences in the other group. The analysis of the group with the response to the bar grasp revealed that the activity in the self-initiated conditions was significantly lower than in the externally initiated conditions (Wilcoxon signed rank test, $p=0.03$). The comparisons of the pairs of the conditions also showed that the average activity before the noise onset was not significantly different between the conditions with high effort after Bonferroni correction ($p=0.03$). The average activity before the noise onset was significantly lower in the Self-LowE condition compared with the Ext-LowE ($p=6*10^{-3}$). Thus, the average activity before the noise onset did not differ significantly between the self- and externally-initiated conditions. The analysis of the group without the response to the bar grasp revealed that the activities before the noise onset were significantly lower in the self- initiated conditions ($p=3*10^{-6}$). Pairwise comparisons of the conditions with the high and low effort also showed significance and lower activity in the self-initiated conditions ($p=7*10^{-8}$ and $p=0.01$). Thus, the average activity before the noise onset was significantly different between the self- and externally-initiated conditions. Therefore, the assumption that the inhibition after the self-initiation was related to the response to the bar grasp was rejected because the first group showed an absence of significant differences in activities before the noise onset; and because the second group showed even stronger differences in the time window. One unit with the response to the bar grasp but without the decrease after the initiation is presented in Figure 41A. One example unit with the response to the bar grasps and with decrease in activity before the noise onset is presented in Figure 41B. Two units without response to the bar grasp but with decrease after the initiation are presented in Figure 36B, Figure 39.

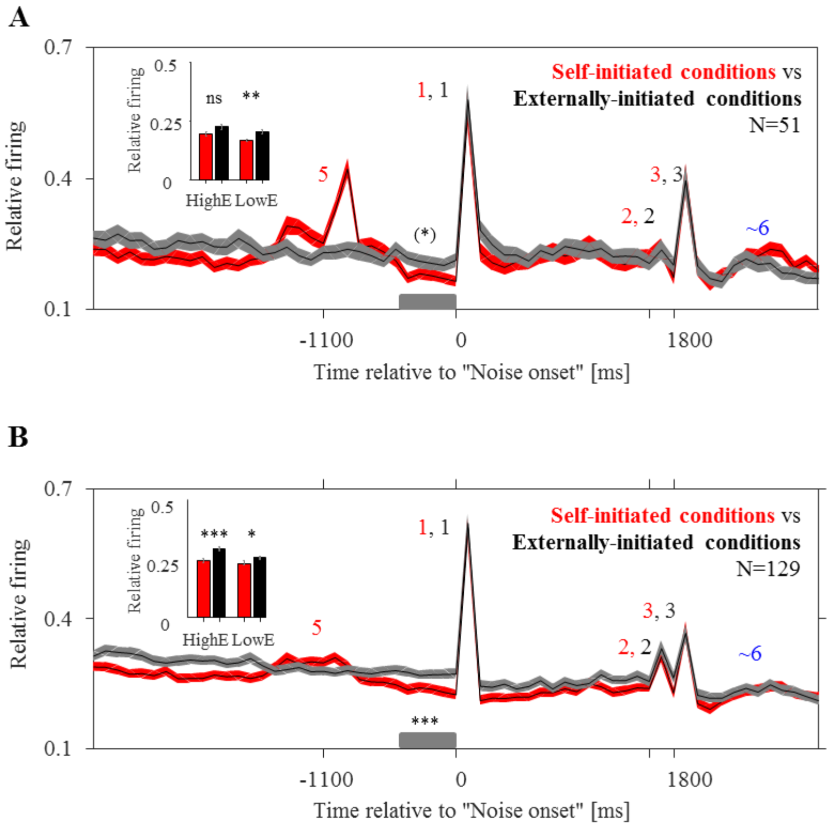

Figure 40. Average neuronal activity of the units that had (A) and did not have (B) response to the bar grasp.

The subplots show PSTHs (bin size 100 ms) with SEM in the self-initiated (red shadow) and in the externally-initiated (black shadow) conditions. Subplot (A) represents the group of units where the response to the bar grasp was significant (comparison of the activity during 100 ms before and after the bar grasp, Wilcoxon signed rank test, $p < 0.05$). Subplot (B) represents the group of units where the response to the bar grasp was insignificant (Wilcoxon signed rank test, $p > 0.05$). Horizontal gray bars indicate the time window where the activities were compared. The stars indicate the level of significance differences (Wilcoxon signed rank test, * - $p < 0.05$, *** - $p < 0.001$). The star in parentheses indicates that not both pairs of conditions differed significantly and had the same direction of differences. The numbers above the curves indicate the time of the acoustical events and movements (red, black and blue colors indicate events in the self- and externally-initiated conditions and conditions with high effort respectively): 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar for the self-initiation, 6. Grasp/Release of the bar for detection. The inserts with the bar graphs show the average neuronal activities during 500 ms before the noise onset in the four conditions (red and black colors indicate self- and externally-initiated conditions respectively). The stars indicate the level of significance differences (Wilcoxon signed rank test, ns - $p > 0.05/2$, * - $p < 0.05/2$, ** - $p < 0.01/2$, *** - $p < 0.001/2$). The whiskers show the SEM of the means.

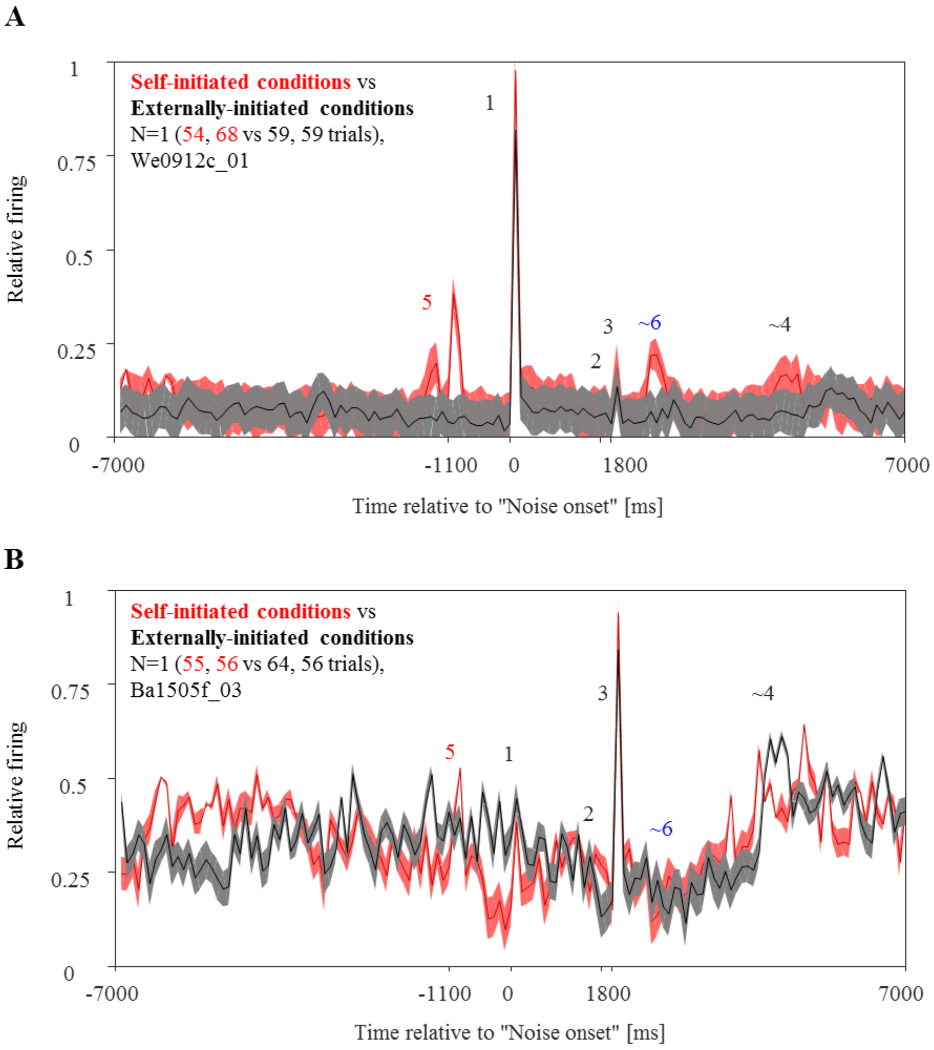

Figure 41. Examples of units without (A) and with (B) decrease in activity after the self-initiation.

The figure presents PSTHs (bin size 100 ms) with SEM of individual units in the self-initiated (red shadow) and externally-initiated (black shadow) conditions. Subplot (A) shows a unit that had a response to the bar grasps and had no decrease in activity before the noise onset in the self-initiated condition. Subplot (B) shows a unit without a response to the bar grasp but with a slow increase after the water delivery and with an inhibition before the noise onset in the self-initiated condition. The numbers above the curves indicate the acoustical events and movements (red, black and blue colors indicate events in the self- and externally-initiated conditions and conditions with high effort respectively): 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset, 5. Grasp of the bar, 6. Grasp/Release of the bar.

In the previous sections we described already that many units had a slow increase in activity after the water delivery (section 3.3.4). After the slow increase, the activity decreased slowly to the time of the next trial. We supposed that the decrease in activity after the bar grasp in the self-initiated conditions could be related to the changes in activity that we observed in some units after the water delivery. In order to test the hypothesis, we selected all units from the sample of 180 that had the significant slow increase in activity after the water delivery (Spearman’s correlation coefficient, $r > 0$ and t-test, $p < 0.05$). The test detected that around half of the units suited the requirements (Figure 42A). The other units of the population were in the second group in which the slow changes after the water delivery were not detected

(Figure 42B). After the selection process, we calculated the differences between the neuronal activities in the self- and externally-initiated conditions in the 500 ms time windows before the noise onsets. The analysis revealed that the activities in the self-initiated conditions were significantly lower in the group where the slow changes were detected (Wilcoxon signed rank test, $p=2*10^{-6}$). Both condition with the high and low effort had significantly lower neuronal activities when they were self-initiated ($p=3*10^{-5}$ and $p=7*10^{-5}$ respectively). Analyses of the group without the slow increase after the water delivery also showed significant differences between the activities with the self- and externally-initiated conditions (p=0.02). The neuronal activity in the Self-HighE condition was also significantly lower than in the Ext-HighE condition ($p=1*10^{-4}$). The neuronal activity in the Self-LowE condition was similar to the activity in the Ext-LowE condition (p=0.41). Thus, not all three comparisons showed the significant results. Therefore, we concluded that the slow changes in activity were related to the decrease after the bar grasp. An example of the unit without slow changes in activity and without decrease after the self-initiation was presented in Figure 41A. The units with the slow changes in activity and with the decrease after the self-initiation were presented in Figure 39A, B and Figure 41B.

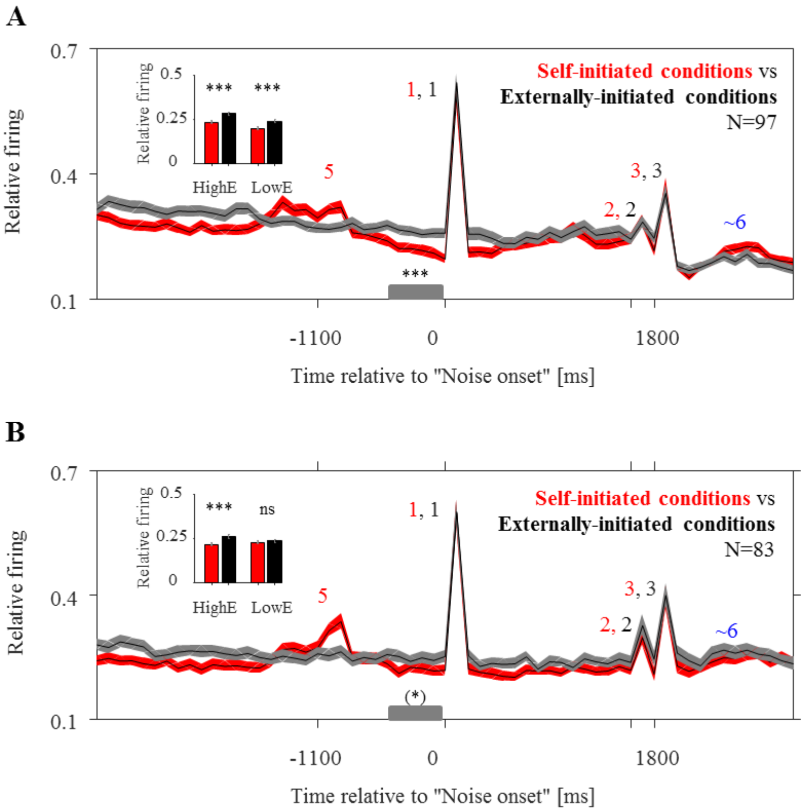

Figure 42. Average neuronal activities of the units that had (A) and did not have (B) significant increase in the activity after the water delivery.

The subplots show PSTHs (with bin size 100 ms) with the SEM of the self-initiated (red shadow) and in the externally-initiated (black shadow) conditions. Subplot (A) represents the group of units where the slow increase in activity after the water delivery was significant (2500 ms after the water delivery, Spearman's correlation coefficient, r > 0, t-test, $p < 0.05$). Subplot (B) represents the group of units where the slow increase in activity after the water delivery was not significant. The horizontal gray bars indicate the time window where the activities were compared. The stars indicate the level of significance differences (Wilcoxon signed rank test, * - $p < 0.05$, *** - $p < 0.001$). The stars in the parentheses indicate that the significance level was reached only between the averaged self- and externally-initiated conditions but not between both pairs in the conditions with high and low effort. The numbers above the curves indicate the time of the acoustical events and movements (red, black and blue colors indicate events in the self- and externally-initiated conditions and conditions with high effort respectively): 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar, 6. Grasp/Release of the bar. The inserts show the averages of the neuronal activities during 500 ms before the noise onset in the self- and externally- initiated conditions (red and black color respectively). The whiskers show the standard deviations of the means. The stars indicate the significance level (Wilcoxon signed rank test, ns - > 0.05/2, *** - p < 0.001/2).

3.4.6. Decrease in the frequency of the mouth movements after the self-initiation of the trials

In the previous section (section 3.3.3), we showed some examples of the frequencies of the mouth movements recorded for the well-trained monkey. We found that in the most of the sessions, the frequency of the mouth movements increased slowly after the water delivery and reached the maximum in ~2500 ms. Additionally, we showed some of the sessions where the frequency of the mouth movements decreased after the self- initiation, after the bar grasps in conditions Self-HighE and Self-LowE. We wondered whether units which were recorded during the sessions also had the decrease in the neuronal activities after the self-initiation.

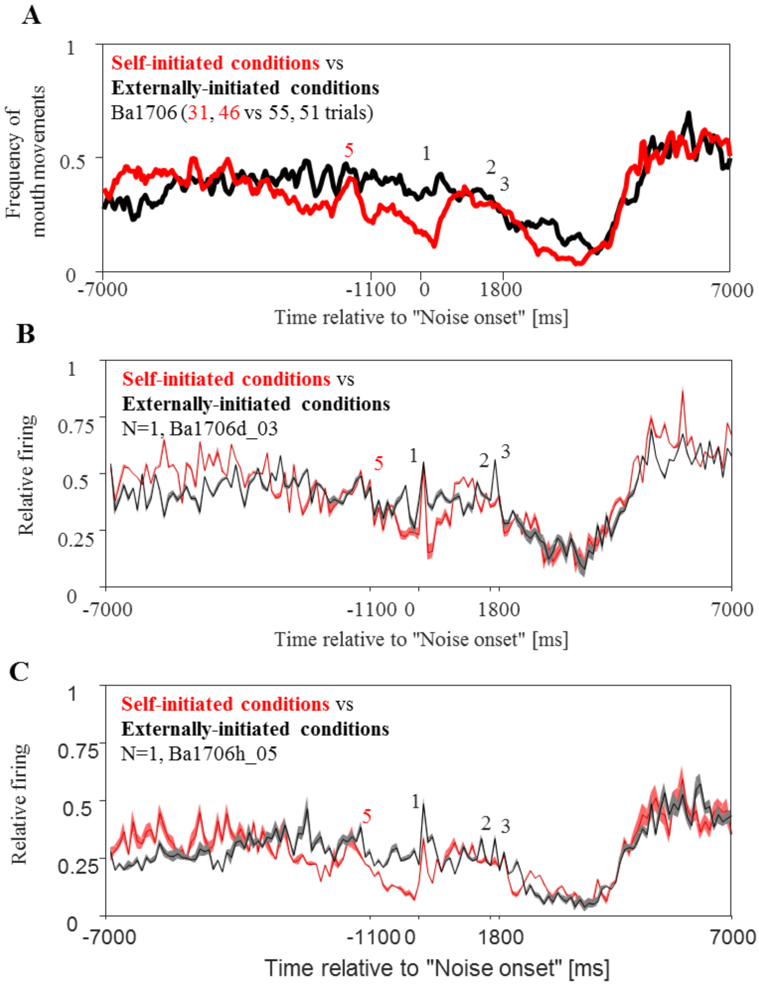

Figure 43. Frequency of the mouth movements (A) and two examples of units (B, C) that were recorded during the same session.

Subplot (A) represents the frequency of the mouth movements during one session, the subplots (B) and (C) show PSTHs (with bin size 100 ms) with the SEM of the two units recorded during this session. The colors of the curves indicate the self-initiated (red) and externally-initiated (black) conditions. Subplot (B) shows a unit that had a decrease in activity after the self- initiation. Subplot (C) shows a unit without a decrease in activity after the self-initiation. The numbers above the curves indicate the time of the acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar (red color).

For this purpose, we analyzed the neuronal activities of all units during the sessions where the decrease was observed after the self-initiation. Having four sessions with the decrease, we were able to analyze 17 units recorded during these sessions (5, 3, 3 and 3 units in each of the session respectively). Between all of them, we observed 9 units (5, 2, 2 and 1 units for each session respectively) with a decrease in the neuronal activity after the self-initiation. One example of the frequency of the mouth movements and two examples of the units recorded during the same session are presented in Figure 43. In the figure, the frequency of the mouth movements was averaged between the two self- and two externally-initiated conditions in order to stress the decrease after the bar grasp (Figure 43A). The two examples of the units had the decrease in the neuronal activity after the grasp of the bar that was similar to changes in the frequency of the mouth movements (Figure 43B-C). From the results we concluded that the presence of the decrease in the frequency of the mouth movements does not obligate the neuronal activity to have the same decrease after the self-initiation.

3.4.7. Level of increase in activity after water delivery did not depend on the sense of agency

Here we controlled whether the level of the slow increase in activity depended on the sense of agency. For this, we compared the average neuronal activities during the 1000 ms period of the maximal strength, i.e., 2000-3000 ms after the water delivery. We found that the average activities during the time period were significantly lower in the self-initiated conditions (Wilcoxon signed rank test, p=0.03). The conditions with high effort also differed significantly, the average activity in the self-initiated condition was lower (Wilcoxon signed rank test, $p=2*10^{-3}$). However, the comparison of the conditions with low effort did not reach the level of significance (p=0.36). Thus, we concluded that the sense of agency did not affect the level of the increase after the water delivery (Supplementary table 4G).

