Predictive coding of global sequence violations in the mouse auditory cortex

Sara Jamali

SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE Cerveau Cognition Comportement

Laboratoire de recherche: Institut de l'Audition, Institut Pasteur

T H È S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ SORBONNE UNIVERSITÉ

Discipline: Neuroscience

Présentée et soutenue par:

Sara JAMALI

le: 30 septembre 2022

Codage prédictif des violations globales de séquences dans le cortex auditif murin

Sous la direction de:

Brice Bathellier – Directeur de Recherche, CNRS

et co-encadré par Timo van Kerkoerle

Membres du jury:

Athena AKRAMI	Rapporteuse
Valérie EGO-STENGEL	Rapporteuse
Karim BENCHENANE	Examinateur
Bechir JARAYA	Examinateur
Brice Bathellier	Directeur de thèse

Acknowledgments

The work of this thesis would not have been possible without my supervisor, Brice Bathellier. Brice, thanks for trusting me and for giving me the chance to work with you and to discover this new field. Thanks for letting me express myself and for being there every time it was necessary. For being passionate and at the same time patient, optimistic and helpful. It was a pleasure to work with you for almost four years. I have learned a lot from you and your vision of the field and the many disciplines it involves. Thanks to Timo van Kerkoerle for co-supervising this work with passion, his great ideas, and his perfect knowledge of literature. For his encouragement, persistence in the face of problems, and his ability to always find a way to solve them. Thanks to Stanislas Dehaene for his helpful comments on this work.

I am grateful to all the members of the Bathellier lab for making the lab a living space. Anthony and Etienne for all the Ph.Tea moments and all the scientific and personal discussions. Sophie and Evan, for sharing their enthusiasm for science. Joanna, for her presence as a good friend and for her emotional support on not-so-easy days. Antonin, for the BAM times, and Simone, for his kindness. Elena, Mathilde, Juliette, Sebastian, Anton, Jacques, and the Unicog members, Marie bellet, Marie and Joachim and also all the NeuroPsi people, Thomas, Aurelie, Margaux, Henri, Sophie, and all the Institut de l'audition colleagues, Marta, Maia, Gabriela, Sedigheh, Assad, Sepideh. It will be a long list to name everyone but thank you all.

I've been able to enjoy this marathon thanks to very special friends in my life. Parvaneh, who introduced me to this field for the very first time. The way she communicates science inspired me a lot. Thanks for being there, helping, and listening to me. Baptiste, thanks for your sense of

humor, for putting up with me in the most stressful times, and for bringing fun and adventure into my life. And all other friends, Fafa, Farzaneh, Ali, Dania, Sara, Bita, Gaël, Hosseine, Hiba, Caroline, Mahsa, Sepideh, Florian, Thibaut, Aurelian, etc.

Last but not least, thanks to my parents, Pari joon, Hassan, and to my brother Amir Arsalan. Thank you so much for your love and your encouragement. My special thanks go to my dad, who has scared me a lot in the past few months but has been very inspiring in my life. Thanks for all our conversations until 4 am, for everything you taught me, from physics, and mathematics, and how to repair my bike, to how responsible I should be in life no matter what I am doing. Thanks for teaching me that nothing is impossible, to dream big, and to chase my dreams. With all these connections and love, you trusted me and let me fly away when I was 19. Thanks for all that, and thanks for so many other things that I can barely put into words.

I would like to thank my committee thesis members of the first and second years of Ph.D. Dr. Christophe pallier, Dr. Yves Boubenec, and Dr. Karim Benchenane for their helpful comments and ideas, and their encouragement.

Finally, I would like to thank the members of the jury who agreed to participate in my defense and to read this manuscript: Dr. Athena Akrami, Dr. Karim Benchenane, Dr. Valérie Ego-stengel, and Dr. Bechir Jaraya.

Table of content

List of figures

Acronyms

A1 : Primary auditory Cortex

AC : Auditory cortex

CN: Cochlear nucleus

DBS : Deep brain stimulation

DD : Deviant-detection

EEG : Electroencephalography

ERP : Event-related potential

EXC : Excitatory neurons

FiOS : Fiber Optic Service

fMRI : Functional magnetic resonance imaging

IC : Inferior colliculus

ISI : Inter-stimulus intervals

MCS : Minimally conscious

MEG : Magnetoencephalography

MGB : Medial geniculate body

MMN : Mismatch negativity

PFC : Prefrontal cortex

PV : Parvalbumin

SNR : Signal-to-noise ratio

SOM : Somatostatin

SSA : Stimulus-specific adaptation

SVM : Support vector machine

V1: Primary visual cortex

VIP : Vasoactive intestinal peptide

VS : Vegetative state

WN : White noise

Abstract

The ability to extract temporal regularities at different time scales in sensory inputs and to detect unexpected deviations from these regularities is a key cognitive ability. The classical auditory oddball paradigm shows that the brain responds to sequence violations at a local time scale, but such responses also occur under anesthesia and therefore seem pre-attentive. In contrast, recent studies in humans and monkeys suggest that when the violation concerns regularities occurring over longer time scales, responses to the violation appear only in conscious, attentive subjects. To investigate whether local and global sequence violation responses exist in the mouse, we recorded from layer 1 to 5 of the auditory cortex using two-photon calcium imaging while mice passively listened to repetitions of 1s-long sequences of five tones. The repeated short sequence contained either a single tone (AAAAA) or a local violation at its end (AAAAB). Purely global violations were generated by presenting occasionally the AAAAA sequence in a block where AAAAB is repeated. We found that a population of neurons in the auditory cortex specifically responds to such purely global violations at the end of the AAAAA sequence. Although sparse, this population contained enough information to predict violations on single trials. A larger fraction of neurons boosted their responses to combinations of local and global violations (AAAAB presented in an AAAAA block). These global responses were resistant to a wide increase of inter-sequence intervals (1.5 s - 30 s) showing that they depend on a long-term prediction system. However, global responses vanished during anesthesia, in line with the observation made in humans and monkeys that global violation signals are present only in awake subjects, conscious of the stimulus in humans. Moreover, we established that vasointestinal peptide (VIP) and parvalbumin (PV) positive neurons encode weakly or not at all global violation signals. However, VIP interneurons

displayed sequence termination responses that were not specific to the stimulus. Last, we explored the potential role of behavioral engagement in global violation coding in a new behavioral task. We found that our behavioral task had little effect on global violation signaling but decreased the salience of local violation signaling, potentially due to an interaction with reward prediction encoding.

These results establish that the mouse brain is able to detect global violations in sound sequences in a subgroup of auditory cortex neurons and pave the way for the study of circuit mechanisms underlying long-term temporal regularity detection.

Resumé

La capacité d'extraire des régularités temporelles à différentes échelles de temps dans les entrées sensorielles et de détecter des écarts inattendus par rapport à ces régularités est une capacité cognitive clé. Le paradigme classique du “oddball” auditif montre que le cerveau réagit aux violations de séquence à une échelle de temps locale, mais de telles réponses se produisent également sous anesthésie et semblent donc pré-attentives. En revanche, des études récentes chez l'homme et le singe suggèrent que lorsque la violation concerne des régularités se produisant sur des échelles de temps plus longues, les réponses à la violation n'apparaissent que chez des sujets conscients et attentifs. Pour déterminer s'il existe une réponse de violation de séquence locale et globale chez la souris, nous avons enregistré de la couche 1 à 5 du cortex auditif à l'aide d'une imagerie calcique à deux photons tandis que les souris écoutaient passivement des répétitions de séquences de 1 s de cinq tons. Les courtes séquences contenaient soit un seul son répété (AAAAA), soit une violation locale à la fin de la séquence (AAAAB). Des violations purement globales peuvent être générées en présentant occasionnellement la séquence AAAAA dans un bloc où AAAAB est répété. Nous avons constaté qu'une population de neurones dans le cortex auditif répond spécifiquement à ces violations purement globales à la fin de la séquence AAAAA. Bien que petite, cette population contenait suffisamment d'informations pour prédire les violations lors d'essais uniques. Une plus grande fraction de neurones répondent à des combinaisons de violations locales et globales (AAAAB présenté dans un bloc AAAAA). Ces réponses globales étaient résistantes à une large augmentation de l'intervalle inter-séquence (1,5 s - 30 s). Par ailleurs, les réponses globales disparaissent pendant l'anesthésie, conformément à l'observation faite chez l'homme et le singe que les signaux de violation globale ne sont présents que chez les sujets éveillés, conscients du stimulus chez l'homme. De plus, nous avons établi que les neurones positifs

au peptide vaso-intestinal (VIP) et à la parvalbumine (PV) codent faiblement ou pas du tout les signaux de violation globale. Cependant, les interneurones VIP ont montré des réponses spécifiques à la terminaison des séquences mais qui n'étaient pas spécifiques à la séquence elle-même. Enfin, nous avons exploré dans une nouvelle tâche comportementale le rôle potentiel de l'engagement comportemental dans le codage global des violations. Nous avons constaté que notre tâche comportementale avait peu d'effet sur la signalisation des violations globales, mais qu'elle diminuait la saillance de la signalisation des violation locale, potentiellement en raison d'une interaction avec l'encodage de prédiction de récompense.