A comparison of the average neuronal activities during the time window between the conditions in individual units showed that 20% of them had a lower average activity in the self-initiated conditions (Wilcoxon signed rank test, $p < 0.05$, Supplementary table 5G). A similar size of the fraction, 10%, showed that the activity was significantly higher in the self-initiated conditions (Wilcoxon signed rank test, $p < 0.05$). Between 180 units, only ~2% showed a lower activity after the water delivery in both self-initiated conditions; and only ~1% showed significantly higher activity in both self-initiated conditions.

3.5. Influence of the level of effort on neuronal activity

Researchers made many attempts to reveal the influence of the engagement in the neuronal activity of the auditory cortex previously. However, the studies compared conditions in which one was reinforced and others were a passive presentation of acoustical stimuli (Atiani et al. 2014, Otazu et al. 2009, Fritz et al. 2005, Shinba et al. 1995 and others). Thus, the presence of the unconditioned stimuli was not considered in their comparisons. In the present study, we considered the presence of unconditioned stimuli and compared conditions with two levels of effort.

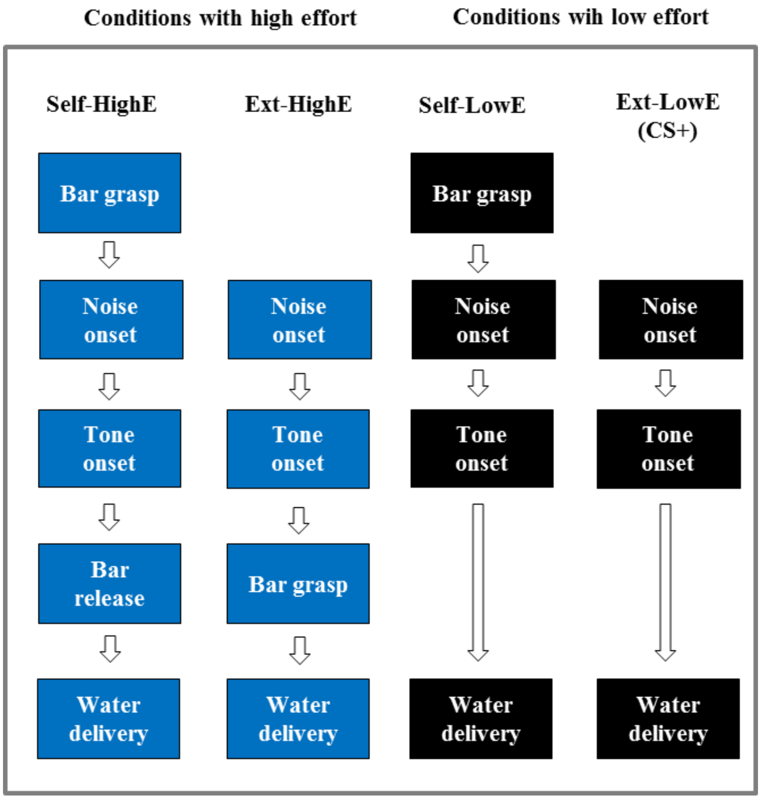

Figure 44. Scheme of the experimental design[5] adapted to show the contrast between the conditions with high and low effort.

The figure shows the sequence of events of the four conditions. The four conditions varied in the level of efforts. Two of the four conditions required high effort (blue blocks) and the other two required low effort (black blocks) from the monkeys. The scheme represents only one type of the acoustical stimuli from the two presented to the monkeys.

In order to control the differences, two monkeys were trained to perform two conditions with high effort and two conditions with low effort (Figure 44). In the conditions with high effort, the monkeys had to detect the onset of the pure tone. Additionally, one of the conditions was required being self-initiated (Self-HighE) and one was initiated externally by the training

---

[5] The scheme of the four conditions was already presented in the Figure 1. The presence scheme is presented on purpose to explain the comparisons of the present section. Note also that the order of the four conditions is changed to stress the differences between the pairs of conditions under the consideration of the present section.

computer (Ext-HighE). The second pair of the conditions with a low level of effort also had one condition which the monkeys initiated themselves (Self-LowE) and another condition that was initiated externally by the training computer (Ext-LowE, previously named CS+). Thus, using the four conditions, we had possibility to reveal the effect of the level of effort in the neuronal activity of the monkeys. The neuronal activities were recorded while the monkeys performed the four conditions. The same sample of 180 units was used for the present comparisons as was introduced in the previous section (Section 3.3, see also Supplementary table 1).

To address the question whether the level of effort had an influence on the neuronal activity, we first compared the average activity in time window of interest in the two conditions with high and two with low effort (Wilcoxon signed rank test, $p < 0.05$). After these comparisons, we also controlled whether the pairs of the conditions (Self-HighE vs Self-LowE and Ext-HighE vs Ext-LowE, Wilcoxon signed rank test, Bonferroni corrected $p < 0.05/2$) had the same direction and significance of the results. The result was accepted as reliable only if the three comparisons had the same directions of differences and were significant.

3.5.1. Responses to the acoustical events did not depend on the level of effort

First, we controlled whether the responses to the acoustical events depended on the level of effort (Figure 45). For this, we compared the average responses to the acoustical events (during 100 ms after the events) between the conditions with high and low effort (Table 8A, a copy of the table is in Supplementary table 6A). The analysis revealed that the responses to the noise onset were similar between the conditions with high and low effort. The responses to the noise onset in the Self-HighE condition were not significantly different from the responses in the Self-LowE condition after Bonferroni correction. The responses to the noise onset in the Ext-HighE and Ext-LowE condition were similar. Thus, the responses evoked by the noise onset did not depend on the level of effort. The responses to the noise offset were insignificantly different between the conditions with low and high effort. Two pairs of the self-initiated and externally-initiated conditions also did not reveal significant differences that indicated that the response to the noise offset did not depend on the level of effort. The responses to the tone onset were significantly higher in the conditions with high effort. Responses in the pairs of conditions were similar. Thus, the response to the tone onset did not depend on the level of effort. The responses to the tone offset were similar in the conditions with high and low effort and in the two pairs of conditions. Thus, the responses to the tone offset did not depend on the level of effort.

We also controlled the influence of the level of effort in each of the 180 units (Table 9A, a copy of the table is in Supplementary table 7A). The analysis revealed that the fractions of units, which had significantly higher responses to the noise onset, noise offset, tone onset and tone offset in the conditions with high effort, were 7-12%. Out of the 180 units, less than 2% had the same direction of significant differences between the three comparisons. The sizes of fractions that had higher responses to the acoustical events in the conditions with low effort were 4-10%. Very small fractions of 0-1% of units showed significant difference with the same directions between the three comparisons.

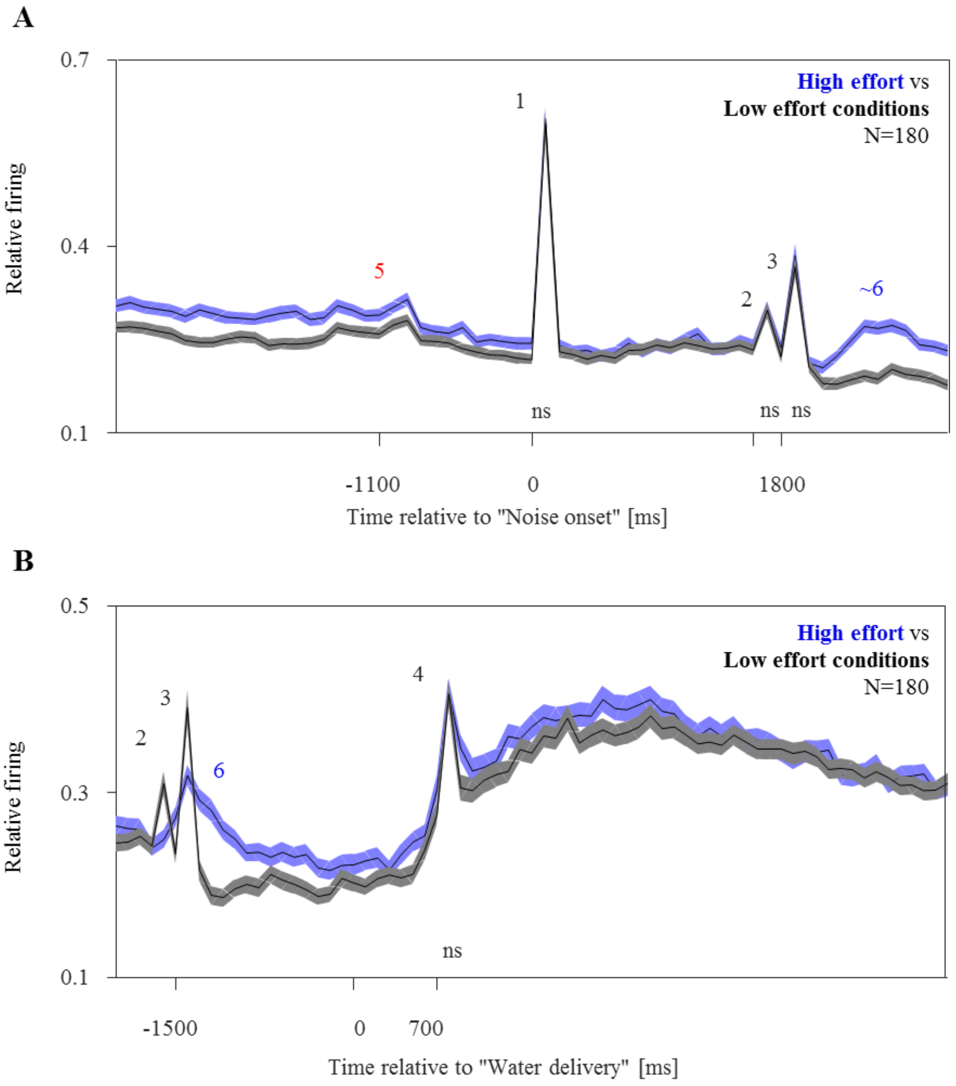

Figure 45. Average neuronal population activities (A and B) in the conditions with high and low effort.

The subplots show PSTHs (with bin size 100 ms) with the SEM in the conditions with high (blue shadow) and low (black shadow) effort. Each curve represents the average of the two conditions: Self-HighE and Ext-HighE (blue color), Self-LowE and Ext-LowE (black color). Subplot (A) shows PSTHs with time relative to the “Noise onset”. Note that the responses to the bar grasp in the self-initiated conditions were averaged with activity without event in the externally-initiated condition. Subplot (B) represents PSTHs with time relative to the “Water delivery”. Symbols at the bottom of the subplot indicate the level of significance between responses to the acoustical events (Wilcoxon signed rank test, ns - $p > 0.05$). The numbers above the curves indicate acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 4. Tone offset, 5. Grasp of the bar for self-initiation, 6. Grasp/Release of the bar for detection.

In order to consider the differences between the baselines between the conditions (see section 3.3.7), we compared the changes of neuronal activities before and after the acoustical events (during 100 ms before and after of the events). The average changes of neuronal activities in the two conditions with low effort were subtracted from the average activities in the two conditions with high effort. The resulting median of the 180 values were compared with zero using the Wilcoxon signed rank test. The analysis revealed that the median calculated for the noise onset was not significantly different from zero (median = -0.02, Table 8B, a copy of the table is in Supplementary table 6B) that indicated that the changes were similar to the conditions with the low effort. The analysis of the pairs revealed that the median calculated for the externally-initiated conditions was negative and significantly different from zero but the median of the self-initiated conditions was not significantly different from zero. Therefore, we concluded that the responses evoked by the noise onset and adjusted to the baseline level did not depend on the level of effort. For the noise offset, the change of the firing was similar between the conditions with the low and high effort (median = $6*10^{-4}$). Neither the pair of the self-initiated conditions, nor the pair of the externally-initiated conditions revealed significant differences of medians from zero. Thus, the responses evoked by the noise offset and adjusted to the baseline level did not depend on level of effort. For the tone onset, the median was not significantly different from zero (median = -0.003). The median for the self- and externally-initiated condition did not differ from zero significantly. That indicated that the responses evoked by the tone onset and adjusted to the baseline level did not depend on the level of effort. For the tone offset, the changes of the activities were significantly smaller in the conditions with high effort (median = -0.01). However, only one pair of the comparisons repeated the result. The pair of the self-initiated conditions showed that the change of the activity was smaller in the conditions with high effort when the comparison of the externally-initiated conditions did not reveal significant differences. From the obtained results, we concluded that the change of the response evoked by the tone offset and adjusted to the baseline level also did not depend on the level of effort.

Table index	Comparison	Time window		Average of High vs Low effort conditions		Pairs of High vs Low effort conditions				Conclusion
		Main	Subtracted from		P-value	Self-	P-value	Ext-	P-value	Significance
Absolute responses evoked by the acoustical events										
A	Noise onset	0-100 ms after event	-	=	0.42	=	0.04	=	0.47	ns
	Noise offset	0-100 ms after event	-	=	0.69	=	0.38	=	0.47	ns
	Tone onset	0-100 ms after event	-	>	0.03	=	0.32	=	0.11	ns
	Tone offset	0-100 ms after event	-	=	0.14	=	0.25	=	0.73	ns
Change of neuronal activity evoked by the acoustical events										
B	Noise onset	0-100 ms after event	-100-0 ms before event	=	0.07	=	0.65	<	$9*10^{-3}$	ns
	Noise offset	0-100 ms after event	-100-0 ms before event	=	0.64	=	0.91	=	0.22	ns
	Tone onset	0-100 ms after event	-100-0 ms before event	=	0.73	=	0.37	=	0.40	ns
	Tone offset	0-100 ms after event	-100-0 ms before event	<	0.04	<	$2*10^{-3}$	=	0.66	ns

Table 8. Effect of level of effort on the responses in the population activity evoked by the acoustical events.

The tables (A) and (B) show the dependence of the response to the acoustical events on the level of effort. Symbols in the conclusion indicate significance level of differences after the three comparisons: ns - p > 0.05/2.

We made similar analysis for each of the 180 units (Table 9B, a copy of the table is in Supplementary table 7B). The analysis revealed that the change of the neuronal activities evoked by the noise onset, noise offset, tone onset and tone offset were larger in the conditions with high effort in 12-18% of units. Out of the 180 units, less than 5% also had significantly larger changes of activities evoked by the acoustical events in the conditions with high effort in all three comparisons. The result revealed 10-18% of units that had significantly smaller changes in the neuronal activities evoked by the acoustical events in the conditions with high effort. Out of the 180 units, only 2-4% showed significantly smaller changes of the activities in the three comparisons.

Table index	Comparison	Time window		High > Low		High < Low	
		Main	Subtracted from	Averaged	Averaged and pairs	Averaged	Averaged and pairs
Absolute responses evoked by the acoustical events							
A	Noise onset	0-100 ms after event	-	7.78 %	1.11 %	7.78 %	0 %
	Noise offset	0-100 ms after event	-	7.78 %	0.56 %	9.44 %	0.56 %
	Tone onset	0-100 ms after event	-	8.33 %	0 %	4.44 %	0.56 %
	Tone offset	0-100 ms after event	-	11.11 %	1.11 %	7.78 %	0.56 %
Change of firing evoked by the acoustical events							
B	Noise onset	0-100 ms after event	-100-0 ms before event	13.33 %	1.67 %	16.11 %	3.33 %
	Noise offset	0-100 ms after event	-100-0 ms before event	13.33 %	2.77 %	10.00 %	2.78 %
	Tone onset	0-100 ms after event	-100-0 ms before event	17.22 %	4.44 %	11.11 %	2.22 %
	Tone offset	0-100 ms after event	-100-0 ms before event	12.78 %	1.67 %	17.78 %	3.33 %

Table 9. Effect of level of effort on the responses of individual units evoked by the acoustical events.

The tables (A) and (B) show the fraction of units evoked responses of which depended on level of effort.

3.5.2. Responses to the noise stimulation adjusted to the baseline level were smaller and the responses to the tone were higher in the conditions with high effort

In the previous subsection we tested whether the responses evoked by the acoustical stimuli depended on the level of effort. We wondered whether the neuronal activity during the full time of the acoustical stimuli, during noise or tone stimulations, depended on the levels of effort.

To test the neuronal activities during the acoustical stimuli, we averaged the activities during the noise (200-1600 ms after the noise onset) and tone stimulations (during 1400 ms before the water delivery). Thus, we did not include the responses evoked by the onset of the stimuli and also excluded the activities related to the detection of the tone onset. The analysis revealed that the average population activity during the noise stimulation did not differ significantly in the condition with high and low effort (Figure 46, Table 10A, a copy of the table is in Supplementary table 6C). None of the pairs showed significant differences between the average activities in the conditions. Therefore, we concluded that the average activities during the noise stimulation did not depend on the level of effort. The average activities during the tone stimulation were significantly higher in the conditions with high effort. Both pairs of the comparisons also showed significant differences between the neuronal activities and the average activities were higher in the conditions with high efforts. Thus, the average activities during the tone stimulation depended on the level of effort. The difference in the neuronal activities during the tone stimulation can be also seen in the figure that represents the average population activity (Figure 45B). Examples of the units where the activity was higher during the tone stimulation in the conditions with high effort are presented in Figure 47A-B.

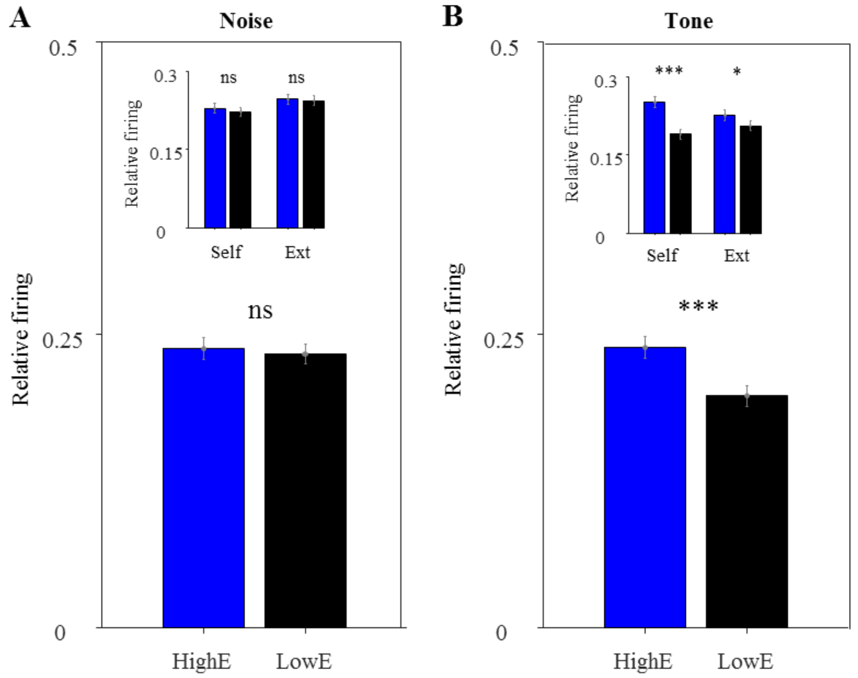

Figure 46. Comparisons of the neuronal activities during the noise (A) and tone (B) stimulation in the conditions with high and low effort.

The bars in (A) indicate average neuronal activities during the noise stimulation (200-1600 ms after the onset). The bars in (B) indicate average activities during the tone stimulation (during 1400 ms before the water delivery). The stars indicate significance levels of differences between the neuronal activates in the self- and externally-initiated conditions (Wilcoxon signed rank test, ns - $p > 0.05$, *** - $p < 0.001$). The stars in parentheses indicate that the difference between the pairs of conditions was insignificant or had different directions. The inserts in (A) and (B) show the average neuronal activities during the noise and tone stimulations respectively in each of the four conditions. The average activities in the conditions with high effort are colored in blue; the activities with the low effort are colored in black. The stars indicate the significance levels of differences between the neuronal activates in the pairs of conditions (Wilcoxon signed rank test, ns - $p > 0.05/2$, * - $p < 0.05/2$, *** - $p < 0.001$). The whiskers in the general plots and in the inserts indicate SEM of the means.

We also calculated the number of units that showed a dependency of the neuronal activities on the level of effort (Table 11A, a copy of the table is in Supplementary table 7C). Out of the 180 units, activities of the ~17 and ~17% were significantly higher and lower, respectively, during the noise stimulation in the condition with high effort. Only ~6 and ~3% had higher and lower activity, respectively, in the conditions with high effort in in all three comparisons. Individual analysis of the units during the tone stimulation revealed that ~38% of units had higher activity in conditions with high effort. Out of the 180 units, the activities of 11% of the units were significantly higher in the conditions with high efforts in all three comparisons. The activities of 8% of units were lower during the tone stimulation in the conditions with high effort. Only ~2% of units had lower activities in the conditions with high effort in all three comparisons.

Table index	Comparison	Time window		Average of High vs Low effort conditions		Pairs of High vs Low effort conditions				Conclusion
		Main	Subtracted from		P-value	Self-	P-value	Ext-	P-value	Significance
Absolute responses evoked during acoustical stimulations										
A	Noise	200-1600 ms after event	-	=	0.84	=	0.57	=	0.83	ns
	Tone	-1400-0 ms before water delivery	-	>	$4*10^{-5}$	>	$6*10^{-9}$	>	0.01	*
Change of firing evoked by the acoustical stimulation										
B	Noise	200-1600 ms after event	-1700 – -700 before trial beginning	<	$4*10^{-7}$	=	0.18	<	$7*10^{-12}$	ns
	Tone	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	>	0.03	>	$6*10^{-10}$	<	0.01	ns

Table 10. Effect of level of effort on the neuronal activities in the population during the acoustical stimulations.

Symbols in the conclusion indicate significance level of differences after the three comparisons: ns - $p > 0.05/2$, * - $p < 0.05/2$.

Table index	Comparison	Time window		High > Low		High < Low	
		Main	Subtracted from	Averaged	Averaged and pairs	Averaged	Averaged and pairs
Absolute responses evoked during acoustical stimulations							
A	Noise	200-1600 ms after event	-	16.67 %	5.56 %	17.22 %	2.78 %
	Tone	-1400-0 ms before water delivery	-	37.78 %	10.56 %	7.78 %	2.22 %
Change of firing evoked by the acoustical stimulation							
B	Noise	200-1600 ms after event	-1700 – -700 before trial beginning	10 %	0 %	30.56 %	5.56 %
	Tone	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	24.44 %	7.78 %	16.11 %	3.33 %

Table 11. Effect of level of effort on the neuronal activities of individual units during the acoustical stimulations.

The tables (A) and (B) show fraction of units neuronal activities of which depended on the level of effort.

The baseline levels differed between the four conditions (Section 3.3.7). We wanted to find whether the neuronal activities during the acoustical stimulations adjusted to the baseline levels depended on level of effort. For this, we compared the averaged *z*-scored activities during the noise stimulation (200-1600 ms after the noise onset) or the averaged *z*-scored activities during the tone stimulation (during 1400 ms before the water delivery). The resulting values for each unit were subtracted between the conditions with high and low effort. Further, the medians of the 180 values were compared with zero. The analysis revealed that the median calculated for the noise stimulation was significantly different from zero (median = -0.34, Table 10B, a copy of the table is in Supplementary table 6D). The median value was negative. That indicated that the changes of the neuronal activities in the conditions with high effort were smaller than in the conditions with low effort. The comparison of the pairs of the conditions showed that the changes of activities in the Self-HighE condition were similar as in the Self-LowE condition; the changes in the Ext-HighE were smaller than in the

Ext-LowE. Thus, the neuronal activities during the noise stimulation adjusted to the baseline level did not depend on the level of effort. For the tone stimulation, the median value was positive and differed from zero significantly (median = 0.21). The median obtained from the comparisons of the externally-initiated conditions was negative and differed significantly from zero. The median obtained from the comparisons of the self-initiated conditions was not significantly different from zero. Thus, the neuronal activities during the tone stimulation adjusted to the baseline level did not depend on the level of effort.

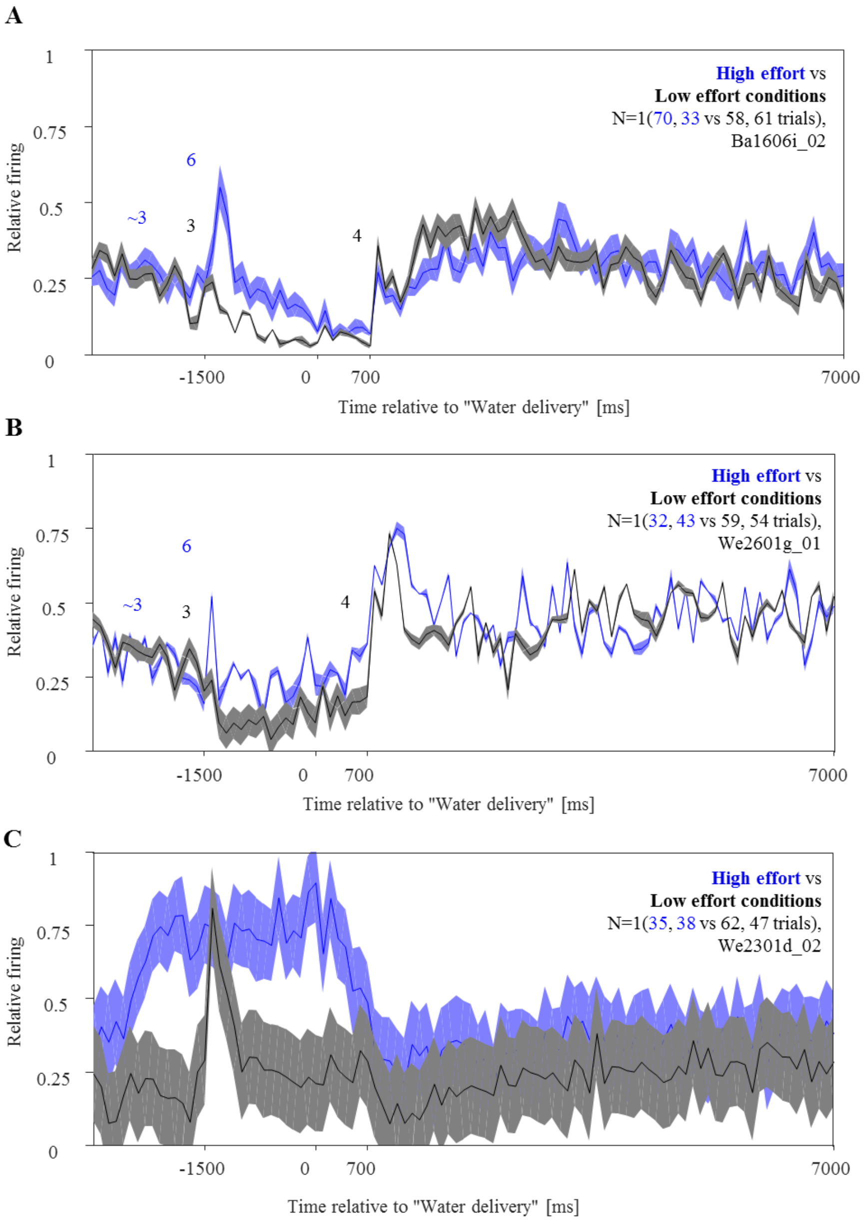

Figure 47. Examples of units where the neuronal activity was higher in the conditions with high efforts during the tone stimulation.

The figure presents PSTHs (bin size 100 ms) with SEM in the conditions with high (blue curves) and low (black curves) efforts. The numbers above the curves indicate the time of the acoustical events and movements: 3. Tone onset, 4. Tone offset, 6. Grasp/Release of the bar for detection. Note that the position of the tone onset in the conditions with high effort (~3) is only estimated because the reaction times varied in each trial.

We also analyzed in how many units of the sample the neuronal activities depended on the level of effort during the acoustical stimuli adjusted to the baseline level (Table 11B, a copy of the table is in Supplementary table 7D). We found that 10% of the units had significantly larger changes of neuronal activities and ~31% of units had significantly smaller changes of activities during the noise stimulation in the conditions with high effort. We found none and ~6% of units that had significantly larger and smaller changes of the neuronal activities in the three comparisons in the conditions with high effort. For the tone stimulation, ~24% of the units had significantly larger and ~16% had significantly smaller changes of the neuronal activities in the conditions with high effort. Around 8% and ~2% of the unit showed larger and smaller changes in activities, respectively, in the condition with high effort in the three comparisons.

3.5.3. Absence of units with a slow increase (or decrease) in the neuronal activities that depended on the level of effort

Previous studies showed that some of the units in the auditory cortex slowly increased or decreased their activity when the sound was associated with a following reinforcement (Abolafia et al. 2011, Fritz et al. 2010, Quirk et al. 1997, Shinba et al. 1995). The design of the present study allowed us to find out whether the neuronal activities increased (or decreased) slowly during the acoustical stimuli, during expectation of the water delivery and depended on the level of effort.

In order to reveal the slow increase or decrease during the noise stimulation, we used the time window 200-1600 ms after the noise onset. For the tone, we considered 1400 ms before the water delivery in order to exclude the influence of the activity changes due to the detection. We controlled whether the activity increased or decreased slowly during the time of the acoustical stimuli in the two conditions with high efforts (Pearson correlation coefficient, $r \geq 0.5$ or $r \leq -0.5$ respectively) but did not increase in the two conditions with low efforts. We controlled the presence of the slow increase (or decrease) in each unit of the population. After that, we compared the numbers of units, which satisfied the request. The presence of the increase (or decrease) in the conditions with high effort was accepted only if the units did not have the increase (or decrease) in the conditions with low effort. For additional control, we calculated percentage of units that had the increase (or decrease) in the conditions with high but not in the conditions with low effort. The dependence of the increase on the level of effort was accepted only if the percentages differed significantly.

The analysis revealed a very small amount of units (less than 5%, Table 12) which would satisfy the requirements described above and showed the slow increase (or decrease) in the conditions with high but not with low effort. Also very few units were detected as having the slow increase (or decrease) in the conditions with low but not with high effort. Thus, the number of units that showed the slow increase (or decrease) during the acoustical stimuli in the conditions with high effort was very similar to the number in the conditions with low effort. From the result, we concluded that the slow increase (or decrease) in the neuronal activities during the acoustical stimuli did not depend on the level of effort.

Change during		Noise stimulation		Tone stimulation	
Only in condition with		High Effort	Low Effort	High Effort	Low Effort
Type of change	Slow increase	0%	0.56%	0%	0.56%
	Slow decrease	0.56%	0.56%	3.89 %	4.44%

Table 12. Percentage of units where the slow increase or decrease was detected during the noise or tone stimulation.

3.5.4. Baseline levels did not depend on the level of effort

Our previous findings that were described in the present study revealed that the level of baseline varied between the four conditions (section 3.4.7). We wondered whether the baseline level depended on the level of effort.

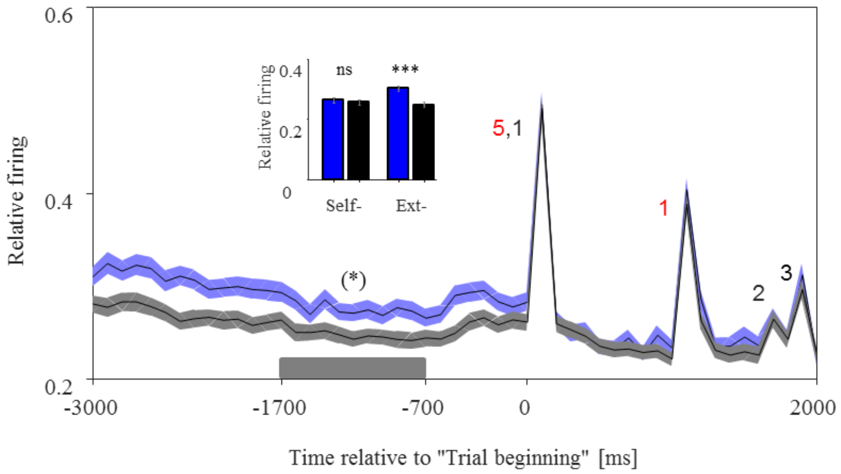

Figure 48. Average population neuronal activity in the conditions with high- and low- effort.

The subplots show PSTHs (bin size 100 ms) with SEM in the conditions with high (blue shadow) and low (black shadow) effort. The time of the PSTHs is related to the beginning of the trial that was the bar grasp in the self-initiated conditions and the noise onset in the externally-initiated conditions. The horizontal gray bar indicates the time interval where the baseline levels were compared. The stars in the parentheses indicate the significance level of differences between the neuronal activities in the marked time window in the conditions with high and low effort (Wilcoxon signed rank test, * - $p < 0.05$). The stars are in brackets because the comparison of one pair of the conditions did not reveal significant results. The numbers above the curves indicate the acoustical events and movements: 1. Noise onset, 2. Noise offset, 3. Tone onset, 5. Grasp of the bar for the self-initiation. The red and black numbers indicate the events in the self-and externally-initiated conditions respectively. The insert shows the average neuronal activities in the marked time window. The blue bars indicate average activities in the conditions with high effort; the black bars indicate the conditions with the low effort. The stars indicate the significance levels of differences (Wilcoxon signed rank test, ns - $p > 0.05/2$, *** - $p < 0.001$). The whiskers indicate the standard deviations of the means.

To address this question, we compared the pair of conditions with the high effort with the pair of the conditions with the low effort. The first event of the self-initiated condition was the bar grasp and the first event of the externally-initiated conditions was the onset of the noise. We showed previously (section 3.3.6) that the neuronal activity responded ~700 ms before the bar grasp. Therefore, we compared the neuronal activities between each pairs of conditions with high and low effort 1700-700 ms before the beginning of the trial (Figure 48). We found that the baseline level in the conditions with high effort was significantly higher than in the conditions with low effort (Wilcoxon signed rank test, p=0.02). However, the pairwise comparison of the conditions revealed that the baseline level in the Self-HighE condition did not differ significantly from the Self-LowE conditions (0.55). The baseline level in the Ext-HighE condition was significantly higher than in the Ext-LowE condition ($p=1*10^{-7}$). We concluded that the level of baseline did not depend on the level of effort (Supplementary table 6E).

Comparisons of the baseline levels and their dependency on the level of effort was also controlled in the 180 individual units (Supplementary table 7E). The analysis showed that ~26% of the units had higher baseline level in the two conditions with high effort (Wilcoxon signed rank test, $p < 0.05$). Only ~7% of the units had higher activity in the three comparisons. In the population, ~7% of units had lower baseline level in the conditions with high effort. Only 2% of the 180 units had lower activity in the three comparisons.

3.5.5. Increase in the neuronal activity after detection correlated with the slow increase in activity after the water delivery

The results described previously (section 3.5.2) showed that the neuronal activities during the tone stimulation were higher in the conditions with high effort. It is worth pointing that the differences in the activities could be due to the detection of the tone onset in the conditions with high effort. Therefore, we controlled whether the same units that had the response after the bar grasp or release also had the higher activity during the tone stimulation.

In order to control the correlation between the responses to the movement and the increase in the activities during the tone stimulation, we found all units that had the responses evoked by the bar grasp or release. We found 59 units with responses to the bar grasp or release (Figure 49A). Further, we analyzed whether the activity during the tone stimulation was significantly different between the conditions with high and low effort. The analysis revealed that the neuronal activity during the tone stimulation was significantly higher in the conditions with high effort (Wilcoxon signed rank test, $p=9*10^{-9}$). Comparisons of the pairs of conditions also showed that the activity was significantly higher in the conditions with high effort (Wilcoxon signed rank test, $p=9*10^{-8}$ for the self-initiated and $8*10^{-4}$ for the externally-initiated conditions). After, we tested the differences between the conditions with high and low effort in the group that did not show the responses evoked by the bar grasp or release. For this, we combined the remaining 121 units (Figure 49B). Similarly as for the group where the responses to the detection were found, we analyzed the activities during the tone stimulation. The activities during the tone stimulation did not differ significantly in the conditions with high and low effort (Wilcoxon signed rank test, p=0.12). The comparison of the pairs of the conditions revealed that the neuronal activities were higher in the self- but similar in externally- initiated conditions with high effort (p=0.002 and p=0.54 respectively). Thus, the group of the units where the responses to the bar grasp or release were not detected did not show differences in the activities during the tone stimulation between the conditions. The result indicates that the differences during the tone stimulation were related to the response to the movements. Two examples presented in Figure 47 represent the two types of units where the activity in the conditions with high effort was higher. All the units had the responses to the bar grasp or release.

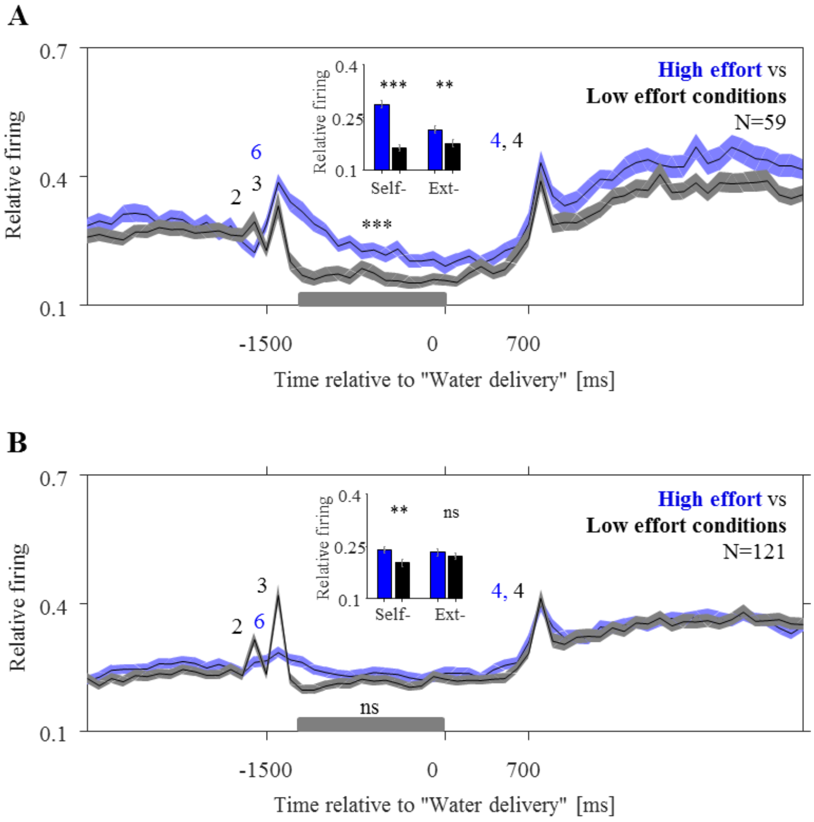

Figure 49. Average neuronal population activity of the units that had (A) and did not have (B) response to the bar grasp or release.