Ces résultats établissent que le cerveau de la souris est capable de détecter des violations globales des séquences sonores dans un sous-groupe de neurones du cortex auditif et ouvre la voie à l'étude des mécanismes dans les circuits sous-jacents à la détection des régularités temporelles à long terme.

1. Introduction

Hearing is a powerful sense, as our ears are sensitive to sounds arriving from any direction whereas vision relies generally on a limited view angle. We can hear the world all around us from every direction, in the dark, when we are not attentive, or even when we are sleeping, which makes the auditory system an early warning system. The auditory system is scanning continuously surrounding sounds and is able to attract our attention to unpredicted events. Our auditory system is therefore organized to process auditory stimuli in an automatic, unconscious manner and at the same time, it can process auditory inputs that require detailed conscious processing.

Many of the sounds surrounding humans and animals have regularities. For example, biological sounds such as footsteps, animal calls, or vocalizations are repetitions of the same sounds or auditory motifs. This is even more common in human-engineered sounds such as the ticking of a clock or musical pieces. Estimating temporal regularities in sounds to predict the next event and detect when regularities are broken is an asset for survival.

Beyond the processing of sounds, predictions are at the core of cognition. It is thought that human subjects continuously run an internal model of the world which predicts what would happen next. Beyond this, humans can think about what would happen if they were to do something, without having really to perform the considered action. So if we predict all the time, the world around us, doesn't it interfere with how we see, feel, and hear the world?

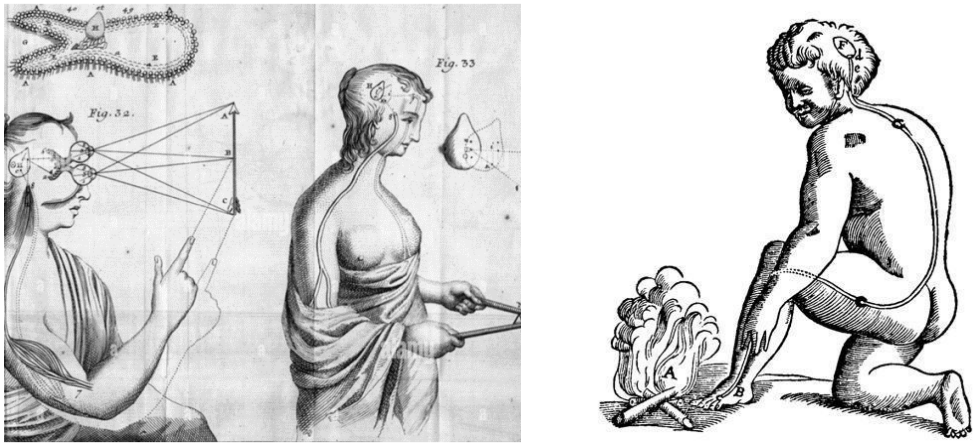


Figure 1.1 : Early descriptions by R. Descartes of the anatomy of three sensory systems*: vision on the left, touch in the middle, and pain on the right. Adapted from (Descartes, René 1664).*

Several centuries ago, the French philosopher René Descartes documented the idea that we have specific pathways for each specific stimulus that we are able to feel as human beings (Descartes, René 1664). He was among the first to describe pathways for touch, vision, and pain (Figure 1.1). Starting from these descriptions, simplified models of brain processing, proposed that a particular stimulus will always produce the same response. However, it rapidly became clear that our brain and even the brain of animals also predicts what is likely to happen next, and uses sensory information to either confirm some predictions or negate them (Helmholtz 1867; Gregory 1980; Neisser 1978). The expectations we have about the world consciously or subconsciously influence how we perceive the world. For example, the brain expects specific colors to be independent of the lighting context, a phenomenon known as color constancy (Foster 2011). Therefore, some intelligently crafted images of strawberries that are playing with background colors and that include no red pixels can give the impression that we see the red color whereas it is not present (Figure 1.2).

While early processing of sensory stimuli in the nervous system is now better described, understanding how the brain generates predictions about external stimuli remains one of the significant challenges of neuroscience. In my thesis, I focused on one particular type of prediction, which can be made from the recurrence of a particular temporal sequence of auditory stimuli, aiming to understand how the auditory cortex processes these temporal regularities at different time scales and how prior expectations influence auditory responses. To introduce my work, I will first briefly review the advantages provided by a predictive brain. Then I will shortly present the hierarchical predictive coding framework and the reasons why it is easier to study signals related to violations of predictions rather than pure prediction signals. Next, I will review the key protocols enabling us to study violations of predictions in humans and what we have learned from them. Finally, I will review some evidence of violation signals in animal models which will be key investigation tools for capturing the network and cellular mechanisms by which the brain computes predictions.

Figure 1.2 : There is no red pixel in this image. *Adapted from Akiyoshi Kitaoka, http://www.psy.ritsumei.ac.jp/~akitaoka/histogram_compression-ECVP2021-ShowTime.html*

1.1. Advantages of predictions

Sensory stimuli are noisy, embedded in more complex scenes, and develop over time. For interacting efficiently with the environment, we need to identify and estimate the spatial and temporal patterning of incoming events. Anticipation of the movement of a car is what allows us to avoid collisions. For an ambush predator, predicting the trajectory of the prey and the estimation of when to attack is crucial for survival (Wearmouth et al. 2014).

There are enormous advantages for the brain to do predictions, which include:

- gains in reaction time by anticipating events even before they reach our sensory receptors.
- gains in the efficiency of sensory inputs filtering, by using predictive priors, to extract information when it is hidden by noise or by other signals, or when it is missing.
- gains in the complexity of neural circuit architecture and the processing. If some inputs are perfectly predictable it is not necessarily needed to encode them completely. This could lead to a compression of the transferred data to what is not predicted. In the visual system, for example, several studies have suggested such compression happens already at the level of the retina (Srinivasan, Laughlin, and Dubs 1982; Schwartz et al. 2007; Schwartz and Berry 2008; Werner, Cook, and Passaglia 2008) which seems to transfer through the optical nerve mostly the action potentials that encode the inputs that cannot be compressed or predicted.

In general, the problem of perception is that sensory stimuli are ambiguous and this is not only a question of noise. Even without any noise, there are fundamental ambiguities. For example, in Figure 1.3, the sensation that we are receiving on our retina from a simple bar can come from

multiple different visual stimuli from different distances and with different tilt angles. Our perceptual system should select from an infinitive set of choices, the one that is more probable and more reasonable, and predictions about the structure of the external environment can be useful for that.

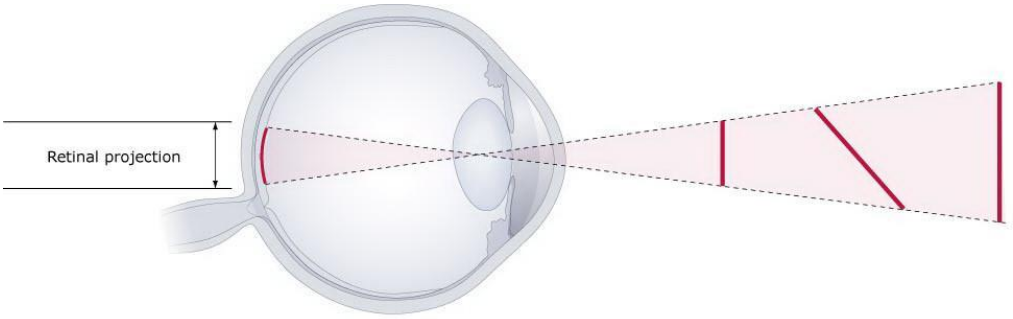


Figure 1.3 : Schematics describing the identical retinal projection of three lines placed at different distances of the eye and with different angles with respect to the optical axis of the eye. *These three lines generate the same projection on the retina and are therefore an ambiguous stimulus in the absence of context or expectation. Adapted from Dale Purves and colleagues, Visual illusions: An Empirical Explanation - Scholarpedia.*

Predictions are also important for reward evaluation. The seminal work of Wolfram Schultz has shown that dopamine neurons behave as a reward prediction system (Schultz 2016). If a conditioned stimulus predicts a reward, these neurons will fire to the conditioned stimulus even before the delivery of the reward. Then, when the reward arrives, if it is exactly the same as predicted, dopamine neurons do not fire. Only if the reward is different, they are firing, which signals the prediction errors. The reward prediction considers both the delay and the amount of received reward at different time scales (Tanaka et al. 2007).

1.2. The predictive coding framework

Predictive coding is a theory that aims to explain how the brain generates perception as a combination of sensory processing and prior knowledge. It is often formalized in the so-called Bayesian framework as a way to calculate a posterior distribution of the state of the world, based on actual input probability distributions and on *a priori* distributions (named 'priors') that represent internal predictions about the state of the world. The posterior distribution can then be used for decision-making, in order to choose one action.

In order to optimize decisions, there is, therefore, a necessity to evaluate the consequences of different actions based on posterior distributions. Moreover, because actions influence both the external world and its perception, each action solicits new observations from the environment. Hence predictions are nested into a perception-action loop (Ernst and Bülthoff 2004). Importantly also, predictions are relevant on very diverse time scales, from a few 10 ms for predicting trajectories of prey to seconds or days and years for more complex human behaviors. Time is thus an important dimension in the predictive coding framework, and different time scales may rely on different processes.

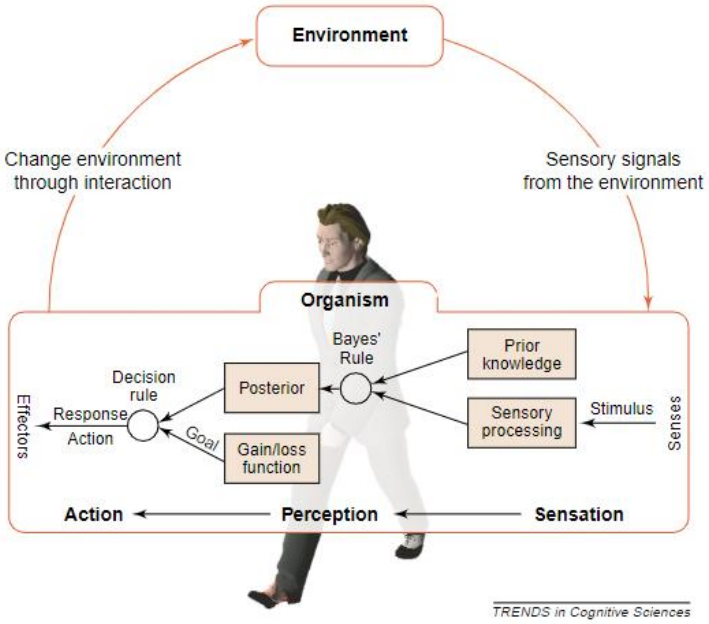


Figure 1.4 : The perception and action loop in the predictive coding framework. *Sensory inputs and prior knowledge allow evaluating the posterior probability of an object or event, based on which a behavioral decision can be made according to the current internal goal of the agent. Decision and action have an effect on the environment and on perception, thereby closing the loop. Adapted from (Ernst and Bülthoff 2004).*

Predictive coding was originally used in efficient transmission starting in the 1950s (Harrison 1952) with data compression algorithms. Algorithms based on this principle are the core of modern compression algorithms such as the ones used in the FiOS (Fiber Optic Service) network to stream video on TVs or in the cell network to bring video contents to the cell phone while minimizing the amount of data to transfer. These algorithms take advantage of the redundancy of the physical world (Figure 1.5) and our brain is likely able to use the same type of approach. The sensory inputs tend to be coherent spatially and temporally, for example, neighboring pixels in an image are likely to be similar across subsequent images as the presence of particular objects in the

field of view of the camera is likely to last for a certain amount of time. Because of these redundancies, encoding all the information at each time point is very inefficient. Therefore, the approach of video compression algorithms is to reduce the amount of data transferred or stored by encoding only the information that violates the simple predictions that can be made from previous images.

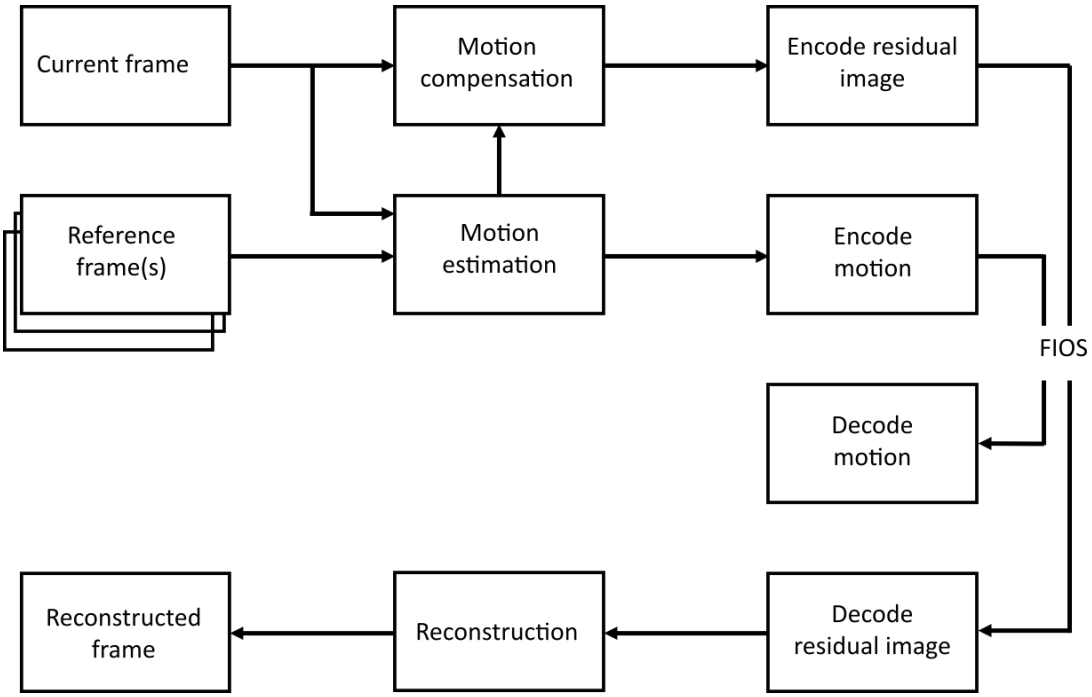


Figure 1.5 : Schematic representing the main computations performed by the FiOS video compression algorithm. *The main steps are an estimation of x-y frame motion which produces a predictable change in the pixel values that can be easily compressed and encodes only differences between the current and reference image after motion compensation. Because images in usual movies change relatively slowly with respect to the actual frame rate this allows efficient compression. According to (Lu et al. 2021).*

Similar ideas can be developed for sensory coding in biological systems. For example, the retina consists of different types of neurons, notably, the photoreceptors that are responsible for detecting light and turning it into electrical signals. Other neurons are called the retinal ganglion cells whose axons constitute the fibers of the optic nerve and transmit photoreceptor signals to our brain. We have about 100 million photoreceptors in each eye but only about 0.5 million retinal ganglion cells. The eye, therefore, likely performs compression of the information received by the receptor.