The subplots show PSTHs (bin size 100 ms) with SEM in the conditions with high (blue shadow) and low (black shadow) effort. Subplot (A) represents the group of units where the responses to the bar grasp or release was significant (comparison of the activity during 100ms before and after the bar grasp, Wilcoxon signed rank test, $p < 0.05$). Subplot (B) represents the group of units where the responses to the bar grasp was insignificant (Wilcoxon signed rank test, $p > 0.05$). The horizontal gray bar indicates the time window in which the neuronal activities were compared. The stars indicate the significance level (Wilcoxon signed rank test, ns - $p > 0.05$, *** - $p < 0.001$). The numbers indicate the acoustical events and movements: 2. Noise offset, 3. Tone onset, 4. Tone offset, 6. Grasp/Release of the bar for detection. The blue and black numbers indicate the events of the conditions with high and low effort respectively. The insert shows the averages of the neuronal activities during the tone stimulation in each of the four conditions. The stars indicate the significance levels (Wilcoxon signed rank test, ns - $p > 0.05/2$, ** - $p < 0.01/2$, *** - $p < 0.001/2$). The whiskers show the standard deviations of the means.

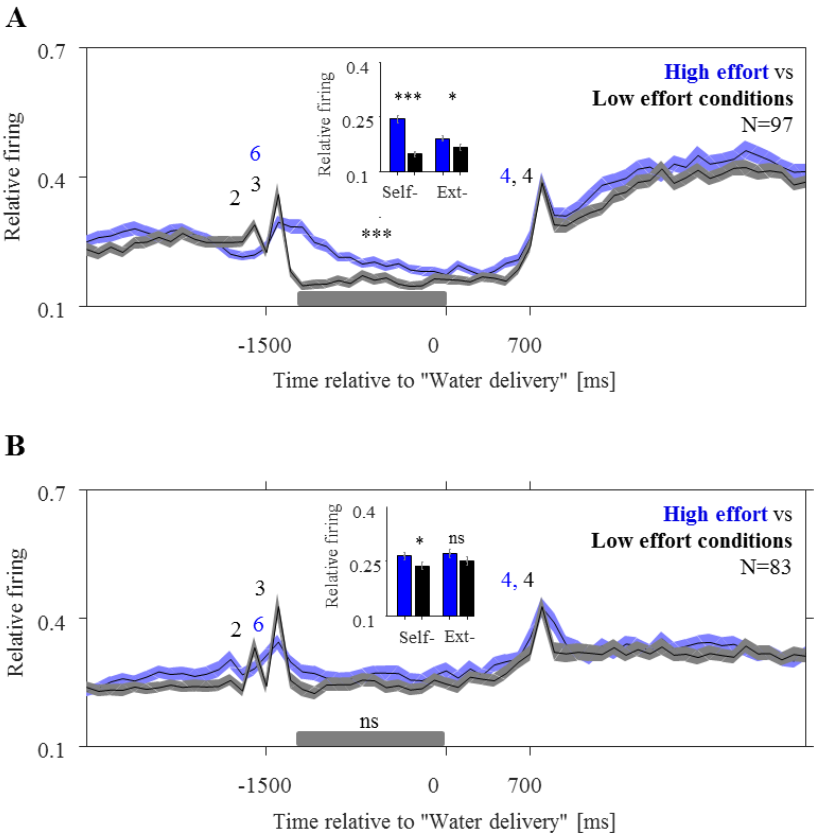

Figure 50. Average neuronal population activity in the units that had (A) and did not have (B) significant increase in activity after the water delivery.

The subplots show the average PSTHs (bin size 100 ms) with SEM in the conditions with high (blue line) and low (black line) effort. Subplot (A) represents the group of units where the slow increase in activity after the water delivery was significant (2500 ms after the water delivery, Spearman's correlation coefficient, $r > 0$, t-test, $p < 0.05$). Subplot (B) represents the group of the units where the slow increase in activity after the water delivery was not significant. The horizontal gray bar indicates the time window in which the activities were compared. The stars indicate the significance level (Wilcoxon signed rank test, ns – $p > 0.05$, *** - $p < 0.001$). The stars in the parentheses indicate that the significant level was reached only between the averaged conditions with high and low effort but not between the pairs of the self- and externally-initiated conditions. The numbers above the curves indicate the acoustical events and movements: 2. Noise offset, 3. Tone onset, 4. Tone offset, 6. Grasp/Release of the bar for detection. The blue and black numbers indicate the events of the conditions with high and low effort respectively. The inserts show the averages of the neuronal activity during the tone stimulation. The whiskers show the standard deviations of the means. The stars indicate significance level of the differences (Wilcoxon signed rank test, ns - $> 0.05/2$, * - $p < 0.05/2$, *** - $p < 0.001/2$).

In the previous sections we reported that more than half of the population had a slow increase in activity after the water delivery independently of the types of the conditions (section 3.3.4). We wondered whether the same units that had the higher neuronal activity during the tone stimulation in the conditions with high effort had the slow increase in activity after the water delivery. To address this question, we grouped all units that had the slow increase in activity after the water delivery (Spearman correlation coefficient, $r > 0$, $p < 0.05$). In the detected 97 units (Figure 50A) and in the remaining 83 units (Figure 50B), we calculated the differences between the activities during the tone stimulation in the conditions with high and low effort. The analysis revealed that between the units in the group with the slow increase in activity after the water delivery, the neuronal activity was higher during the tone stimulation in the conditions with high effort (Wilcoxon signed rank test, $p=4*10^{-8}$). Both pairs of the conditions, the self- and externally-initiated also had higher activity in the conditions with high effort ($p=5*10^{-9}$ and p=0.02). Thus, the group of the units with the slow increase in activity after the water delivery also had the higher neuronal activity during the tone stimulation in the conditions with high effort. In the second group, without the increase in activity after the water delivery, the activity during the tone stimulation was similar between the conditions with high and low effort (Wilcoxon signed rank test, p=0.20). Only one pair of the externally-initiated conditions had a significantly higher activity during the tone stimulation in the condition with high effort (Wilcoxon signed rank test, p=0.02) when the pair of the self-initiated conditions did not reveal significant differences between the conditions with high and low efforts (p=0.26). Thus, the units without the slow increase in activity after the water delivery did not show significantly higher activity during the tone stimulation in the conditions with high effort. Two examples of the units where the slow increase after the water delivery and differences between neuronal activities in the conditions with high effort were detected were presented in the Figure 47. From the results obtained from the group division, we concluded that the higher activity during the tone stimulation was correlated with the slow increase in activity after the water delivery.

3.5.6. Increase in the frequency of the mouth movements after detection of the tone onset

Previously (section 3.3.3), we showed some example sessions of the frequencies of the mouth movements. We found that in the most of the recorded sessions, the frequency of the mouth movements increased slowly after the water delivery and reached maximum in ~2500 ms. We also found some sessions where the frequency of the mouth movements increased at the time when the monkey detected the tone onset. The increase was independent of the type of the detection, grasp or release of the bar. We wondered whether the units which were recorded during the sessions with the increase in the frequency of the mouth movements after detection also had the increase in the neuronal activity at this time.

For this purpose, we analyzed the neuronal activities of all units that were recorded during the session where the increase was detected. During the four sessions with the increase, 14 units were recorded (1, 3, 5 and 5 units in each of the four sessions). Between the units, we observed 10 units (0, 1, 2 and 5 units in each session respectively) that had a higher neuronal activity after detection, during the tone stimulation. One example of the frequency of the mouth movements and one example of the unit recorded during the session is presented in Figure 51. In the figure, the frequency of the mouth movements were averaged between the two conditions with high and two with low effort in order to stress the increase in the time after the detection of the tone onset (Figure 51A). The example units showed the response evoked by the movements and also higher neuronal activity during the tone stimulation (Figure 51B). Thus, the presence of such decrease in the neuronal activity of only some units in one session indicates that the presence of the increase in the frequency of the mouth movements did not obligate the increase in the neuronal activity. However, the frequency of the mouth movements, or some accompanying processes of the water collection, might be reflected in the neuronal activity of some units.

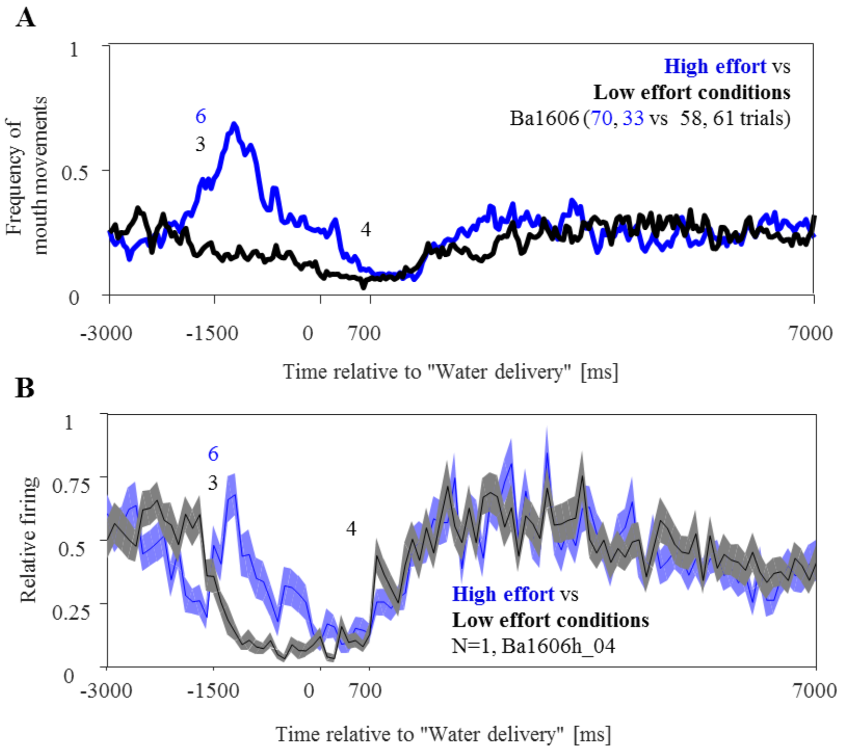

Figure 51. Frequency of the mouth movements (A) and one examples of unit (B) that was recorded during the same session.

Subplot (A) represents frequency of the mouth movements during one session; subplot (B) represents PSTHs (bin size 100 ms) with SEM of unit recorded during this session. Colors of the curves indicate the levels of effort in the conditions: high (blue) and low (black). Subplot (B) shows a unit that had the increase in activity after detection, grasp or release of the bar, and higher activity during the tone stimulation. The numbers indicate the time of the acoustical events and movements: 3. Tone onset, 4. Tone offset, 6. Gasp/Release of the bar for tone detection. The blue and black numbers indicate events in the conditions with high and low effort respectively.

3.5.7. Level of increase after water delivery did not depend on the level of effort

Lastly, we wanted to find whether the increase in the neuronal activity after the water delivery depended on the level of effort. To address the question, we compared the average neuronal activities during the period of the maximum increase (2000-3000 ms after the reward delivery). Using the Wilcoxon signed rank test, we found that the average activity in the time window did not differ significantly between the conditions with high and low effort (p=0.18). The pair of the self- initiated conditions did not differ significantly (p=0.36), when the pair of the externally-initiated conditions had a higher activity in the condition with high effort ($p=5*10^{-3}$). Thus, the increase after the water delivery did not depend on the level of effort (Supplementary table 6F).

We also controlled in how many units the increase after the water delivery depended on the level of effort (Supplementary table 7F). A comparison of the neuronal activities in the individual units showed that ~22% of them had higher activity in the conditions with high effort. Between 180 units, around ~7% of units showed higher averaged neuronal activities in the two conditions with high effort and in both conditions with high efforts. In the population, ~10% of units showed that activity was significantly higher in the time window in the conditions with low effort. Only ~1% of 180 units showed significantly higher activity in the conditions with the low effort in the three comparisons.

4. Discussion

4.1. Summary of the results

The aim of the study was to find the influences of different forms of engagement including the meaning of a sound, the presence of unconditioned stimuli alone, the effect of the sense of agency and the level of effort on the neuronal activity in the primary auditory cortex.

In order to answer these complex questions, we first analyzed how the individual factors, such as the unconditioned stimuli alone and the acoustical stimuli without meaning, changed the activity in the primary auditory cortex of the low- and high-trained monkeys. We compared the effects of the individual factors with their pairing, i.e., with Pavlovian conditioning (sections 3.1 and 3.2). We found that the presence of the unconditioned stimuli led to a slow increase after the water delivery and to an upward shift of the baseline level in the groups of the low- and well-trained monkeys. We also found that the acoustical stimuli without meaning led to higher responses to the noise onset in the group of the well-trained monkeys.

Considering the effect of the individual factors, we analyzed the neuronal activities in the well-trained monkeys between the instrumental and passive condition (section 3.3). We found that the baselines of the neuronal courses in the three instrumental and the one passive condition were generally similar. The neuronal activities differed in short intervals of the courses due to the variations in the required movements in the conditions.

Lastly, we compared the effect of the sense of agency and levels of effort (sections 3.4 and 3.5). The main effects of the sense of agency were the decrease in the neuronal activity after the self-initiation and lower response during the noise stimulation. The main effect of the effort was the higher neuronal activity after detection (during the tone stimulation) in the conditions with high effort.

4.2. The effect of the sound meaning and of the unconditioned stimuli alone

4.2.1. Slow modulations related to the presence of unconditioned stimuli

We found surprisingly that such a small number of studies seek to control the effect of the unconditioned stimuli in the neuronal activity of the sensory cortexes. The only one attempt to find the effect in the auditory cortex was made by Ide and colleagues (2012). They revealed a tonic response in the neuronal activity related to the presence of aversive stimuli but only after animals have previously learned the association between the acoustical and aversive stimuli. Another study made by Rowland and colleagues (1985) showed the effect of pleasant unconditioned stimuli in the visual cortex. The pleasant stimuli were the electrical stimulation of the rat’s forebrain bundle. The study reported a slow increase in the neuronal activities before the pleasant stimuli in the animals that also learned to associate them with visual signals. Thus, we are the first who report the effect of the positive unconditioned stimuli alone on the neuronal activity in the primary sensory cortex of animals.

Before the experiment, we supposed to find some changes in the background activities that were related to the presence of the unconditioned stimuli due to the rich connections with the primary auditory cortex (Scheich et al. 2011). Indeed, we found a shift of the baseline in the neuronal activity in all conditions with unconditioned stimuli, in both groups of monkeys with and without previous experience. We suppose that the slow changes in activities in the conditions with water delivery could be explained by responses of the units to the mouth movements that were made by the monkeys in order to collect the liquid drops (frequency of the mouth movements). As we described in the results, the frequency of the mouth movements were similar to the neuronal activities of some units recorded in the low- and in the well-trained monkeys. Probably, the units that showed the correlations were multisensory units of the primary auditory cortex that were highly sensitive to the somatosensory signals coming from the mouth. Previous studies that were made by our colleagues and many studies of the last decades showed that somatosensory perceptions led to responses in the neuronal activities of the primary (Brosch et al. 2005, Lakatos et al. 2007) and secondary auditory cortexes (Fu et al. 2003, Schroeder et al. 2001, Stein and Meredith 1993, Vaadia et al. 1982). Indeed, the mouth movement of the water restricted monkeys during the collection of liquid drops can be considered as important somatosensory (and motor) events and, certainly, were very familiar for the adult animals. Because the events of the mouth movements occurred many times per each trial and at different moments from trial to trial, the averaging between the trials will rather lead to smooth changes in the neuronal activity than to some phasic changes. Note that the courses of the mouth movements in the three conditions were very similar to the average neuronal activities of the population in the three conditions in the low- and in the well-trained monkeys. These similarities support our hypothesis.

One could assume that the slow modulations and the shift of the baseline in the neuronal activities correlated with mastication responses in the auditory cortex. Indeed, a few units, of whose activities correlated with mastication, were found in the auditory system (Hage and Jürgens 2006). But none of the studies were made in the primary auditory cortex. Although we cannot clarify certainly whether the slow changes were or were not related to the mastication. We can assume only that the mastication responses led to an increase in somatosensory inputs in the brain.

Considering the fact that the recordings were made in the auditory cortex, the appearance of the slow modulations in the neuronal activities might be explained by the sounds of the licking, swallowing and others movements produced by the monkeys. Some studies showed that the self-produced sounds led rather to inhibition in the neuronal activity in the auditory cortex (Creutzfeldt et al. 1989, Eliades and Wang 2003, 2005). Based on these studies, we assume to have rather a downward shift of the baseline level but the results showed an upward shift. Also, the present study of the sense of agency and the effort, which was presented in sections 3.4-3.5 and will be discussed in more details in sections 4.4-4.5, showed some contradictions of the effect described by other studies (Creutzfeldt et al. 1989, Eliades and Wang 2003, 2005).

Some studies reported slow modulations in the neuronal activities related to the expectations of the reinforcement in the animals trained to some instrumental conditions. One of the studies was made by Komura and colleagues (2001) in the rat's auditory thalamus. Rats had to lick a spout after an acoustical stimulus to obtain a drop of water. During the expectation, the neuronal activity in the thalamus had a slow increase after the onset of the acoustical stimulus and continuing before the water delivery. These increase disappeared when the water delivery was cancelled. The authors interpreted the increase due to the expectation of the reinforcement. It is interesting that the increase in the neuronal activity disappeared together with the licking of the spout. Thus, probably, the increase in the neuronal activity in the thalamus was also related to the motor/somatosensory inputs as in the present study. Other studies also obtained the slow modulations related to the expectation (timing) of the reinforcement but in the rats' primary visual cortex (Chubykin et al. 2013, Huertas et al. 2015, Namboodiri et al. 2015, Shuler and Bear 2006). After visual stimuli, rats had to lick a spout a curtain number of times in order to get a water delivery. During the trials, three types of the changes in the neuronal activity were described: the sustained increase and sustained decrease before the water delivery and (prolonged) peak after the water delivery. Thus, the authors interpreted that the slow changes in the neuronal activities were related to the timing of the reinforcement. We also seek for the three types of the changes in the neuronal activities but observed only a mix of the second and the third types. In our study, the slow increase after the water delivery could be also described as a decrease before the water delivery smoothly passing into the (prolonged) peak. Probably, the slow changes in the neuronal activities described in the rats' visual cortex were more related to the somatosensory input during the required licking before and after the water delivery rather than to the expectation (timing) of the reinforcement. It is difficult to say for sure because the mouth movements and the expectations of the water delivery highly dependent on each other. Also, it is important to note that the recordings in the auditory thalamus (Komura et al. 2001) and the primary visual cortex (Shuler and Bear 2006 and later studies of the same group) were made during animals performed instrumental conditions. Based on these, slow changes in the neuronal activities were interpreted to be related to the expectation. In our study, the slow modulations were obtained even during the passive conditions. Therefore, we showed that the slow modulations in the neuronal activity appeared independently of learning of the instrumental conditions and, therefore, independently of the expectation (timing) but only due to the presence of water. Moreover, we controlled the effect of expectation and did not observe any changes related the process.

Previous studies of our lab have also revealed slow modulations, also related to the prediction error of the water delivery (Brosch et al. 2011b, 2015). The peaks after the water delivery described in the study of Brosch and colleagues were not similar to the slow modulations observed in the present study. We suppose that the reason of the differences was due to the differences in design and requirements to the animals between the two studies (Fanselow and Wassum 2016). The present design embraced variety of the requirements that depended on the conditions. Between the six conditions that the well-trained monkeys were presented with, the requirements varied between a bar grasp and waiting, a bar release and waiting, a bar holding and waiting and just waiting. In the studies of Brosch et al. (2011b, 2015), the monkeys had only one requirement for all conditions, i.e., a bar release. Thus, the differences in the requirements to the animals led to differences in general strategies of the animals in the two studies. In turn, the different strategies might explain the differences in the neuronal activities (Gorkin 2017, Gorkin and Shevchenko 1995).

The slow changes in the neuronal activities in the well-trained monkeys varied between the two conditions with the unconditioned stimuli. In the conditions where the water delivery was predicted by the acoustical stimuli, the slow increase after the water delivery smoothly decreased to the beginning the next trial and was minimal during the tone stimulation. We found that the frequency of the mouth movement repeated the courses of the neuronal activities. Thus, the expectations of the well-trained monkeys to get the water varied during the trial and were the lowest during the tone stimulations. In contrast, the slow changes in the neuronal activities and frequencies of the mouth movements in the conditions where the water delivery was predicted by an acoustical stimulation was rather constant during the trials in the group of the low-trained monkeys. In this group, we also found very poor evidence of the association between the acoustical stimuli and water delivery. The constant expectation of the water delivery and the high neuronal activity throughout the entire trial that were observed in the low-trained monkeys took, probably, a lot of energy from the animals (Deitmer 2001, Fanselow and Wassum 2016, Siegel et al. 1999). Thus, we can interpret the leveling of the neuronal activity and the decrease in the frequency of the mouth movements at the time of the acoustical stimulation as optimization of behavior and energy saving of the well-trained monkeys.

4.2.2. Neuronal responses to sounds with different meaning

In the present study we also controlled the effects of the sound meaning in the neuronal activities of the barely and highly experienced monkeys. We revealed the effect in the neuronal population of the group of monkeys with high experience in which the responses evoked to the sounds with meaning were lower than the responses to the sounds without meaning. We observed similar effects in only some units in the population of units in the monkeys with few experience. We also had only some sessions in which the monkeys showed their association of the sounds with the reinforcements, therefore, there was very poor evidence of their understanding of the meaning.

Neuronal recordings in the primary auditory cortex during pairing of the conditioned (acoustical) stimuli and unconditioned stimuli, Pavlovian conditioning, were also made by some other research groups previously. They showed that some neurons fired stronger and some weaker to the acoustical stimuli after learning of their meaning than to the same acoustical stimuli without meaning (Diamond and Weinberger 1989, 1986, Ohl and Scheich 1996, 1997, Weinberger et al. 1984, Weinberger and Diamond 1987). Thus, not only the higher but also the lower response to the acoustical stimuli with meaning was considered as a possible effect. Suga and Ma (2003) explained the lower response with a model in which the tonotopical map changed. They described that when a sound was meaningful, responses of a neuron were higher to the central frequency but lower to all neighboring frequencies. However, the model did not explain how the changes in the tonotopical map can vary as fast as during one session as it was reported in the present study.

Similarly to the present study, some researchers described lower neuronal responses in the auditory cortex to the acoustical stimuli with meaning while animals performed instrumental conditions compared with the same acoustical stimuli presented to the passively listening animals (Lee and Midlebrooks 2011, Otazu et al. 2009). Similar effects was observed in the olfactory (Fontanini and Katz 2006) and visual cortexes (Shuler and Bear 2006) to the meaningful stimuli with corresponding modalities. In contrast, many studies showed that the responses to the acoustical stimuli in the auditory (Atiani et al. 2014, Hubel et al. 1959, Shinba et al. 1995) and other sensory cortexes (Maunsell and Cook 2002) increased with meaning. The study made by Abolafia and colleagues (2011) described both effects that were found in different units, the higher and lower responses to the meaningful acoustical stimuli.

We were surprised to find that much more units in the population had lower responses to the acoustical stimuli with meaning. Higher responses to acoustical stimuli with meaning would be much more intuitive. However, it is still unknown what the way of brain is to indicate the meaning of sounds or some other stimuli. After revealing both, the higher and lower responses to the stimuli with meaning, some researchers rather focused their attention on the higher responses to the sounds with meaning (for review read Ohl and Scheich 2005). Nevertheless, it may turn out that the lower responses evoked by a stimulus is a marker of meaning for the brain. For instance, the results of our study showed only the part of the units that responded significantly lower to the sound with meaning. There were also a lot of units that had no differences between the responses to these sounds. Thus, the comparison of the differences in one subgroup of units with no change in another subgroup could be a key factor for indicating the meaning of the acoustical stimuli. Thus, probably, the differences of changes between subpopulations of units indicate importance of acoustical stimuli (Fishman et al. 2001, 2012, 2017, Fritz et al. 2007a, 2007b, Knyazeva et al. 2018, Shamma et al. 2011). An additional bonus to this way of functioning would be the energy saving for an organism (compared to the situation where an important acoustical stimulus leads to even higher responses). A similar reasoning of the lower responses to the acoustical stimuli with higher meaning was also given by other researchers (Ghose 2004, Ghose et al. 2002).

Some previous publications of colleagues (Ohl and Scheich 2005, Weinberger 2004, Weinberger 2007) revealed discrepancies in the results and interpretations between the studies. When the studies of one research group reported the prevalence and importance of the higher responses related to the sound meaning, another research group reported the prevalence of the lower responses related to the sound meaning and importance of taking into account both changes. The present study also showed two types of the changes (Supplementary table 2), higher and lower responses depending on the compared conditions and on the interest time windows. Curiously, the lower increases of the responses to the acoustical stimuli was rather related to their meaning; when higher responses were observed to the acoustical stimuli in the conditions with presence of unconditioned stimuli in the neurons of the monkeys with few experiences. The effect of the presence of the unconditioned stimuli was not considered in the previous studies of the colleagues (Ohl and Scheich 2005, Weinberger 2004, Weinberger 2007) and could be a possible reason of the discrepancy in the results and interpretations.

A group of studies (Atiani et al. 2014, David et al. 2012, Fritz et al. 2005, 2007c, 2010, Lu et al. 2016, Yin et al. 2014) which recorded spectrotemporal receptive fields showed that the acoustical stimuli that were matched to the best frequency of the unit had a higher variability and depended on the meaning. As follows, the acoustical stimuli that activate the neighborhood of the best frequency were less variable. The recordings of the present study were made in the primary auditory cortex, units of which respond higher to the pure tones (Rauschecker and Tian 2004 but see also Kajikawa et al. 2011). However, we found that the responses to the noise onset, in comparison to the responses to the pure tone onset, were higher and were the most sensitive to the meaning. That occurred despite to our intention to adjust the frequency of the pure tones to the best frequency of the neurons. The explanation of the higher response to the noise in the population level might be due to the fact that we could never reach an ideal situation when all the recorded units of a session were fully adjusted to the best frequencies. Due to the physical properties of the noise, the combination of the large range of the frequencies, the probability of the responses to a noise might be higher than to a pure tone in the primary auditory cortex (Moller 2006). Simultaneously with the higher probability of the response, we supposed that one of the frequencies in the range of the noise was closer to the best frequency of a neuron than the best frequency of the pure tone. Due to the properties of the noise and features of the adjustment to the best frequencies, we could expect higher variation of the response to the noise than to the pure tone depending on the meaning. Also, we cannot exclude that the first acoustical stimuli in the row, the noises in the type “tone followed by noise”, was more important for the animals than the second acoustical stimuli, the pure tones, despite the pure tone was synchronized with the water delivery. Another possible reason of higher response to the noise onset, instead of the tone onset or offsets, is the second-order conditioning (Fanselow and Wassum 2016) that can be also described as a direct association (Sechenov 1863) between the noise and reinforcement of the well-trained monkeys in the present design. Lastly, we propose another explanation. Probably, the reason of the higher sensitivity of the units to the noises with different meaning was persistence of the noise and change of the pure tone from session to session (Denton and Kruschke 2006, Fanselow and Wassum 2016, Van Hamme and Wasserman 1994, Le Pelley et al. 2016). Thus, the noise was more reliable stimuli predicting the water delivery because it was present in each session compared with the pure tones that varied between all sessions.

An exciting topic whether the positive and negative reinforcement lead to the activation of different structures in the brain (Fanselow and Wassum 2015, Fritz et al. 2005, Ilango et al. 2010, Kluge et al. 2011, Martin-Soelch et al. 2007, Morison and Salzman 2009, Pi et al. 2013, Yin et al. 2014) was never regarded as an effect that might lead to different effects in the sensory cortexes. More than fifty years ago it was already supposed that the type of the reinforcement should be differentiated in some ways in the brain (Konorski 1967). In the studies of some groups, the reinforcement was positive, such as the delivery of sucrose or a drop of water. The responses to the acoustical stimuli with meaning were lower than to the same stimuli without meaning (David et al. 2012, Fontanini and Katz 2006, Lee and Middlebrooks 2011, Otazu et al. 2009, Shuler and Bear 2006). In contrast, in the studies of the Shamma’s group (Atiani et al. 2014, David et al. 2012, Fritz et al. 2005 and 2010, Yin et a. 2014), who used an analogue of the “go/no go” paradigm, the punishment for the wrong performed “no go” condition in the experiment was usually negative, a tail shock. The researcher devoted the readers' attention to the higher responses to the acoustical stimuli with meaning. However, some of their manuscripts showed but did focus the readers' attention on the differences between the responses to the "go" acoustical stimuli, which allowed drinking water unhindered, and the same acoustical stimuli during passive listening. Responses to the "go" stimuli were lower compare with the responses to the same stimuli during passive stimulation. We also found one study that disagreed with our assumption that the positive reinforcement lead to the lower responses when the negative leads to the higher responses. Thus, Abolafia and colleagues (2011) showed that the responses to the acoustical stimuli with meaning had higher responses compared with the response to the same stimuli without meaning. Thus, we suppose that the effect of the type of reinforcement must be controlled in further studies.

Beside the responses to the onsets/offsets of the acoustical stimuli, some studies reported differences between the responses during the full time of the acoustical stimulations (Atiani et al. 2014, David et al. 2012, Fritz et al. 2010, Yin et al. 2014). In their studies, ferrets were trained to stop licking at a target tone to prevent a tail shock. The same tones were presented before and after the training session. The researchers found that the responses during the tone stimulations were higher during the instrumental conditions compared with the passive listening. Some similar results were described in the auditory (Abolafia et al. 2011, Shinba et al. 1995), gustatory (Fontanini and Katz 2006) and somatosensory cortexes (Pantoja et al. 2007).

In the present study, we showed that both groups of monkeys had lower average neuronal activities during the noise and tone stimulations in the conditions with presence of unconditioned stimuli. Because the group of the monkeys with little experience showed rather absence of the association between the acoustical stimuli and reinforcement, we assumed that the differences in the neuronal activities between the conditions were not related to the sound meaning. Probably, the reason of these differences was more related to the slow modulations in the neuronal activities in the conditions with presence of reinforcement.

4.2.3. Omitted analyses, limitations of the present study and further interesting questions regarding to the presence of unconditioned stimuli and meaning of sound

One of the omitted topics of the present study is the changes in the neuronal activities and in the behavior that could be observed during some time after the switches from one condition to another during one session, i.e., the extinction and acquisition as it was shown by Quirk and colleagues (1997), Rowland and colleagues (1985) and others. We deliberately skipped the analysis because we supposed that our design is unsuitable for the analysis. We believe that a more appropriate design should be implemented with a much higher number of trials during one condition, with only one type of acoustical stimuli. Ideally, the chronicle recording of the same units should be performed from the first to the last sessions (for instance, as it was made in the studies of Bondar' et al. 2014 and Lovell et al. 2014). Additionally to the larger number of sessions, one could conduct all video recordings of the mouth movements and, probably, size of the pupils (Kahneman 1973, Kahneman and Beatty 1966) simultaneously with the neuronal recordings. Without all of these requirements, we could not take responsibility to interpret the results.

Unexpected water delivery is a very important and pleasant event for a thirsty animal that might change activity in the ventral tegmental area (VTA, Hollerman and Schultz 1998, Kim et al. 2016, Puschman et al. 2013, Rowland et al. 1985, Schultz 2002), substantia nigra (Ljunberg et al. 1991, 1992, Romo and Schultz 1990, Scheich et al. 2011, Schultz 1998, Tobler et al. 2003, Waelti et al. 2001), nucleus accumbence (Bowman et al. 1996, Cameron and Carelli 2012, Cartoni et al. 2016, Chernyshev and Weinberger 1998, Tremblay et al. 1998), orbitofrontal cortex (Morrison and Salzman 2009, Rolls 2004, Rolls and Deco 2016, Rolls and Grabenhorst 2008) and other structures (Bouret and Richmond 2009, Fanselow and Wassum 2015, Rolls 2016). Previous studies revealed a correlation between the dopamine neurotransmitters and the neuronal activities in the primary auditory cortex (Bao et al. 2001, Huang et al. 2016b, Lou et al. 2014, Weis et al. 2012). One study reported an absence of any change in the neuronal activity during acoustical stimulation with and without VTA stimulation (Bao et al. 2001); other two studies revealed lower responses evoked by the acoustical stimuli (Lou et al. 2014) and lower spontaneous activity (Huang et al. 2016b) in the auditory cortex when the VTA was stimulated. Therefore, the activation of these structures and the presence of changes in the neuronal activity of the auditory cortex might be highly related. There are a lot of question that we ask ourselves in order to understand the mechanisms and interconnections between the reinforcement-related structures and auditory cortex. Thus, for instance, a very important step to continue the study would be to make parallel recordings in the neuronal activities in a dopamine structure and in the auditory cortex.

A further question that could enrich the described topics is whether the neuronal and behavioral activities will show different results when each new trial is reinforced with a lower probability than 100%. The reason of the interest relates to possible overlapping of the neuronal activity between the two neighboring reinforced trials. Thus, we cannot be sure of whether the increase after the water delivery was related to the present trial or rather with expectations of the next trial.

We suppose that other primary cortexes also have the rich connections with subcortexes and with other cortexes of the brain as the primary auditory cortex. We revealed that the slow modulations related to the presence of the unconditioned stimuli occurred not only during the pairing of the acoustical stimuli with water delivery but also during the regular delivery of water; occurred not only in the well-trained monkeys but even in the low-trained monkeys. Due to these facts, we assumed that similar modulations will arise in any of the sensory cortexes when water, or some unconditioned stimuli, are present. However, if this assumption is correct, the upward shift in the baseline level in all sensory cortexes of the brain will take a lot of energy from the organism. Therefore, a question arose whether the units in the primary visual cortex, for instance, will be also very active during the combination of the acoustical stimuli with water delivery.

4.3. Motor/somatosensory responses were the main reason of the differences between the four conditions

4.3.1. Toward differences in the neuronal activities between Pavlovian and instrumental conditionings

In the present study we recorded the neuronal activities in conditions, in which the well-trained monkeys were presented with repeated acoustical stimuli paired with the water deliveries, which could be in principal named Pavlovian conditioning. However, Pavlovian conditioning implicates a condition in which a subject is presented a conditioned stimulus paired with reinforcement (Pavlov 1927, Sprague and Epstein 1983) and, what is very important for differentiation from the instrumental condition (Bouton 2007, Konorski J. 1948, 1967, Skinner 1938, Thorndike 1901), excludes the presence of any special voluntary behavior in order to get/avoid the reinforcement. Because the well-trained monkeys were highly trained to get the reinforcement in the instrumental conditions in the present study and because the same acoustical stimuli were presented during the pairing with the unconditioned stimuli, we expected to find and also revealed that the monkeys actively inhibited their motor behavior in the condition (see section 3.2.1). The inhibition of the motor behavior might be considered as a special behavior that had to be learned. However, the specialization of behavior formed to obtain a reward is more proper for instrumental conditions. Therefore, we accepted the combination of the acoustical stimuli paired with water delivery as a “passive condition”, as a condition located at the junction between Pavlovian and instrumental.

The features of the neuronal activity observed during the pairing of the acoustical stimuli with reinforcement in the well-trained monkeys was generally similar to the neuronal activity recorded during the same condition in the low-trained monkeys. Because for the low-trained monkeys the condition was absolutely new, it did not require any special effort from the monkeys and might be considered as Pavlovian condition only. We expected to see more and more similarities between the neuronal activities in the two groups with an increase in the number of sessions for the low-trained monkeys (Lubow and Gewirtz 1995, Lubow and Moore 1959). In other words, we consider that the presence of additional effort of the animals in the group of the well-trained monkeys during the presentation of paired acoustical stimuli and water did not lead to the changes in the neuronal activities. Rather, the higher experience of the well-trained monkeys and stronger association of the acoustical stimuli with the water delivery led to the differences in the neuronal activities between the two groups of the monkeys.

The comparison of the courses in the neuronal activities did not reveal large differences between the one passive and three instrumental conditions in the well-trained monkeys. Taking into account the consideration and the results of the neuronal activities in the passive condition that were mainly similar between the low- and well-trained monkeys, we considered that the neuronal activity in Pavlovian conditioning was generally similar to the neuronal activity in the instrumental conditions.

4.3.2. Phasic responses related to the motor/somatosensory events

The only intervals in which the courses of the neuronal activities had differences were related to the movements of the monkeys, leading to the additional motor and somatosensory inputs into the brain. Thus, there were the responses to the bar grasps and releases at different moments in the instrumental conditions and absence of such responses in the passive condition.

We observed that the more important bar releases related to the detection evoked responses in a higher fraction of units compare with the less important bar releases after the water deliveries. However, we did not observe that the responses to the more important bar releases related to detection evoked different responses than the bar releases after the water delivery. The results of the response size were partly inconsistent with the previous results showed by Germain and Lamarre (1993). They showed that the response to the more important movements led to higher response in the premotor cortex of monkeys.