One argument for the need of compression in neurons is the transfer of information over long distances by means of action potentials that cost energy. Our brain uses around 20 % of our total body's energy, but this high energy budget amounts to only 0.1 spikes/sec/neuron on average (with large disparities across neurons). There is therefore a need to reduce the number of action potentials that encode a particular signal. This is one argument that supports the idea that the brain uses predictive coding to optimize the energy budget of information transmission (Srinivasan, Laughlin, and Dubs 1982).

Predictive coding is not only related to coding or energy efficiency issues. It is thought that the brain constantly makes probabilistic inferences. These inferences are accessible to children as young as a few months old (Teglas 2007 Xu and Garcia 2008). In general, the human brain has to implement advanced mechanisms of probabilistic reasoning as a part of its elementary, automatic, and unconscious operations. Predictive coding theories suggest that the brain constructs an internal model of the external world based on the sensory inputs and that this model allows building inferences to anticipate sensory inputs.

How could these inferences be implemented in the cortical circuit? To start addressing this question, Karl Friston and colleagues have suggested a model, in the framework of the gaussian predictive coding, that specifies the bidirectional exchange of information in different regions of the brain. In Figure 1.6, each system represents a cortical region made of columns that traverse the different architectonic layers of the cortex. In this model, Friston et al. interestingly suggested an interpretation for the role of some layers of the cortex. The first idea is that the pyramidal neurons of deep layers implement the hidden causes of the behavior and feedback to lower-level cortical areas. Each φ in each cortical region corresponds to a group of neurons that encode conditional expectations of causes at each level. The descending connections are necessary to implement predictive coding. The ascending connections which are coming from superficial layers of the cortex transfer the violations of predictions to higher-level cortical areas. For example, in Figure 1.6, the 2nd level causes (Φ_2), will predict the inputs of the 1st level. So in the 1st level sensory representation ξ_1 receives the predictions coming from Φ_2 and the sensory input (u). Then the subtraction of sensory inputs and predictions generates the prediction errors which are transmitted to the 2nd level. This operation is iterated across processing stages. At each node, the model optimizes Bayesian inferences. This optimization algorithm is a part of the broader "free energy minimization" principle suggested by Friston (Friston 2005; Bastos et al. 2012) which postulates that the brain minimizes the complexity of its state defined through a free energy function.

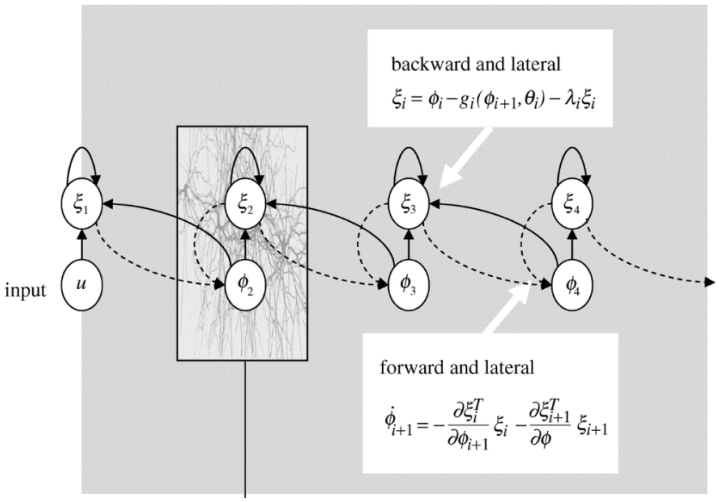


Figure 1.6 : Schematic summarizing the Friston model for predictive coding in the cortex. *Φ_i and ξ_i represent, at each encoding level, the prediction and the sensory surprise representation respectively. Adapted from (Friston 2005).*

In summary, predictive coding frameworks postulate that neurons in the brain, in particular in the cortex, encode violations of predictions rather than exact representations of the external inputs. The ascending signals are not only the sensory signals but the sensory signals from which what can be predicted at that level is subtracted. As a consequence, if there is a perfect prediction of sensory inputs, there is no ascending transmission, and representations equate to predictions. Therefore, the main distinctive feature of predictive coding models with respect to sensory processing models is the existence of signals encoding violations of predictions. This has two consequences for the study of predictive coding-like processing in the brain. First, it is important to identify the information that is likely predicted by the brain. Repeated patterns are usually employed for this purpose, based on the hypothesis that brain circuits will learn to predict the patterns through their repetition. Second, for a given prediction, the most straightforward way to reveal the existence of a prediction is to evidence signals that correspond to violations of predictions. Most studies on the subject have taken this approach using different protocols which I will briefly review in the next part.

1.3. Key protocols identifying prediction violations in humans

1.3.1. Mismatch negativity (MMN)

The first and most famous protocol which suggested an encoding of prediction violations in the brain is "Mismatch negativity" or MMN. First described by Näätänen, MMN is an event-related potential (ERP) captured during an oddball task typically in the auditory domain (Mäntysalo and Näätänen 1987; R. Näätänen and Alho 1997; Risto Näätänen 1990). When we listen to the repetition of a frequent sound, the presentation of a deviant sound which varies in pitch will evoke extra negativity in the electrical brain signals measured with electroencephalography (EEG). This response has a peak latency of approximately 100–250 ms and exhibits the strongest intensity in temporal and frontal areas of topographic scalp maps (Sams et al. 1985). In this protocol, the two tones X and Y are played to the subject in two different blocks, in which they have very different occurrence probability. In one block, X is presented very often while Y has only rare occurrences. In the other block, X is rare and Y is common. It is typically observed that the same tone played in a common and rare context leads to very different ERP response amplitudes. The difference in the ERP response to the same tone when it is common (standard tone) and when it is rare (deviant tone) is the additional negativity identified in the MMN protocol. Interestingly, this additional negativity is highest over the frontocentral areas. MMN is known to arise from pre-attentive processing as it is present in subjects non-attentive to the sounds (Alain, Woods, and Ogawa 1994; R. Näätänen et al. 1989, 197; R. Näätänen, Gaillard, and Mäntysalo 1978), and it is slightly attenuated in attentive subjects (Muller-Gass, Stelmack, and Campbell 2005). MMN response persists also partially under anesthesia (Koelsch et al. 2006;

Quaedflieg et al. 2014) and sleep (Nashida et al. 2000; Strauss et al. 2015). The deviant sound can differ not only in spectral-contents but also differ in duration, location, intensity and gap. The latency of MMN is inversely related to, and its amplitude is positively related to the magnitude of the difference between the standard and the deviant stimulus. MMN is sensitive to the inter-stimulus intervals (ISI) and disappears when the ISI is longer than a few seconds (Mäntysalo and Näätänen 1987). MMN can elicit not only with auditory stimuli but also in other senses such as vision (Kimura et al. 2009) and somatosensory sensation (Kekoni et al. 1997).

The origin of MMN signals is debated. The most common interpretation is that it is a comparison of the current stimulus and a memory trace that encodes a history of the played frequent stimuli in the temporo-prefrontal network (Doeller et al. 2003; Risto Näätänen 2003). Some studies indicate that one memory process that is key in MMN is a form of habituation called stimulus-specific adaptation (SSA), which means an activity-dependent decrease of neuronal or synaptic responsiveness. In this framework, the rare stimuli activate “fresh afferents” that are not adapted, and result in a bigger response than the frequent stimuli (Ulanovsky, Las, and Nelken 2003; May and Tiitinen 2010). In view of the fact that SSA reduces the neuronal responses to redundant stimuli, it can be considered as a preliminary predictive system (Fiorillo 2008) that reduces the amount of action potential and/or synaptic resources that are dedicated to a predictable event. Yet, this mechanism does not really constitute a prediction mechanism capable of predicting any kind of sensory input motif, that can be deduced from a learned or innate internal model of the environment.

There is in fact evidence that part of the MMN signal arises from an active predictive system, distinct from passive adaptation, in particular in frontal areas (Garrido et al. 2009;

Wacongne, Changeux, and Dehaene 2012). In order to support this evidence and to disambiguate active prediction signals from adaptation, it was proposed to search for more specific evidence of active predictions. Further sound sequence protocols were proposed for this purpose:

- First, a predicted stimulus should generate a violation response when the stimulus does not arise (omission). As a matter of fact, omissions also produce violation responses that cannot be explained by adaptation (Todorovic et al. 2011).
- MMN-like responses can be evoked when more complex predictable patterns are violated. For example, if the alternation "ABABAB" is the standard stimulus, a repetition of the same tone "AA" elicits a violation response (Horvath & Winkler, 2004, Summerfield, 2008), which cannot be explained with a simple habituation model because it is one of the repeated sounds that trigger the novelty response.

1.3.2. P300

Another higher-level novelty response detected in human ERPs is the P300. P300 was divided into two distinct late-positive components, P3a and P3b (N. K. Squires, Squires, and Hillyard 1975). These two components can be detected when subjects perform a deviant detection task with two deviant stimuli. One of the deviants is very similar to the frequent tone and the subjects are asked to detect this stimulus in a hard detection task. The other deviant is very different from the frequent stimulus and is used as a distractor. The first component, P3a is evoked by the distractor only, and lasts a short time over frontal regions (K. C. Squires et al. 1976). This component is not attention-dependent. The second component, P3b is more sustained, evoked by the target stimuli only, over the centro-posterior regions and it is attention dependent (N. K. Squires, Squires, and Hillyard 1975; Polich 2007; Dehaene and Changeux 2011). P3b component

in contrast to P3a component (Pegado et al. 2010) is not sensitive to the ISIs exceeding tens of seconds and it has been obtained even with an ISI of 10 min (Wetter, Polich, and Murphy 2004; Rohaut and Naccache 2017). P300 responses like MMN are often termed "novelty" or "prediction error signals" and are both evidence for prediction systems in the brain, matching one feature of the predictive coding model: the existence of signals for the violation of predictions.

1.3.3. Local-global

In the predictive coding approach suggested by Karl Friston, there is not only one predictive system but a hierarchy of them, in which we have serial cortical processing and each of these systems generate predictions. To test this hypothesis and to disambiguate whether the predictive or adaptation models can better explain the occurrence of the novelty responses in the brain, (Bekinschtein et al. 2009) proposed a paradigm called "local-global". In this paradigm, the authors aimed to study how the different regions in the brain respond to stimulus sequences with different levels of regularities, occurring at different time scales. Two sequences of 5 short tons were used: Either 5 identical tones (XXXXX) or four identical tones followed by a fifth different tone (XXXXY). These sequences are played in two different blocks of 125 sequences. In one of the blocks, the XXXXX is the common sequence, presented around 80 % of the times and the XXXXY is the rare sequence, presented at around 20 % of the times. This block is very similar to the MMN paradigms. The key distinctive feature of the local-global paradigm arises in the second block where the "XXXXY" sequence is the common sequence, presented at around 80% of the times, and the "XXXXX" sequence is the rare sequence, presented at around 20% of the times. So in this block, the sequence "XXXXY" becomes predictable at the global level and the presence of the repetition of five identical tones is a violation of this prediction. Therefore, there are two levels

at which there are regularities in this protocol, which can have distinct associated predictions and violations. First, at the local temporal scale, within the 5-tone sequence, the first four repeated tones "XXXX" create a regularity that can be violated by a "local change" with a "Y" tone in the "XXXXY" sequence. Second, when the 5-tone sequence differs from the repeated 5-tone sequence this constitutes a violation of an expectation that is based on a longer time scale regularity by a "global change" of the sequence structure.

Local-global paradigm has been extensively studied in humans (Bekinschtein et al. 2009; Wacongne et al. 2011; El Karoui et al. 2015; Nourski et al. 2018; Strauss et al. 2015; Faugeras et al. 2012; Basirat, Dehaene, and Dehaene-Lambertz 2014; Chennu et al. 2016; Recasens and Uhlhaas 2017) and non-human primates (L. Uhrig, Dehaene, and Jarraya 2014; Chao et al. 2018; L. Wang et al. 2015; Bellet et al. 2021; Lynn Uhrig et al. 2016; Tasserie et al. 2022) under different conditions and different imaging techniques with similar results. Bekinschtein et al. recorded the human brain activities with EEG and fMRI, and later (Wacongne et al. 2011) recorded with EEG and MEG simultaneously. In healthy attentive subjects, they detected two distinct potentials for the "local" and "global" violations. First, they observed an early response in the auditory cortex picked between 100 and 200 ms from the onset of the local deviant "Y" which is the novelty response known as MMN. This response persists even in the "XXXXY" block where the "Y" tone was predictable at the global level. However, this early response was stronger when the "XXXXY" sequence was presented rarely. This difference can be explained by a stronger adaptation when the Y tone is more frequently presented. From a predictive coding point of view, the transition probabilities from X to Y is higher in the block where "XXXXY" is common than in the block where "XXXXY" is rare. Hence, the transition is more surprising when "XXXXY" is rare. After this early response, there is also a response in a late time window over the auditory cortex

(Wacongne et al. 2011; El Karoui et al. 2015) and the precentral cortex (as for P3b response) to the global deviants in both blocks, both to rare XXXXY and rare XXXXX (Figure 7). The late response in the auditory cortex is suggested to be a feedback from precentral areas (Chao et al. 2018). Interestingly, the response to the global violations (P3b-like wave) almost disappeared when the subjects were not attentive to the sound in different conditions, i.e. mind-wandering, distracted with a visual task and during sleep (Strauss et al. 2015). This suggests that the subjects need to be aware of the global change to be able to detect the P3b. But the response to the local deviant (MMN) was not affected by these manipulations in awake, conscious subjects (Bekinschtein et al. 2009; Chennu et al. 2016) and was just reduced during sleep (Strauss et al. 2015). Applying the same paradigm to non-communicating vegetative (VS) or to minimally conscious (MCS) patients which have been asked to be attentive to the sound, one observes a detectable, but reduced global violation response in MCS patients and no response in VS patients (Bekinschtein et al. 2009). Under propofol anesthesia, electrocorticographic recordings in neurosurgical patients in four regions: core auditory cortex, non-core auditory cortex, auditory-related, and PFC, indicated that the local deviant effect (MMN) was absent outside of the auditory cortex under deep anesthesia and the global effect vanished totally in all regions. In the awake state both effects could be found in all regions (Nourski et al. 2018).