In the present study we also showed that the sizes of the responses in the primary auditory cortex did not depended on the type of the movements, i.e., detection and initiation. The result was also inconsistent with the previous studies made by Romo and Schultz (1990). They showed that the responses to the movements related with the initiation were higher than the response evoked by detection in substantia nigra of monkeys.

We suppose that an additional reason of the higher number of units that responded to the motor acts with importance might be related to the mouth movements of the monkeys. We observed them to be synchronized with the bar grasps/releases in some sessions. It was found that polymodality, the increase in inputs, may lead to larger change in the neuronal activity (Bizley et al. 2007, Fu et al. 2003, Ghazanfar et al. 2005, Kayser et al. 2008, Lakatos et al. 2007). Thus, a probable reason of the higher response might be related to the additional sensory input from monkeys' mouth.

The present design of the study put monkeys in a situation where they preferred to use their left hand in order to initiate or detect acoustical stimuli in the trials. Due to methodological advantages, we made recording in the right side of the monkeys' auditory cortexes. Thus, a question arises whether we will find the same or very similar responses to the bar grasp/releases in the left auditory cortex when the monkeys perform the conditions with the same left hand.

4.4. Effects of the sense of agency

4.4.1. Towards the inhibition after the self-initiation

The inhibition of the neuronal activity after the initiation of the trials was one of the main effects of the sense of agency. We found it interesting that the frequency of the mouth movements also decreased after the self-initiation in some of the sessions. The coincidence of the decrease in the frequency of the mouth movements and in the neuronal activity at the same time led us to the idea of a relation between the two processes. As it is known, somatosensory stimuli activate neurons in the primary and secondary auditory cortexes (Brosch et al. 2005, Hoefer et al. 2013, Lakatos et al. 2007, Niwa et al. 2012b, Ro et al. 2013, Vaadia et al. 1982). Thus, a somatosensory response might be a good explanation of the upward shift in the neuronal activity after the water delivery in the auditory cortex, when the mouth moved more often; and of the decrease in the neuronal activity after the self-initiation, when the mouth moved rarely in half of the sessions. One of the studies reported that more than half units in the secondary auditory cortex responded to somatosensory stimulation of the head and neck and only fifth part of the units responded to the somatosensory stimuli from limbs (Fu et al. 2003). Although the research was conducted in the secondary auditory cortex, an analogue might be also observed in the primary auditory cortex. Thus, it gives the answer why we observed the correlations in the decrease in only part of the population. It is worth noting that we did not observe a good correlation between the responses to the bar grasps and the inhibition after the self-initiation but we found a correlation between the slow changes in the neuronal activity, which we related to the mouth movements, and the self-initiation. If the assumption about the distributions of units responding, or copying a signal, to the somatosensory stimuli from limbs and head are indeed correct than it will explain the correlations between the slow increase in the neuronal activity after the water delivery, the same increase in the mouth movements, the decrease in the frequency of the mouth movements after the self-initiation and the same decrease in the neuronal activity. Taking into account the proposed explanation, we found intriguing that the research that inspired us for the present study also explained the inhibition in the neuronal activity during vocalization in the marmosets through the activation of the somatosensory inputs (Eliades and Wang 2003, 2005). In contrast to our assumption, they concluded that the increase in somatosensory input led to the decrease in activity in the auditory cortex.

A more common explanation for the inhibition after the self-initiation is the motor-related signal in the auditory system, named efference copy or corollary discharge (Schneider and Mooney 2015). The theory says that the inhibition occurs due to the signal transmitted from the motor cortex. It assumes that the signal is necessary for the feedback and/or for the threshold decrease and enhancement of the sounds detection. Most of the studies that revealed similar inhibition (Buran et al. 2014, Carcea et al. 2017) explained the decreases after the self-initiation using the theory. Indeed, the connections between the motor cortex and the auditory cortex were confirmed by many researchers (Schneider et al. 2014, Schneider and Mooney 2015, Zhou et al. 2014). However, the theory does not explains why we also observed an increase of the neuronal activity some seconds and minutes later when the animals collected the water in the same trials and detected the sound in other conditions despite to the facts that these acts also had a motor, self-initiated nature (see sections 3.2.5, 3.3.4, 3.5.5). The increase in the neuronal activity after detection or discrimination was observed not only in the present study but in many studies of other groups (more details will be described in the section 4.5).

The enhancement of the perception that is explained by the corollary discharge becomes even less clear if we take into account the existence of the cochlear reflex (stapedius muscle contraction). The described property of the reflex is the protection of the auditory system from the self-initiated and potentially dangerous noises (Carmel and Starr 1964, Mukerji et al. 2010, Salomon and Starr 1963, Simmons 1964). The reflex must be triggered when a subject self-initiates a movement or an act that lead to the increase in the threshold in the auditory system. However, it is unclear how the cochlear reflex and the inhibition for the enhancement may function together.

4.4.2. Responses to the acoustical stimuli

The average neuronal activity during the acoustical stimulation, the noise, was lower when the monkeys initiated the sounds for themselves. We explained the origin of the differences in the activities due to the decrease of neuronal activity after the self-initiation before the acoustical stimulation (compare Figure 42A and B). We assumed that the inhibition after the self-initiation recovered for some time after the acoustical stimulation and that led to the lower activity in the self-initiated condition. Due to the absence of the decrease in the neuronal activity in the externally-initiated conditions, the level of the activity during the acoustical stimulation was higher.

Many other studies reported weaker responses (Creutzfeldt et al. 1989, Eliades and Wang 2003, Kirzinger and Jürgens 1991, Müller-Preuss and Ploog 1981) or lower MEG/EEG potentials and BOLD levels (Curio et al. 2000, Daliri and Max 2016, Horváth and Burgyán 2013, Houde et al. 2002, Martikainen 2005, Papanicolaou et al. 1986, Schafer and Marcus 1973) evoked by the presentation of the self-initiated sounds compare with the replayed or externally-initiated sounds. The different methods need different explanations, therefore possible interpretations are analyzed separately below.

We cannot judge whether the differences in BOLD levels and MEG/EEG potentials evoked by the acoustical stimulation indicate that the same differences and ratios will be observed in the activities of individual neurons. The reason of the issue is the unclear relationships between the responses in the neuronal activity of the individual cells or small clusters used in the studies and BOLD signals (Logothetis et al. 2001, Logothetis and Wandell 2004). Similarly, we cannot judge the relationships between the resulting potentials of the EEG/MEG and the neuronal activities recorded in the cells. The only fact we are convinced in is that the differences between the potentials recorded using EEG/MEG or BOLD signals indicate to the differences in the responses in the area including the auditory cortex. However, it is not obvious and is not known yet what kind of differences we should expect.

A direct comparison of the previous results, in which the difference was shown in the neuronal activities between the self-produced vocalizations and replays (Creutzfeldt et al. 1989, Eliades and Wang 2003, Kirzinger and Jürgens 1991, Müller-Preuss and Ploog 1981), and the present study is complex. The reason of this is the use of replays in the studies of our colleagues. In the present study, the externally-initiated sounds were potentially interesting for the animal because it predicted the water delivery directly or indirectly, depending on the condition and monkeys' performance; when the externally-initiated sounds in the referred studies were unimportant for the subjects. Therefore, the question rose whether the differences that were observed by our colleagues were caused by the change of the sense of agency or by the meaning of the acoustical stimulations.

We would like to highlight a recently published study made by Carcea and colleagues (2017) on rats. They used the “go/no go” paradigm. In one task, the rats had to initiate the trials with a nose poke and wait for the tone onset that could be a target or no target. In another task, the rats waited for the tone onset that also could be target or non-target. The advantage of the study was the possibility to compare the neuronal activities in the self- and externally-initiated trials and consider the level of effort and sound meaning, the same as it was made in the our study. The acoustical stimulus was artificial in both tasks and was generated by a training computer. In contrast to our study, (1) the rats were not required to prolong their poke until and during the tone stimulation; (2) the tone stimulation lasted only 100 ms. Similar to the results of the present study, Carcea and colleagues (2017) reported the decrease in activity after the self-initiation before the tone onset. Contrary to the present study, the colleagues reported smaller response to the tone onset in the self-initiated compare with the externally-initiated conditions. We assume that the differences between the results of the present study and the study of the Carcea et al. (2017) might be due to the differences in the methods mentioned above.

4.5. Effect of effort

4.5.1. Increase in the neuronal activity after detection

One of the main finding related to the level of effort was the higher neuronal activity during tone stimulations after detections in the conditions with high effort. What we first supposed and controlled in the study was the correlation of the increase with the motor or somatosensory response after the bar grasps/releases but the results showed no correlation. Then we controlled the correlation of the increase after the detection with the slow changes in the neuronal activity and found a correlation. Interestingly, some of the sessions showed that the mouth movements also had a higher frequency after the detection. However, only some units recorded during these sessions had a similar increase in the neuronal activities as the mouth movements. The similar activity in some of the units might be explained by the somatosensory inputs from the mouth to the primary auditory cortex. However, such inputs were shown only for the secondary but not for the primary auditory cortex (Fu et al. 2003, was described in more details in the section 4.4.1).

From another point of view, the increase in the frequency of the mouth movements, which was synchronized with the detection, might be a marker for the higher expectations of the water delivery for the monkeys. Such connections were already described in the primary auditory cortexes by other groups who showed a gradual increase in the activity after a signal and before the following reinforcement (Shinba et al. 1995, Quirk et al. 1997). The gradual increase in the neuronal activity were also described in the sub-cortical structures (Komura et al. 2001, Metzger et al. 2006) and in other sensory cortexes (Amador et al. 2000, Pantoja et al. 2007, Shuler and Bear 2006). In the present study, we showed the absence of any gradual increase but only a temporal higher activity during the tone presentation before the water delivery depending on the level of effort.

4.5.2. Decrease in the neuronal activity and animals' engagement

There were already approaches to find the effect of the engagement in the neuronal activity in the primary auditory cortex. In contrast to our study, the previous studies described an increase of the responses to the acoustical stimuli with the engagement (David et al. 2012, Fritz et al. 2005, Fritz et al. 2007b, 2007c, Fritz et al. 2010, Hubel et al. 1959, Lu et al. 2016, Yin et al. 2014). Some studies also showed that the engagement led to lower responses to the acoustical stimuli (Otazu et al. 2009, Zhong et al. 2016). The present study did not reveal any differences between the responses to the acoustical events in the conditions with different engagement, i.e., levels of effort. We assume that the reason of the differences in the cited studies was due to the presence of the unconditioned stimuli and not due to the change of the engagement.

The present study showed an increase in the activity during the tone stimulation that was observed after the detection. A similar effect was observed by one group of researchers (Atiani et al 2014, Fritz et al. 2010) after the presentation of a sound that indicated a negative reinforcement and the animal had to stop licking. The authors interpreted the result as the effect of attention in the primary and higher auditory cortexes. As was already explained previously (4.5.1), we concluded that the increase after the detection, during the tone stimulation, was rather related to some additional input from motor/somatosensory areas rather than the different response to the tone stimulation. The authors (Atiani et al 2014, Fritz et al. 2010) revealed the increase in the neuronal activity after comparison of the instrumental condition with a passive listening. In our study, the effect was observed between the conditions with two level of effort. Therefore, the results of the present study expanded the previous and showed that the similar effect is present between two conditions that are differed with only one factor, i.e., with the level of effort.

4.6. Conclusion

In the present study, we concluded that the meaning of sound and presence of the unconditioned stimuli changed the neuronal activity drastically. Both effects intensified with the experience of the animals. Also we revealed the effects of the sense of agency and of levels of effort on the neuronal activity in the primary auditory cortex of monkeys. These results were consistent with and expanded upon previous studies showing that the neurons in the primary auditory cortex are sensitive to engagement. Moreover, we have shown that different forms of engagement should be considered separately because they influence the neuronal activity in the primary auditory cortex unequally.

5. References

Abolafia JM, Martinez-Garcia M, Deco G, Sanchez-Vives MV. (2011). Slow modulation of ongoing discharge in the auditory cortex during an interval-discrimination task. *Front. Integr. Neurosci.* 5(60):1-6.

Amador N, Schlag-Rey M, Schlag J. (2000). Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. *J. Neurophysiol*. 84(4):2166-70.

Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, Fritz JB. (2014). Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82:486-499.

Bao S, Chan VT, Merzenich MM. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412(6842):79-83.

Bizley JK, Nodal FR, Bajo VM, Nelken I, King AJ. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb. Cortex* 17(9):2172-89.

Blake DT, Heiser MA, Caywood M, Merzenich MM. (2006). Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. *Neuron* 52(2):371-81.

Blake DT, Strata F, Churchland AK, Merzenich MM. (2002). Neural correlates of instrumental learning in primary auditory cortex. *Proc. Natl. Acad. Sci. USA*. 99(15):10114-9.

Bondar' IV, Vasil'eva LN, Badakva AM, Miller NV, Zobova LN, Roshchin VIu. (2014). Quality of neuronal signal registered in the monkey motor cortex with chronically implanted multiple microwires. *Zh. Vyssh. Nerv. Deiat. Im. IP Pavlova* 64(1):101-12.

Boudreau CE1, Williford TH, Maunsell JH. (2006). Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *J. Neurophysiol.* 96(5):2377-87.

Bouret S, Richmond BJ. (2009). Relation of locus coeruleus neurons in monkeys to Pavlovian and operant behaviors. *J. Neurophysiol*. 101(2):898-911. doi: 10.1152/jn.91048.2008.

Bouton ME. (2007). Learning and behavior: A contemporary synthesis. Sunderland, MA: Sinauer Associates.

Bowman EM, Aigner TG, Richmond BJ. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol*. 75(3):1061-73.

Brosch M, Scheich H. (2008). Tone-sequence analysis in the auditory cortex of awake macaque monkeys. *Exp. Brain Res*. 184(3):349-61.

Brosch M, Selezneva E, Scheich H. (2005). Nonacoustical events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.* 25 (29):6797-6806.

Brosch M, Selezneva E, Scheich H. (2011a). Formation of associations in auditory cortex by slow changes of tonic firing. *Hear. Res.* 271:66-73.

Brosch M, Selezneva E, Scheich H. (2011b) Representation of reward feedback in primate auditory cortex. *Front. Syst. Neurosci.* 5(5):1-12.

Brosch M, Selezneva E, Scheich H. (2015). Neuronal activity in primate auditory cortex during the performance of audiovisual tasks. *Eur. J. Neurosci*. 41(5):603-14. doi: 10.1111/ejn.12841.

Budinger E, Laszcz A, Lison H, Scheich H, Ohl FW. (2008). Non-sensory cortical and subcortical connections of the primary auditory cortex in Mongolian gerbils: bottom-up and top-down processing of neuronal information via field AI. *Brain Res.* 1220:2–32.

Buran BN, von Trapp G, Sanes DH. (2014). Behaviorally gated reduction of spontaneous discharge can improve detection thresholds in auditory cortex. *J. Neurosci*. 34(11):4076-81. doi: 10.1523/JNEUROSCI.4825-13.2014.

Cameron CM, Carelli RM. (2012). Cocaine abstinence alters nucleus accumbens firing dynamics during goal-directed behaviors for cocaine and sucrose. *Eur. J. Neurosci*. 35(6):940–951. doi: 10.1111/j.1460-9568.2012.08024.x

Carcea I, Insanally MN, Froemke RC. (2017). Dynamics of auditory cortical activity during behavioural engagement and auditory perception. *Nature Commun*. 8:14412. doi: 10.1038/ncomms14412.

Carmel PW, Starr A. (1964). Non-acoustic Factors influencing Activity of Middle Ear Muscles in Waking Cats. *Nature* 202:195–196. doi:10.1038/202195a0.

Cartoni E, Balleine B, Baldassarre G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci. Biobehav. Rev*. 71:829-848. doi: 10.1016/j.neubiorev.2016.09.020.

Chen Y, Martinez-Conde S, Macknik SL, Bereshpolova Y, Swadlow HA, Alonso JM. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nat. Neurosci*. 11(8):974-82. doi: 10.1038/nn.2147.

Chernyshev BV, Weinberger NM. (1998). Acoustic frequency tuning of neurons in the basal forebrain of the waking guinea pig. *Brain Res*. 793(1-2):79-94.

Chubykin AA, Roach EB, Bear MF, Shuler MG. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77(4):723-35. doi: 10.1016/j.neuron.2012.12.039.

Creutzfeldt O, Ojemann G, Lettich E. (1989). Neuronal activity in the human lateral temporal lobe. I. Responses to speech. *Exp. Brain Res*. 77(3):451-75.

Curio G, Neuloh G, Numminen J, Jousmäki V, Hari R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Hum. Brain Mapp*. 9(4):183-91.

Daliri A, Max L. (2016). Modulation of Auditory Responses to Speech vs. Nonspeech Stimuli during Speech Movement Planning. *Front. Hum. Neurosci*. 10:234. doi: 10.3389/fnhum.2016.00234. eCollection 2016.

David SV, Fritz JB, Shamma SA. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. *PNAS* 109(6): 2144-2149.

Deitmer JW. (2001). Strategies for metabolic exchange between glial cells and neurons. *Respir. Physiol*. 129(1-2):71-81.

Denton SE, Kruschke JK. (2006). Attention and salience in associative blocking. *Learn. Behav. 34:*285–304

Diamond DM, Weinberger NM. (1984). Physiological plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response: II. Secondary field (AII). *Behav. Neurosci*. 98(2):189-210.

Diamond DM, Weinberger NM. (1986). Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. *Brain Res*. 372(2):357-60.

Diamond DM, Weinberger NM. (1989). Role of context in the expression of learning-induced plasticity of single neurons in auditory cortex. *Behav. Neurosci*. 103(3):471-94.

Driver J, Noesselt T. (2008). Multisensory Interplay Reveals Crossmodal Influences on ‘Sensory-Specific’ Brain Regions, Neural Responses, and Judgments. *Neuron* 57(1):11–23. doi: 10.1016/j.neuron.2007.12.013.

Eliades SJ, Wang X. (2003). Sensory-motor interaction in the primate auditory cortex in the self-initiated vocalizations. *J. Neurophysiol*. 89(4):2194-207.

Eliades SJ, Wang X. (2005). Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cereb. Cortex*. 15(10):1510-23.

Fanselow MS, Wassum KM. (2015). The Origins and Organization of Vertebrate Pavlovian Conditioning. *Cold Spring Harb. Perspect. Biol*. 8(1):a021717. doi: 10.1101/cshperspect.a021717.

Farrer C, Frith CD. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* 15(3):596-603.

Fishman YI, Kim M, Steinschneider M. (2017). A Crucial Test of the Population Separation Model of Auditory Stream Segregation in Macaque Primary Auditory Cortex. *J. Neurosci.* 37(44):10645-10655. doi: 10.1523/JNEUROSCI.0792-17.2017.

Fishman YI, Micheyl C, Steinschneider M. (2012). Neural mechanisms of rhythmic masking release in monkey primary auditory cortex: implications for models of auditory scene analysis. *J. Neurophysiol*. 107(9):2366-82. doi: 10.1152/jn.01010.2011.

Fishman YI, Reser DH, Arezzo JC, Steinschneider M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear. Res*. 151(1-2):167-187.