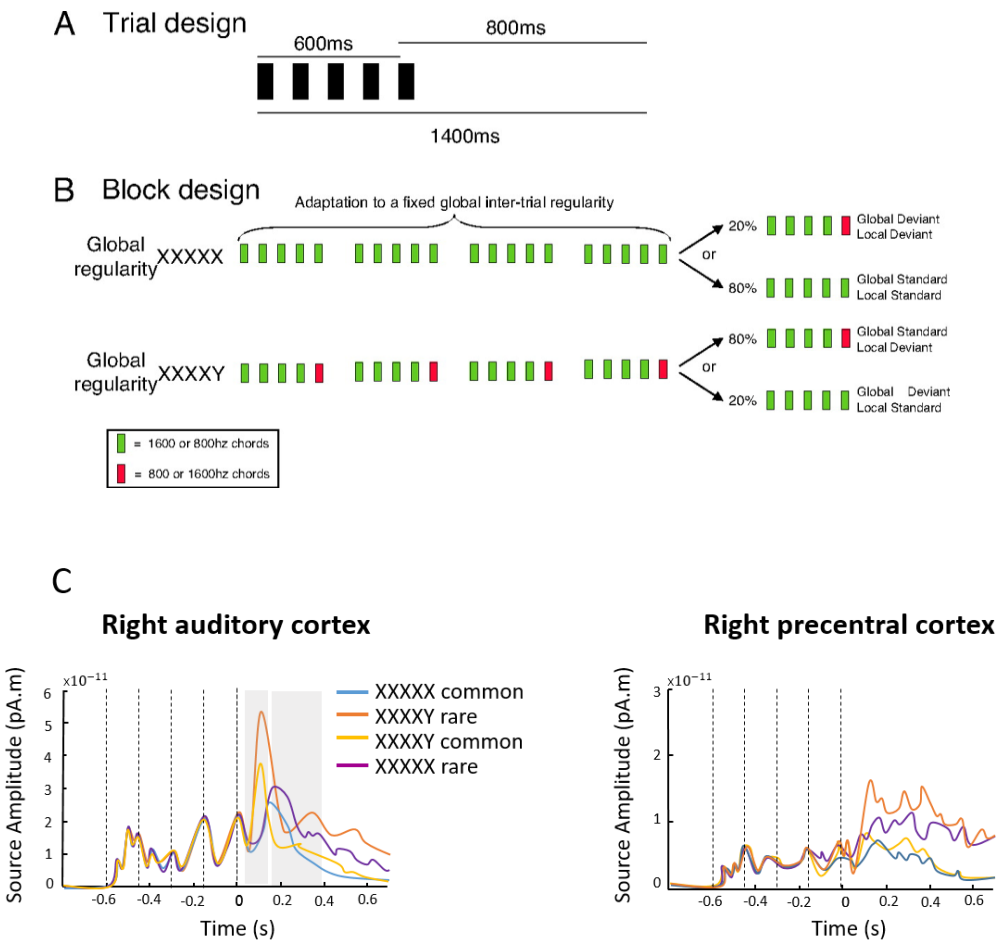


Figure 1.7 : Local and global paradigm. *Top. Design of the local and global paradigm:* ***A)*** *Tones X and Y are played in short sequences involving 5-tones.* ***B)*** *Each block consists of a series of these sequences where the regularity of the presentation is modulated at the local (sequence-level) and global (block-level) levels. XXXXY sequences contain a violation of the local regularity (within sequence time-scale). In blocks where XXXXY and XXXXX are commonly presented, XXXXX and XXXXY sequences contain a violation of global regularity respectively (across sequence time scale).* ***C)*** *Typical MEG responses in auditory and precentral cortex. Modified from (Wacongne et al. 2011; Bekinschtein et al. 2009).*

This suggests that the detection of global violations is a signature of conscious processing (Bekinschtein et al. 2009; Wacongne et al. 2011). In the end, the novelty responses that occurred with the omitted sequences were examined with repetitions of the first four tones "XXXX" presented rarely (10%) in both "XXXXY" and "XXXXX" blocks. These rare omitted sequences elicited an early response in both blocks which peaked at 100ms after the onset of the omitted sound with a topography similar to MMN. A late response was observed in the auditory cortex and also in the precentral cortex coinciding with the P3b signals as occurred with the global violations (Wacongne et al. 2011). Hence, one can hypothesize that the global violation responses observed in these studies correspond to the detection of deviance from the existent rule in the working memory.

1.4. Violation signals in animal models

1.4.1. Rodent

Visual system

As mentioned in section 1.1, several studies have suggested that the retina is a predictive system. In line with this idea, it was reported that the retina can recognize the complex pattern of fast stimulations (from 5Hz) and detect violations of predicted sequences. Previous work (Schwartz and Berry 2008; Schwartz et al. 2007; Werner, Cook, and Passaglia 2008) have presented a sequence of periodic dark flashes at 12Hz to the isolated retina of the mouse and the salamander and recorded spikes extracellularly. Many ganglion cells evoked strong responses to randomly omitted flashes. They observed a variety of firing profiles (Figure 1.8a). Some cells responded only at the beginning and at the end of sequences ("start-end"). Some cells had a

sustained activity during the sequences and also fired at the end to the omission ("sustained"), and some cells were responding only at the end of the sequence ("end-only"). (Werner, Cook, and Passaglia 2008) argued that the omitted flash activities might be only a rebound response to the end of the stimulus by recording the same cells during a long continuous flash (Figure 1.8b). But (Schwartz et al. 2007; Schwartz and Berry 2008) highlighted that the latency of the omitted responses depends on the frequency of previous flashes and also this omitted response is shifted when the last flash in the sequence arrives earlier or later (Figure 1.8c). So the omitted responses predicted the time when the omitted stimulus was supposed to be delivered. By using the light flashes instead of dark, they show that an individual ganglion cell fires selectively only at the end of one of the dark or the light flashes (to the omission) and not both. They also detected a violation response to the more complex pattern by changing the intensity and the interval between the flashes.

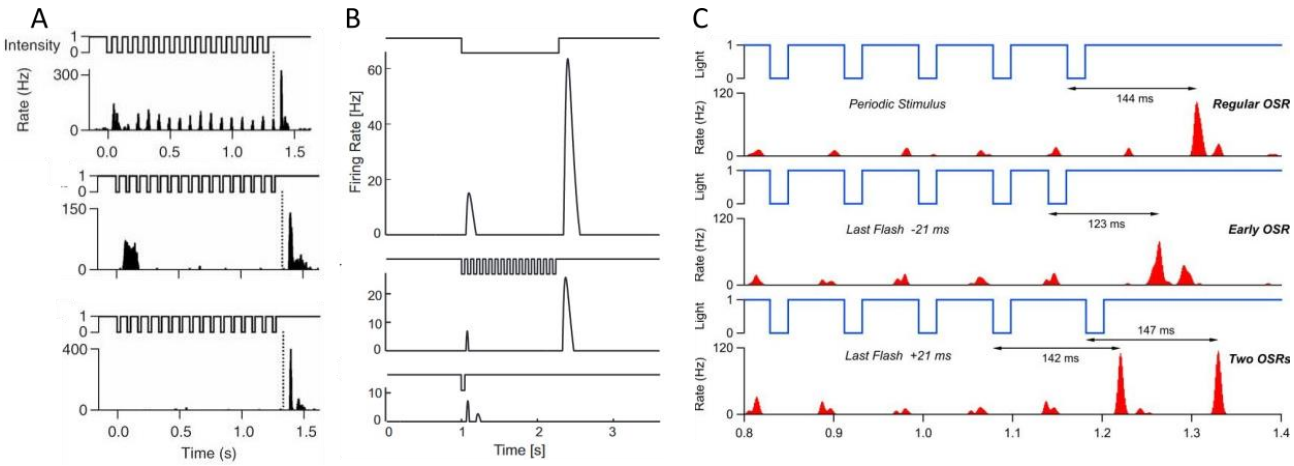


Figure 1.8: Ganglion cell responses to the end of the sequence or to omitted flashes. A) *Different patterns of firing rate recorded in the isolated salamander and mouse retinas during periodic dark flashes presentations recorded extracellularly with a multielectrode array.* ***B)*** *The same as -A- recorded during the presentation of a long sequence (top), a periodic short sequence (middle), and*

only one sequence (bottom) of dark flashes. ***C)*** *A single flash interval can update the prediction of a cell's timing to the omission response. Adapted from (Schwartz et al. 2007; Werner, Cook, and Passaglia 2008; Schwartz and Berry 2008).*

Even if fast periodic regularities can be detected as early as the retina and be probably amplified in the other subcortical regions the slower and more cognitive regularities are expected to arise in the neocortex.

Several studies show some evidence in favor of a predictive coding framework in rodent visual cortex during continuous visual stimulation, sometimes linked to locomotion. In particular, it was shown that some prediction signals at the neuronal level arise before the onset of the stimuli. The amplitude of this predictive signal (Fiser et al. 2016) is increasing with experience when the animals are exposed to the same sequence several times and can do predictions more confidently (Figure 1.9). The violation of these predictions with a rare stimulus or with an omission produces error signals (G. B. Keller, Bonhoeffer, and Hübener 2012; Gavornik and Bear 2014; Fiser et al. 2016; A. J. Keller, Roth, and Scanziani 2020). As proposed in hierarchical predictive coding, these studies observed the predictions (state units) and prediction errors (error units) are coded in distinct population representations. These studies also suggested that there are distinct populations codes for positive (unpredictable stimulus) and negative (omission) prediction errors (O'Toole, Oyibo, and Keller 2022; Attinger, Wang, and Keller 2017; Jordan and Keller 2020) and recently they demonstrated that these neurons have distinct transcriptional identities (O'Toole, Oyibo, and Keller 2022; Attinger, Wang, and Keller 2017; Jordan and Keller 2020). More interestingly, as the predictive coding hypothesis suggests, if the incoming stimulus matches the predicted signal, there

will be no new encoded information afterward. In contrast, if the incoming sensory stimulus differs from the prediction, there will be a higher activity corresponding to the prediction errors. Experimental observations show that when the prediction signal is high (more informative), the response to an expected stimulus (visual response) is lower and the response to a mismatched or omitted stimulus is higher when the predicted stimulus is violated (Fiser et al. 2016).