Fontanini A, Katz DB. (2006). State-dependent modulation of time-varying gustatory responses. *J. Neurophysiol.* 96:3183-3193.

Fritz JB, David SV, Radtke-Shuller S, Yin P, Shamma SA. (2010). Adaptive, behaviorally gates, persistent encoding of task-relevant information in ferret frontal cortex. *Nature Neurosci.* 13(8):1011-1021.

Fritz JB, Elhilali M, David SV, Shamma SA. (2007a). Auditory attention-focusing the searchlight on sound. *Curr. Opin. Neurobiol.* 17(4):437-55.

Fritz JB, Elhilali M, David SV, Shamma SA. (2007b). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear. Res.* 229(1-2):186-203.

Fritz JB, Elhilali M, Shamma S. (2005). Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hear Res*. 206(1-2):159-76.

Fritz JB, Elhilali M, Shamma SA. (2007c). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J. Neurophysiol.* 98(4):2337-46.

Fritz J, Shamma S, Elhilali M, Klein D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6(11):1216-23.

Fu KMG, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE. and Schroeder CE. (2003). Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci*. 23(20):7510-7515.

Germain L, Lamarre Y. (1993). Neuronal activity in the motor and premotor cortices before and after learning the associations between auditory stimuli and motor responses. *Brain Res*. 611(1):175-9.

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci*. 25(20):5004-12.

Ghose GM. (2004). Learning in mammalian sensory cortex. *Curr. Opin. Neurobiol.* 14(4):513-8.

Ghose GM, Yang T, Maunsell JH. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* 87(4):1867-88.

Gilat E, Perlman I. (1984). Single unit activity in the auditory cortex and the medial geniculate body of the rhesus monkey: behavioral modulation. *Brain Res*. 324:323–333, doi:10.1016/0006-8993(84)90043-X.

Gorkin AG. (2017). Some features of the structure of individual experience revealed by the activity of specialized neurons. Fundamental and applied research in modern psychology. Edited by Zhuravlyov AL, Koltzova VA. Moscow, 1531-1537.

Gorkin AG, Shevchenko DG. (1995). Distinction in activity of neurons of the limbic cortex depending on different strategies of training. *IP Pavlov J. High. Nerv. Act.*45(1):90-100.

Hage SR, Jürgens U. (2006). On the role of the pontine brainstem in vocal pattern generation: a telemetric single-unit recording study in the squirrel monkey. *J. Neurosci.* 26(26):7105-15.

Hocherman S, Benson DA, Goldstein J, Heffner HE, Hienz RD. (1976). Evoked unit activity in auditory cortex of monkeys performing a selective attention task. *Brain Res.* 117:51–68, doi:10.1016/0006-8993(76)90555-2.

Hoefer M, Tyll S, Kanowski M, Brosch M, Schoenfeld MA, Heinze HJ, Noesselt T. (2013). Tactile stimulation and hemispheric asymmetries modulate auditory perception and neural responses in primary auditory cortex. *Neuroimage* 79:371-82. doi: 10.1016/j.neuroimage.2013.04.119.

Hollerman JR, Schultz W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1(4):304-9.

Horváth J, Burgyán A. (2013). No evidence for peripheral mechanism attenuating auditory ERPs to self-induced tones. *Psychophysiol.* 50(6):563-9. doi: 10.1111/psyp.12041.

Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. (2002). Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14(8):1125-38.

Huang Y, Brosch M. (2016). Neuronal activity in primate prefrontal cortex related to goal-directed behavior during auditory working memory tasks. *Brain Res.* 1640(Pt B):314-27. doi: 10.1016/j.brainres.2016.02.010.

Huang Y, Matysiak, Heil P, König R, Brosch M. (2016a). Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. *Elife* 5. pii: e15441. doi: 10.7554/eLife.15441.

Huang Y, Mylius J, Scheich H, Brosch M. (2016b). Tonic effect of the dopaminergic ventral midbrain on the auditory cortex of awake macaque monkeys. *Brain Struct. Funct. 221(2):969-77. doi: 10.1007/s00429-014-0950-2.*

Hubel DH, Henson CO, Rupert A, Galambos R. (1959). Attention units in the auditory cortex. *Science* 129(3358):1279-80.

Huertas MA, Hussain Shuler MG, Shouval HZ. (2015). A Simple Network Architecture Accounts for Diverse Reward Time Responses in Primary Visual Cortex. *J. Neurosci.* 35(37):12659-72. doi: 10.1523/JNEUROSCI.0871-15.2015.

Ide Y, Takahashi M, Lauwereyns J, Sandner G, Tsukada M. (2012). Fear conditioning induces guinea pig auditory cortex activation by foot shock alone. *Cogn. Neurodyn.* 7(1): 67–77. doi: 10.1007/s11571-012-9224-y.

Ilango A1, Wetzel W, Scheich H, Ohl FW. (2010). The combination of appetitive and aversive reinforcers and the nature of their interaction during auditory learning. *Neuroscience.* 166(3):752-62. doi: 10.1016/j.neuroscience.2010.01.010.

Jackson PL, Decety J. (2004). Motor cognition: A new paradigm to study self-other interactions. *Curr. Opin. Neurobiol.* 14:259–63.

Jeannerod M. (2003). The mechanism of self-recognition in human. *Beh. Brain Res.* 142:1-15.

Kaas JH, Hackett TA. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U S A.* 97(22):11793-9.

Kahneman D. (1973). Attention and effort. Prentice –Hall, Inc., Englewood Cliffs, New Jersey.

Kahneman D, Beatty J. (1966). Pupil diameter and load on memory. *Science* 154(3756):1583-5.

Kajikawa Y, Camalier CR, de la Mothe LA, D'Angelo WR, Sterbing-D'Angelo SJ, Hackett TA. (2011). Auditory cortical tuning to band-pass noise in primate A1 and CM: a comparison to pure tones. *Neurosci. Res.* 70(4):401-7. doi: 10.1016/j.neures.2011.04.003.

Kayser C, Petkov C, Logothetis NK. (2008).Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18(7):1560-74. doi: 10.1093/cercor/bhm187.

Kim Y, Simon NW, Wood J, Moghaddam B. (2016). Reward Anticipation Is Encoded Differently by Adolescent Ventral Tegmental Area Neurons. *Biol. Psychiatry*. 79(11):878-86. doi: 10.1016/j.biopsych.2015.04.026.

Kirzinger A, Jürgens U. (1991). Vocalization-correlated single-unit activity in the brain stem of the squirrel monkey. *Exp. Brain Res*. 84(3):545-60.

Kluge C, Bauer M, Leff AP, Heinze HJ, Dolan RJ, Driver J. (2011). Plasticity of human auditory-evoked fields induced by shock conditioning and contingency reversal. *Proc. Natl. Acad. Sci. USA* 108(30):12545-50. doi: 10.1073/pnas.1016124108.

Knyazeva S, Selezneva E, Gorkin A, Aggelopoulos NC, Brosch M. (2018). Neuronal correlates of auditory streaming in monkey auditory cortex for tone sequences without spectral differences. *Front. Integr. Neurosci*. 12:4. doi: 10.3389/fnint.2018.00004

Komura Y, Tamura R, Uwano T, Nishijo H, Kaga K, Ono T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* 412:546-49.

Konorski J. (1948). Conditioned Reflexes and Neuron. Organization. Cambridge University Press. Cambridge, UK.

Konorski J. (1967). Integrative Activity of the Brain. University of Chicago Press, Chicago, IL.

Kraus N, Disterhoft JF. (1982). Response plasticity of single neurons in rabbit auditory association cortex during tone-signalled learning. *Brain Res*. 246(2):205-15.

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53(2):279-92.

Lee CC, Middlebrooks JC. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nat. Neurosci*. 14(1):108-14. doi: 10.1038/nn.2713.

Le Pelley ME, Mitchell CJ, Beesley T, George DN, Wills AJ. (2016). Attention and associative learning in humans: An integrative review. *Psychol. Bull*. 142(10):1111-40. doi: 10.1037/bul0000064.

Ljungberg T, Apicella P, Schultz W. (1991). Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain. Res*. 567(2):337-41.

Ljungberg T, Apicella P, Schultz W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol*. 67(1):145-63.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412(6843):150-7.

Logothetis NK, Wandell BA. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol*. 66:735-69.

Lou Y, Luo W, Zhang G, Tao C, Chen P, Zhou Y, Xiong Y. (2014). Ventral tegmental area activation promotes firing precision and strength through circuit inhibition in the primary auditory cortex. *Front. Neural Circuits*. 8:25. doi: 10.3389/fncir.2014.00025.

Lovell JM, Mylius J, Scheich H, Brosch M. (2014). Hearing in action; auditory properties of neurons in the red nucleus of alert primates. *Front. Neurosci*. 8:105. doi: 10.3389/fnins.2014.00105. eCollection 2014.

Lu K, Xu Y, Yin P, Oxenham AJ, Fritz JB, Shamma SA. (2016). Temporal coherence structure rapidly shapes neuronal interactions. *Nat. Commun*. 8:13900. doi: 10.1038/ncomms13900.

Lubow RE, Gewirtz JC. (1995). Latent inhibition in humans: data, theory, and implications for schizophrenia. *Psychol. Bull*. 117(1):87-103.

Lubow RE, Moore AU. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *J. Comp. Physiol. Psychol.* 52:415-9.

Namboodiri VM, Huertas MA, Monk KJ, Shouval HZ, Shuler HMG. (2015). Visually cued action timing in the primary visual cortex. *Neuron* 86(1):319-30. doi: 10.1016/j.neuron.2015.02.043.

Numminen J, Salmelin R, Hari R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neurosci. Lett.* 265(2):119-22.

Martikainen MH, Kaneko K, Hari R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb. Cortex*. 15(3):299-302.

Martin-Soelch C, Linthicum J, Ernst M. (2007). Appetitive conditioning: neural bases and implications for psychopathology. *Neurosci. Biobehav. Rev.* 31(3):426-40.

Maunsell JH and Cook EP. (2002). The role of attention in visual processing. Philos Trans R *Soc. Lond. B. Biol. Sci.* 357(1424):1063-72

Metzger RR, Greene NT, Porter KK, Groh JM. (2006). Effects of reward and behavioral context on neural activity in the primate inferior colliculus. *J. Neurosci*. 26(28):7468-76.

Moller A. (2006). Hearing: Anatomy, Physiology, and Disorders of the Auditory System.Academic Press, 2nd Edition.

Morrison SE, Salzman CD. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.* 29(37):11471-11483.

Müller-Preuss P, Ploog D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Res.* 215(1-2):61-76.

Mukerji S, Windsor AM, Lee DJ. (2010). Auditory brainstem circuits that mediate the middle ear muscle reflex. *Trends Amplif*. 14(3):170-91. doi: 10.1177/1084713810381771.

Mylius J, Happel MKF, Gorkin AG, Huang Y, Scheich H, Brosch M. (2014). Fast transmission from the dopaminergic ventral midbrain to the sensory cortex of awake primates. *Brain Struct. Funct.* 220(6):3273-94. DOI 10.1007/s00429-014-0855-0.

Niwa M, Johnson JS, O’Connor KN, Sutter ML. (2012a). Active engagement improves primary auditory cortical neurons’ ability to discriminate temporal modulation. *J. Neurosci*. 32:9323–34. doi:10.1523/JNEUROSCI.5832-11.2012.

Niwa M, Johnson JS, O’Connor KN, Sutter ML. (2012b). Activity related to perceptual judgment and action in primary auditory cortex. *J. Neurosci*. 32:3193–210. doi:10.1523/JNEUROSCI.0767-11.2012.

Ohl FW, Scheich H. (1996). Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field Al) of the alert Mongolian gerbil. *Eur. J. Neurosci*. 8(5):1001-17.

Ohl FW, Scheich H. (1997). Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *J. Comp. Physiol. A.* 181(6):685-96.

Ohl FW, Scheich H. (2005). Learning-induced plasticity in animal and human auditory cortex. *Curr. Opin. Neurobio.* 15:470–77.

Otazu GH, Tai LH, Yang Y, Zador AM. (2009) Engaging in auditory task suppresses responses in auditory cortex. *Nat. Neurosci*. 12(5):646-54.

Pantoja J, Ribeiro S, Wiest M, Soares E, Gervasoni D, Lemos NA, Nicolelis MA. (2007). Neuronal activity in the primary somatosensory thalamocortical loop is modulated by reward contingency during tactile discrimination. *J. Neurosci.* 27(39):10608-20.

Papanicolaou AC, Raz N, Loring DW, Eisenberg HM. (1986). Brain stem evoked response suppression during speech production. *Brain Lang*. 27(1):50-5.

Pavlov IP. (1927). Conditional Reflexes. Editor: Christopher D. Green. York University, Toronto, Ontario.

Pi HJ, Hangya B, Kvitsiani D, Sanders J, Huang ZJ, Kepecs A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* 503(7477):521-4. doi: 10.1038/nature12676.

Puschmann S, Brechmann A, Thiel CM. (2013). Learning-dependent plasticity in human auditory cortex during appetitive operant conditioning. *Hum. Brain Mapp.* 34: 2831-2851.

Quirk GJ, Armony JL, LeDoux JE. (1997). Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* 19(3):613-24.

Rauschecker JP, Tian B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J. Neurophysiol*. 91(6):2578-89.

Richmond BJ, Sato T. (1987). Enhancement of inferior temporal neurons during visual discrimination. *J. Neurophysiol*. 58(6):1292-306.

Ro T, Ellmore TM, Beauchamp MS. (2013). A neural link between feeling and hearing. *Cereb. Cortex.* 23(7):1724-30. doi: 10.1093/cercor/bhs166.

Rolls ET. (2004). The functions of the orbitofrontal cortex. *Brain Cogn*. 55(1):11-29.

Rolls ET. (2016). Reward Systems in the Brain and Nutrition. *Annu. Rev. Nutr*. 36:435-70. doi: 10.1146/annurev-nutr-071715-050725.

Rolls ET, Deco G. (2016). Non-reward neural mechanisms in the orbitofrontal cortex. *Cortex* 83:27-38. doi: 10.1016/j.cortex.2016.06.023.

Rolls ET, Grabenhorst F. (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Prog. Neurobiol*. 86(3):216-44. doi: 10.1016/j.pneurobio.2008.09.001.

Romo R, Schultz W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch in the self-initiated arm movements. *J. Neurophysiol*. 63(3):592-606.

Rowland V, Gluck H, Sumergrad S, Dines G. (1985). Slow and multiple unit potentials in trace and temporal conditioning controlled by electrical reward in the rat. *Electroencephalogr. Clin. Neurophysiol*. 61(6):559-68.

Salomon G, Starr A. (1963). Electromyography of middle ear muscles in man during motor activities. *Acta. Neurol. Scand*. 39:161-8.

Schafer EW, Marcus MM. (1973). Self-stimulation alters human sensory brain responses. *Science* 181(4095):175-7.

Schanze T. (2017). Removing noise in biomedical signal recordings by singular value decomposition. *Curr. Direc. Biomed. Eng*. 3(2): 253–256.

Scheich H. (1991). Auditory cortex: comparative aspects of maps and plasticity. *Curr. Opin. Neurobiol.*1(2):236-47.

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW, Selezneva E, Stark H, Tischmeyer W, Wetzel W. (2011). Behavioral semantic of learning and crossmodal processing in auditory cortex: The semantic processor concept. *Hear. Res.* 271: 3-15.

Schneider DM, Mooney R. (2015). Motor-related signals in the auditory system for listening and learning. *Curr. Opin. Neurobiol*. 33:78-84. doi: 10.1016/j.conb.2015.03.004.

Schneider DM, Nelson A, Mooney R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513(7517):189-94. doi: 10.1038/nature13724.

Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol*. 85(3):1322-7.

Schultz W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol*. 80(1):1-27.

Schultz W. (2002). Getting formal with dopamine and reward. *Neuron* 36:241-263.

Scott BH, Malone BJ, Semple MN. (2007). Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. *J. Neurosci.* 27:6489–99, doi:10.1523/JNEUROSCI.0016-07.2007.

Sechenov I. (1863). Reflexes of the brain. The MIT Press (March 15, 1965).

Sejnowski TJ, Poizner H, Lynch G, Gepshtein S, Greenspan RJ. (2014). Prospective Optimization. *Proc. IEEE Inst. Electr. Electron. Eng.* 102(5). doi: 10.1109/JPROC.2014.2314297.

Selezneva E, Oshurkova E, Scheich H, Brosch M. (2017). Category-specific neuronal activity in left and right auditory cortex and in medial geniculate body of monkeys. *PLoS One* 12(10):e0186556. doi: 10.1371/journal.pone.0186556. eCollection 2017.

Shamma SA, Elhilali M, Micheyl C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 34(3):114-23. doi: 10.1016/j.tins.2010.11.002.

Shimada S, Hiraki K, Oda I. (2005). The parietal role in the sense of self-ownership with temporal discrepancy between visual and proprioceptive feedbacks. *NeuroImage* 24(4):1225-1232. https://doi.org/10.1016/j.neuroimage.2004.10.039.

Shinba T, Sumi M, Iwanami A, Ozawa N, Yamamoto K. (1995). Increased neuronal firing in the rat auditory cortex associated with preparatory set. *Brain Res. Bull.* 37(2):199-204.

Shuler MG, Bear MF. (2006). Reward timing in the primary visual cortex. *Science* 311:1606-09.

Siegel JG, Agranoff BW, Albers RW, Fisher SK, and Uhler MD. (1999). Basic Neurochemistry, 6th edition. Molecular, Cellular and Medical Aspects. ISBN-10: 0-397-51820-X.

Simmons FB. (1964). Middle ear muscle acoustic reflex as index of cochlear sensitivity in auditory experiments: Some technical notes. *J. Aud. Res.* 4(4):255-260.

Skinner BF. (1938)."The Behavior of Organisms: An Experimental Analysis", New York: Appleton-Century-Crofts.

Spitzer H, Richmond BJ. (1991). Task difficulty: ignoring, attending to, and discriminating a visual stimulus yield progressively more activity in inferior temporal neurons. *Exp. Brain Res.* 83(2):340-8.

Sprague JM and Epstein AN. (1983). Progress in psychobiology and physiological psychology. Volume 10. Academoc Press, New York. Pp: 198-276.

Stein BE and Meredith MA. (1993). The merging of the senses. Boston: MIT.

Suga N, Ma X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci*. 4(10):783-94.

Suga N, Xiao Z, Ma X, Ji W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36:9–18.

Szabo J, Cowan WM. (1984). A stereotaxic atlas of the brain of the cynomolgus monkey (Macaca fascicularis). *J. Comp. Neurol*. 222(2):265-300.

Thorndike EL. (1901). "Animal intelligence: An experimental study of the associative processes in animals". Psychological Review Monograph Supplement. 2: 1–109.

Tobler PN, Dickinson A, Schultz W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23(32):10402-10.

Tremblay L, Hollerman JR, Schultz W. (1998). Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J. Neurophysiol*. 80(2):964-77.

Vaadia E, Gottlieb Y and Abeles M. (1982). Single-unit activity related to sensorimotor association in auditory cortex of a monkey. *J. Neurophysiol.* 48 (5):1201-1213.

Van Hamme LJ, Wasserman EA. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learn. Motiv*. 25:127–151.

Waelti P, Dickinson A, Schultz W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412(6842):43-8.

Weinberger NM. (2004). Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* 5(4):279-90.

Weinberger NM. (2007). Auditory associative memory and representational plasticity in the primary auditory cortex. *Hear. Res.* 229(1-2):54-68.

Weinberger NM, Diamond DM. (1987). Physiological plasticity in auditory cortex rapid induction by learning. *Prog. Neurobio.* 29:1-55.

Weinberger NM, Hopkins W, Diamond DM. (1984). Physiological plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response: I. Primary field (AI). *Behav. Neurosci*. 98(2):171-88.

Weis T, Puschmann S, Brechmann A, Thiel CM. (2012). Effects of L-dopa during auditory instrumental learning in humans. *PLoS One* 7(12):e52504. doi: 10.1371/journal.pone.0052504.

Yin P, Fritz JB, and Shamma SA. (2014). Rapid Spectrotemporal Plasticity in Primary Auditory Cortex during Behavior. *J. Neurosci*. 34(12):4396–4408. doi: 10.1523/JNEUROSCI.2799-13.2014.

Zhong R, Ma L, Qin L. (2016). Engaging in a tone-detection task differentially modulates neural activity in the auditory cortex, amygdala, and striatum. *Sci. Rep*. 7(1):677. doi: 10.1038/s41598-017-00819-z.

Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, Tao HW, Zhang LI. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.* 17(6):841-50. doi: 10.1038/nn.3701.

6. Supplementary materials

6.1. Supplementary tables

Sample size	Group of monkeys	Compared conditions						Controlled effects
		CS-	US	CS+	Ext-HighE	Self-HighE	Self-LowE	
		Average of trials with SEM						
75 units	Low-trained	63.3±0.8	51.6±1.5	62.4±0.5	-	-	-	1. Meaning of sound, 2. Presence of unconditioned stimuli
97 units	Well-trained	44.8±1.5	45.4±1.3	45.4±1.1	-	-	-	
180 units	Well-trained	-	-	44.1±0.6	41.4±0.7	45.3±0.6	46.2±0.5	1. Sense of agency, 2. Attention (level of effort)

Supplementary table 1. Sample sizes used in the analyses, conditions during which they were recorded and the effects that were controlled.

The table shows the number of trials in each condition with the sound type “Tone followed by noise” only. The numbers of trials were similar in the conditions “Tone only”.

Table index				Low-trained monkeys			Well-trained monkeys		
	Comparison	Time window		CS+ vs CS-		Conclusion	CS+ vs CS-		Conclusion
		Main	Additional time window	CS+ vs CS-	P-value	Significance	CS+ vs CS-	P-value	Significance
A (Sections 3.1.1 and 3.2.2)	Noise onset (absolute)	0-100 ms after event	-	>	0.02	*	<	$3*10^{-8}$	***
	Noise offset (absolute)	0-100 ms after event	-	>	$7*10^{-5}$	***	=	0.19	ns
	Tone onset (absolute)	0-100 ms after event	-	=	0.15	ns	<	$6*10^{-3}$	**
	Tone offset (absolute)	0-100 ms after event	-	>	0.02	*	<	0.03	*
B (Sections 3.1.1 and 3.2.2)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	=	0.07	ns	<	$8*10^{-12}$	***
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	=	0.38	ns	=	0.33	ns
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	=	0.11	ns	=	0.06	ns
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	=	0.98	ns	=	0.24	ns
C (Sections 3.1.2 and 3.2.3)	Noise (absolute)	200-1600 ms after event	-	>	$3*10^{-7}$	***	=	0.14	ns
	Tone (absolute)	200-1500 ms after tone onset	-	>	$4*10^{-6}$	***	=	0.09	ns
D (Sections 3.1.2 and 3.2.3)	Noise (change)	200-1600 ms after noise onset	-1000-0 ms before noise onset	=	0.11	ns	<	$4*10^{-9}$	***
	Tone (change)	200-1500 ms after tone ones	-1000-0 ms before noise onset	<	$3*10^{-3}$	**	<	$3*10^{-3}$	**
E (Sections 3.1.4 and 3.2.5)	Baseline	-1000-0 ms before the noise onset	-	>	$1*10^{-8}$	***	>	$1*10^{-3}$	**

Table index	Comparison	Time window		CS+ vs US		Conclusion	CS+ vs US		Conclusion
		Main	Additional time window	CS+ vs US	P-value	Significance	CS+ vs US	P-value	Significance
F (Sections 3.1.4 and 3.2.5)	Baseline	-4300 – -3300 ms before water delivery	-	=	0.09	ns	=	0.26	ns

Supplementary table 2. Results of the effect of the sound meaning and the presence of the unconditioned stimuli in the population activity.

The stars indicate the level of significance: ns - p > 0.05, * - p < 0.05; ** - p < 0.01; *** - p < 0.001. The green color indicates the significant differences observed in the population activities between the conditions.

Table index				Low-trained monkeys		Well-trained monkeys	
	Comparison	Time window		CS+ vs CS-		CS+ vs CS-	
		Main	Additional time window	CS+ > CS-	CS+ < CS-	CS+ > CS-	CS+ < CS-
A (Sections 3.1.1 and 3.2.2)	Noise onset (absolute)	0-100 ms after event	-	21.3%	10.7%	4.1%	26.8%
	Noise offset (absolute)	0-100 ms after event	-	20%	5.3%	7.2%	15.5%
	Tone onset (absolute)	0-100 ms after event	-	25.3%	12%	7.2%	15.5%
	Tone offset (absolute)	0-100 ms after event	-	21.3%	10.7%	5.2%	18.6%
B (Sections 3.1.1 and 3.2.2)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	9.3%	30.7%	5.1%	40.2%
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	13.3%	21.3%	18.6%	24.7%
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	16%	30.7%	21.6%	22.7%
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	21.3%	22.7%	16.5%	23.7%
C (Sections 3.1.2 and 3.2.3)	Noise (absolute)	200-1600 ms after noise onset	-	60%	6.7%	24.7%	28.9%
	Tone (absolute)	200-1500 ms after tone onset	-	54.7%	4%	32.0%	22.7%
D (Sections 3.1.1 and 3.2.2)	Noise (change)	200-1600 ms after noise onset	-1000-0 ms before noise onset	13.3%	14.7%	7.2%	38.1%
	Tone (change)	200-1500 ms after tone onset	-1000-0 ms before noise onset	9.3%	12%	11.3%	30.9%
E (Sections 3.1.4 and 3.2.5)	Baseline	-1000-0 ms before noise onset	-	58.7%	4%	30.1%	11.3%

Table index	Comparison	Time window		CS+ vs US		CS+ vs US	
		Main	Additional time window	CS+ > US	CS+ < US	CS+ > US	CS+ < US
F (Sections 3.1.4 and 3.2.5)	Baseline	-4300- -3300 ms before water delivery	-	28%	6.7%	12.4%	23.7%

Supplementary table 3. Results of the effect of the sound meaning and the presence of the unconditioned stimuli in individual units.

The green color indicates the significant differences observed in the population activities between the conditions.

Table index	Comparison	Time window		Average of Self- vs Ext- initiated conditions		Pairs of Self- vs Ext-initiated conditions				Conclusion
		Main	Additional time window		P-value	HighE	P-value	LowE	P-value	Significance
A (Section 3.4.1)	Noise onset (absolute)	0-100 ms after event	-	<	0.03	=	0.67	<	$6*10^{-3}$	ns
	Noise offset (absolute)	0-100 ms after event	-	=	0.86	=	0.75	=	0.31	ns
	Tone onset (absolute)	0-100 ms after event	-	=	0.54	=	0.57	=	0.23	ns
	Tone offset (absolute)	0-100 ms after event	-	<	$7*10^{-3}$	<	$4*10^{-3}$	=	0.12	ns
B (Section 3.4.1)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	>	$2*10^{-4}$	>	$4*10^{-5}$	=	0.46	ns
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	=	0.57	=	0.96	=	0.32	ns
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	>	$1*10^{-4}$	=	0.07	>	$2*10^{-4}$	ns
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	<	0.03	=	0.07	=	0.50	ns
C (Section 3.4.2)	Noise (absolute)	200-1600 ms after event	-	<	$1*10^{-3}$	<	0.02	<	0.01	*
	Tone (absolute)	-1400-0 ms before water delivery	-	=	0.60	>	0.02	=	0.03	ns
D (Section 3.4.2)	Noise (change)	200-1600 ms after event	-1700 – -700 before trial beginning	=	0.07	=	0.03	<	$4*10^{-8}$	ns
	Tone (change)	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	>	$7*10^{-10}$	>	$7*10^{-12}$	<	$1*10^{-3}$	ns
E (Section 3.4.4)	Baseline	-1700 – -700 before trial beginning	-	<	0.01	<	$2*10^{-6}$	=	0.94	ns
F (Section 3.4.5)	Before the noise onset	-500-0 before noise onset	-	<	$4*10^{-7}$	<	$1*10^{-8}$	<	$4*10^{-4}$	***
G (Section 3.4.7)	After water delivery	2000-3000 ms after water delivery	-	<	0.03	<	$2*10^{-3}$	=	0.36	ns

Supplementary table 4. Results of the effect of the sense of agency in the population activity.

The stars indicate the dependence and the minimal level of significance of the three comparisons: ns - $p > 0.05/2$, * - $p < 0.05/2$, ** - $p < 0.01$, *** - $p < 0.001$. The red color emphasizes the positive results in the population level.

Table index	Comparison	Time window		Self > Ext		Self < Ext	
		Main	Additional time window	Averaged	Averaged and pairs	Averaged	Averaged and pairs
A (Section 3.4.1)	Noise onset (absolute)	0-100 ms after event	-	7.22%	0%	10.56%	1.11
	Noise offset (absolute)	0-100 ms after event	-	6.11%	0%	5.56%	0.56%
	Tone onset (absolute)	0-100 ms after event	-	6.11%	0%	4.44%	0%
	Tone offset (absolute)	0-100 ms after event	-	2.78%	0%	10.56%	0%
B (Section 3.4.1)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	17.78%	3.89%	7.22%	1.67%
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	13.33%	3.33%	14.44%	0.57%
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	23.33%	5.56%	10%	1.67%
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	8.33%	5%	17.78%	1.67%
C (Section 3.4.2)	Noise (absolute)	200-1600 ms after event	-	11.67%	2.22%	26.67%	6.11%
	Tone (absolute)	-1400-0 ms before water delivery	-	24.44%	1.67%	16.67%	2.22%
D (Section 3.4.2)	Noise (change)	200-1600 ms after event	-1700 – -700 before trial beginning	13.33%	1.67%	26.11%	3.33%
	Tone (change)	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	25.56%	6.67%	20%	2.22%
E (Section 3.4.4)	Baseline	-1700 – -700 before trial beginning	-	7.78%	0%	17.22%	2.78%
F (Section 3.4.5)	Before the noise onset	-500-0 before noise onset	-	8.33%	0%	31.11%	6.67%
G (Section 3.4.7)	After water delivery	2000-3000 ms after water delivery	-	10%	1.11%	20%	2.22%

Supplementary table 5. Results of the effect of the sense of agency in the population in individual units.

The red color emphasizes the positive results in the population level.

Table index	Comparison	Time window		Average of High vs Low effort conditions		Pairs of High vs Low effort conditions				Conclusion
		Main	Additional time window		P-value	Self-	P-value	Ext-	P-value	Significance
A (Section 3.5.1)	Noise onset (absolute)	0-100 ms after event	-	=	0.42	=	0.04	=	0.47	ns
	Noise offset (absolute)	0-100 ms after event	-	=	0.69	=	0.38	=	0.47	ns
	Tone onset (absolute)	0-100 ms after event	-	>	0.03	=	0.32	=	0.11	ns
	Tone offset (absolute)	0-100 ms after event	-	=	0.14	=	0.25	=	0.73	ns
B (Section 3.5.1)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	=	0.07	=	0.65	<	$9*10^{-3}$	ns
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	=	0.64	=	0.91	=	0.22	ns
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	=	0.73	=	0.37	=	0.40	ns
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	<	0.04	<	$2*10^{-3}$	=	0.66	ns
C (Section 3.5.2)	Noise (absolute)	200-1600 ms after event	-	=	0.84	=	0.57	=	0.83	ns
	Tone (absolute)	-1400-0 ms before water delivery	-	>	$4*10^{-5}$	>	$6*10^{-9}$	>	0.01	*
D (Section 3.5.2)	Noise (change)	200-1600 ms after event	-1700 – -700 before trial beginning	<	$4*10^{-7}$	=	0.18	<	$7*10^{-12}$	ns
	Tone (change)	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	>	0.03	>	$6*10^{-10}$	<	0.01	ns
E (Section 3.5.4)	Baseline	-1700 – -700 before trial beginning	-	>	0.02	=	0.55	>	$1*10^{-7}$	ns
F (Section 3.5.7)	After water delivery	2000-3000 ms after water delivery	-	=	0.18	=	0.36	>	$5*10^{-3}$	ns

Supplementary table 6. Results of the effect of level of effort in the population activity.

The stars indicate the dependence and the minimal level of significance of the three comparisons: ns - $p > 0.05/2$, * - $p < 0.05/2$, ** - $p < 0.01$, *** - $p < 0.001$. The blue color emphasizes the positive results in the population level.

Table index	Comparison	Time window		High > Low		High < Low	
		Main	Additional time window	Averaged	Averaged and pairs	Averaged	Averaged and pairs
A (Section 3.5.1)	Noise onset (absolute)	0-100 ms after event	-	7.78%	1.11%	7.78%	0%
	Noise offset (absolute)	0-100 ms after event	-	7.78%	0.56%	9.44%	0.56%
	Tone onset (absolute)	0-100 ms after event	-	8.33%	0%	4.44%	0.56%
	Tone offset (absolute)	0-100 ms after event	-	11.11%	1.11%	7.78%	0.56%
B (Section 3.5.1)	Noise onset (change)	0-100 ms after event	-100-0 ms before event	13.33%	1.67%	16.11%	3.33%
	Noise offset (change)	0-100 ms after event	-100-0 ms before event	13.33%	2.77%	10.0%	2.78%
	Tone onset (change)	0-100 ms after event	-100-0 ms before event	17.22%	4.44%	11.11%	2.22%
	Tone offset (change)	0-100 ms after event	-100-0 ms before event	12.78%	1.67%	17.78%	3.33%
C (Section 3.5.2)	Noise (absolute)	200-1600 ms after event	-	16.67%	5.56%	17.22%	2.78%
	Tone (absolute)	-1400-0 ms before water delivery	-	37.78%	10.56%	7.78%	2.22%
D (Section 3.5.2)	Noise (change)	200-1600 ms after event	-1700 – -700 before trial beginning	10%	0%	30.56%	5.56%
	Tone (change)	-1400-0 ms before water delivery	-1700 – -700 before trial beginning	24.44%	7.78%	16.11%	3.33%
E (Section 3.5.4)	Baseline	-1700 – -700 before trial beginning	-	26.11%	7.22%	7.22%	1.67%
F (Section 3.5.7)	After water delivery	2000-3000 ms after water delivery	-	22.20%	7.22%	10.0%	1.11%

Supplementary table 7. Results of the effect of level of effort in the individual units.

The blue color emphasizes the positive results in the population level.

6.2. Abstract

The effect of engagement on the neuronal activity in the primary auditory cortex (A1) has been previously shown. However, the experimental designs of previous studies did not consider possible combined influences of separate forms of engagement. We proposed three forms of engagement, including the presence of unconditioned stimuli that may change the meaning of a sound, the sense of agency (subjective awareness of control of one's own volitional action), and the level of effort. The aim of the present study was to identify how these forms of engagement influence the neuronal activity in A1.

We first analyzed how the individual factors, such as the unconditioned stimuli alone or the acoustical stimuli alone, changed the activity in A1 in monkeys (*Macaca fascicularis*). We compared the effects of the individual factors with their pairing (the Pavlovian conditioning). The three passive conditions were presented to a group of monkeys that were previously trained to instrumental conditions (well-trained monkeys) and to a group without such training (low-trained monkeys). To separate the individual effects of engagement and to prevent the monkeys from learning to react to temporal cues, two types of the acoustical stimuli were used: "tone only" and "tone followed by noise". We reported only the results of the comparisons in the case of "tone followed by noise". To reveal the effects of the sense of agency and level of effort, we compared neuronal activity in A1 recorded while the well-trained monkeys performed three instrumental conditions and were presented one passive condition (the Pavlovian conditioning). Two of these four conditions required self-initiation of the acoustical stimuli with a grasp of a metal bar, whereas the other two conditions were initiated externally by a training computer. To reveal the effect of the level of effort, two of the four conditions required a high effort by requiring the detection of the pure tone, and the other two required a low effort, without such detection.

The effects of sound meaning and of unconditioned stimuli were based on 75 and 97 units recorded during a presentation of the three passive conditions to the low- and well-trained monkeys, respectively. We found that the presence of the unconditioned stimuli led to slow modulations in the neuronal activity throughout the entire trial in both groups of monkeys. The fact that the slow modulations were present in both groups indicated that the changes were not related to the previous experience of the monkeys. The acoustical stimuli without meaning led to higher responses to the noise onset in the well-trained monkeys.

The effects of agency and level of effort were based on 180 units recorded while the monkeys performed in the three instrumental conditions and were presented with one passive condition. We found that the slow modulations in the three instrumental conditions and the one passive condition were very similar. The neuronal activities differed only in short intervals within the time courses due to variations in the required movements in the conditions. The main effects of the sense of agency were a decrease in neuronal activity after the self-initiation and a smaller neuronal activity during the noise stimulation. The main effect of the level of effort was a higher neuronal activity after detection of the tone in the conditions requiring high effort.

These results were consistent with and expanded upon previous studies showing that the neurons in A1 are sensitive to engagement. Moreover, we have shown that different forms of engagement have different effects on the neuronal activity in A1.

6.3. Selbstständigkeitserklärung

Hiermit erkläre ich, dass die hier vorliegende Dissertation mit dem Thema

„**Effects of different forms of engagement on the neuronal activity in the monkey's primary auditory cortex** "

selbstständig verfasst, nicht schon als Dissertation verwendet und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, den 23.03.2018

Stanislava Knyazeva

6.5. List of scientific publications

Knyazeva S, Selezneva E, Gorkin A, Aggelopoulos NC, Brosch M. (2018). Neuronal correlates of auditory streaming in monkey auditory cortex for tone sequences without spectral differences. *Front. Integr. Neurosci*. 12:4. doi: 10.3389/fnint.2018.00004

Selezneva E, Deike S, Knyazeva S, Scheich H, Brechmann A, Brosch M. (2013). Rhythm sensitivity in macaque monkeys. *Front. Syst. Neurosci*. 7:49. doi: 10.3389/fnsys.2013.00049.

Knyazeva SI, Loginova NA, Loseva EV. (2012). Anxiety level and body weight changes in rats living in overpopulated cages. *Bull. Exp. Biol. Med*. 154(1):3-6.