In this study (Fiser et al. 2016), mice were exposed to the "ABABA" sequences around 109 times while the mice are navigating a virtual tunnel. After training, a group of neurons responds in anticipation to the stimulus in an A- or B- selective manner (Figure 1.9a). These neurons were termed 'predictive neurons'. They could decode the identity of the gratings before they arrive. Therefore, these predictive neurons are stimulus-dependent. They also observed neurons that respond with a delay and which would correspond to visually driven neurons. When the last stimuli are replaced in different contexts (Figure 1.9b), the decoding accuracy varied with the level of confidence related to the prediction, the accuracy was higher if a stimulus was presented 100% of the time compared to when it is presented 90% of the time and replaced by another stimulus the 10% of the time. The same study also reported a group of neurons that had a strong response to the omission but were not responding much during the A or B stimuli. (Zmarz and Keller 2016) show also that the mismatch neurons have a receptive field and respond almost only to a mismatch happening in one particular position in the virtual tunnel.

Several studies in rodents indicate that e contextual information is received from the higher processing areas to the primary temporal cortices. They also demonstrate prediction related inhibitory modulations due to interneurons in the local circuits of the visual cortex (Fiser et al. 2016; Hamm et al. 2021; A. J. Keller, Roth, and Scanziani 2020; Leinweber et al. 2017). These

results are also in line with the hierarchical predictive coding model which postulates that the predictions originate in particular cortical areas and are exchanged through cortico-cortical projections. As these long-range cortico-cortical projections are perturbed during anesthesia, the predictions are canceled out or reduced (A. J. Keller, Roth, and Scanziani 2020; Polterovich, Jankowski, and Nelken 2018; Parras et al. 2017).

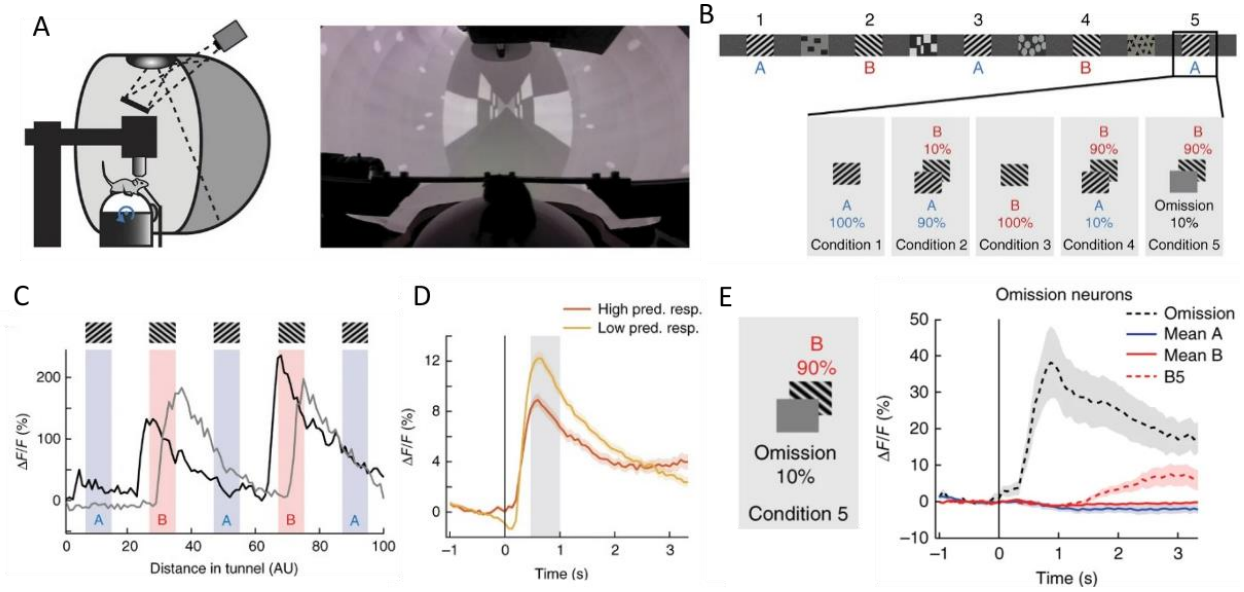


Figure 1.14 : Spatial sequence expectations in mouse visual cortex. A) *Mouse circulating in a virtual tunnel.* ***B)*** *Stimulus presentation in the tunnel on both walls. The stimulus A and B are shown when the animal reaches their positions 1-5 in the tunnel.* ***C)*** *Responses of two B-selective neurons in one traversal of the tunnel. A first neuron responds in anticipation to the B stimulus (black line, predictive neuron) and the other neuron responds after the presentation of the B (gray line, visual neuron).* ***D)*** *Low prediction events produce larger responses for the same stimulus.* ***E)*** *Omission responses in omission-specific neurons. Adapted from (Fiser et al. 2016).*

The mouse visual cortex (V1) neurons can learn sequences of oriented gratings and develop predictions that are highly informative about spatial, ordinal and temporal elements of familiar sequences. A study reported that responses to a familiar sequence of four visual grating orientations noted ABCD become significantly higher after the presentation of only 200 repetition of the sequence during training compared to a control group exposed to random permutations of the same sequence (DCBA, CDAB, etc) (Gavornik and Bear 2014). This result, however, is not in line with the predictive coding framework. If the second grating is omitted after a presentation of A stimulus (i.e. for an A-CD sequence) a very similar response to the B grating was detected as if the missing stimulus could be predicted. This effect seems to appear in the thalamo-recipient layer 4 first with a short latency. But if this omission happens after an E (E-CD) grating as there is no established prediction only a late bump is detected. These responses are not only affected by the serial order but also by timing. If during the learning phase the mice are exposed to a "long A-short B-long C-short D" their response in the test phase will be significantly higher to the learned sequence than a very similar novel "short A-long B- short C-long D" sequence. These information are not transferred between the eyes (Gavornik and Bear 2014), which suggests this learning happened in a region where the information from the two eyes is separable. The combination of orientation selectivity, the short latency omission responses in L4 and the eye selectivity suggests that the underlying plasticity occurs in the mouse visual cortex.

Auditory system

Signature of predictive coding has been also long studied in the auditory system, not only in humans with the mismatch negativity (MMN) and P300 protocols but also in animal models, with measurements of activity at single cell resolution. This resolution brought further evidence

for the phenomenon of stimulus-specific adaptation (SSA) which I mentioned in the paragraph dedicated to mismatch negativity. SSA is a reduction in the response to a sound when it is repeated at a particular time interval as compared to when it is presented more rarely. It arises in experiments that aim at creating expectations by repeating a stimulus and then using a violation of the regularity in order to demonstrate the presence of the expectation (Ulanovsky, Las, and Nelken 2003). In these experiments, one possible explanation for violation response is the passive SSA mechanism by which the less adapted oddball stimulus recruits more activity than the more adapted standard stimulus. For example, MMN protocols in human experiments are based on protocols that typically produce stimulus-specific adaptation (Winkler, Denham, and Nelken 2009). MMN-like paradigms are commonly studied in animals by playing two different stimuli in two blocks where one of the stimuli is common and the other is rare. Then, one measure to which extent the contextual effect of the block changes the response of the brain to the same stimulus. SSA has been studied along the auditory system. Its effect is widespread and robust in the auditory cortex (Ulanovsky, Las, and Nelken 2003; Taaseh, Yaron, and Nelken 2011; Hershenhoren et al. 2014) but can be found at several stations of the auditory pathway. SSA has not been observed in cochlear nuclei (Ayala and Malmierca 2012). It starts to be present in the non-lemniscal part of the "inferior colliculus (IC)" (Malmierca et al. 2009; Zhao et al. 2011; Duque and Malmierca 2014; Shen, Zhao, and Hong 2015), and also in the non-lemniscal part of the "medial geniculate body (MGB)" (Anderson, Christianson, and Linden 2009; Antunes et al. 2010). There is some evidence that SSA responses in the subcortical areas could be partly due to the feedback from the auditory cortex (Lesicko et al. 2022), in this recent study in contrast to the studies mentioned above the SSA effect could be found in the central and shell IC under cortico-collicular inactivation, suggesting that the repetition-suppression in IC is unaffected by the AC inputs and it is likely a fatigue of bottom-up

connections rather than predictive coding. However, the deviant enhancement was reduced or abolished indicating the involvement of the top-down connections in deviant detection (Lesicko et al. 2022). Therefore, in the fast lemniscal descending pathway from cochlear nuclei to the auditory cortex, SSA is weakly expressed before the auditory cortex, suggesting that potentially SSA is amplified in the auditory cortex through inhibitory networks (Park and Geffen 2020; Natan et al. 2015; Yarden, Mizrahi, and Nelken 2022; Hamm and Yuste 2016; Chen, Helmchen, and Lütcke 2015). SSA has been discovered for the first time in the lemniscal division of the auditory cortex (Ulanovsky, Las, and Nelken 2003; Ulanovsky et al. 2004). SSA has been explained with a simple adaptation model in two separate channels in several studies (Taaseh, Yaron, and Nelken 2011; Hershenhoren et al. 2014; Nelken 2014; Dhruv and Carandini 2014). Yet some protocols suggest the existence of complex mechanisms for SSA in the cortex which involve a combination of information from different frequency channels. For example, a recent study used, instead of pure tones, two wideband tone clouds with the same frequencies but with distinct temporal sequences of tones. A model of sound processing in which cortical neurons respond in a narrowly-tuned fashion to the instantaneous frequency of the sound predicts that adaptation in synapses or neurons will be the same for the two wideband clouds. Therefore, according to this model, there should be no SSA for these stimuli. However experimental measurements with such tone clouds demonstrate an early and strong violation response in the auditory cortex and a weak and late violation response in IC and MGB of anesthetized rat (Harpaz et al. 2021). This suggests that, at the cortical level but not in early subcortical responses, adaptation is specific to the full spectro-temporal information that defines perceptually distinct stimuli, potentially due to network phenomena (Harpaz et al. 2021). Alternatively, this experiment could reveal active predictions in the cortex that involve a memory of the full sound. Another study indicates that temporal features in sounds are subject to

SSA in the cortex (Awwad, Jankowski, and Nelken 2020). In this study, the oddball sound is a gap within a white noise. The authors show that the response to rare gaps is higher than the response to common gaps in the auditory cortex of anesthetized rat (Awwad, Jankowski, and Nelken 2020).

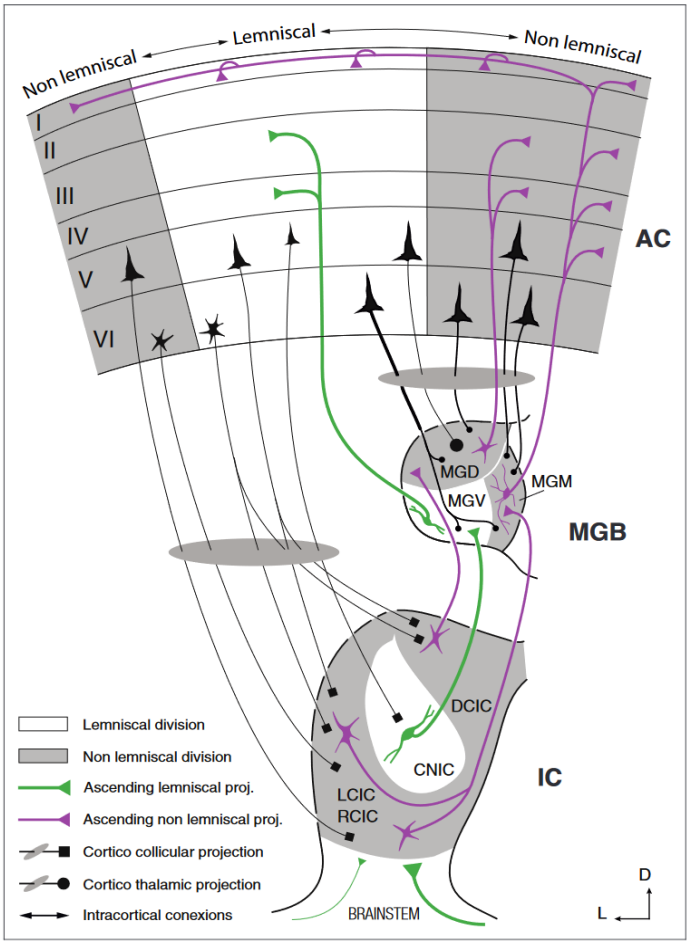


Figure 1.15 : Auditory pathway in rats demonstrating the major feedback and feedforward projections *between inferior colliculus (IC), medial geniculate body of the thalamus (MGB), and auditory cortex (AC). The lemniscal projections shape a straightforward, bottom-up pathway to the auditory cortex. In the lemniscal pathway, the deviant detentions could be observed strongly at the AC stage and not before. The non-lemniscal divisions receive information from the lemniscal core that they are enveloping and the top-down input from the AC and other non-lemniscal subcortical divisions. Adapted from (Malmierca, Anderson, and Antunes 2015).*

To distinguish predictions from adaptation, a useful control is the "many standard control". This consists in playing a tone with the same rarity as the SSA protocol but this time with many other rare tones as a context instead of the common tone used in SSA protocols. In this control, rare tones are not surprising anymore but adaptation should be identical. In the anesthetized animals, the response to the deviant does not differ from the control animals (Taaseh, Yaron, and Nelken 2011) in line with an adaptation model. In the awake animals, the response to the rare stimulus in a "common-rare" block is larger than the response to the same stimulus in a "many-standard control" block (Polterovich, Jankowski, and Nelken 2018) indicating that part of the response seen in the SSA protocol in cortex reflects active predictions while another part may reflect adaptation, as suggested for MMN in humans omission responses are another hint that a prediction system is at play during stimulus repetitions. Omission responses are absent in anesthetized animals but one can observe some omission responses in awake animals (Li et al. 2017; Audette, Zhou, and Schneider 2021).

Another interesting observation indicates that the oddball responses seen in SSA protocols go beyond simple adaptation phenomena. The importance of the sounds played during the SSA protocol can be changed for the animal by fear conditioning a tone. Strikingly, after conditioning this can cancel or reduce the contrast between the same tone presented commonly or rarely if the tone is associated with a foot shock (Yaron et al. 2020). The inverse effect has been found for the second tone not associated with the foot shock.

In summary, neurons in the auditory system are sensitive to contextual information. They are sensitive to a stimulus as a whole and not to its components. Cortical SSA is richer than subcortical SSA. Importantly, SSA and MMN are based on the same stimulation protocol. The

term SSA is used in animals in brain areas dedicated to sensory processing because the repeated sounds clearly produce adaptation in specific neurons. However, SSA and MMN responses naturally share many properties. One major difference is that the SSA phenomenon is rather focused on short latency responses while the responses observed in MMN have several components with distinct latencies. The core SSA effect seems to not be sensitive to NMDA antagonist receptors contrary to some parts of the MMN response (Koelsch et al. 2006; Quaedflieg et al. 2014; Nourski et al. 2018). The SSA effect can be reduced by changing the meaning of the sounds for animals with behavior.

Recently in a kind of similar visiomotor coupling experiment of mainly G.B Keller and colleagues, (Audette, Zhou, and Schneider 2021) conducts a movement-based coupling by sound task. Where the mouse should push a lever down to some position with an interval of longer than 200 ms between each push and receive a drop of water. They played a sound at a precise time at each movement of the mouse forelimb even those that don't result in a reward. This was called "closed-loop sound generating lever experiment". They detect predictive neurons in the auditory cortex L2/3 and L5 that make expectations about the frequency and timing of the expected sound and the prediction-error neurons that respond to the violation of this stimulus. These prediction-error neurons were not driven by sound and they were only responding to the violations.

1.4.2. Primate

To understand whether the primate brain can detect a violation at the global level, at a longer timescale in the sound sequences, brain activity was recorded in macaque monkeys during the "Local-global" paradigm used for humans with a few adaptations for primates. The local-global paradigm includes two orders of temporal regularities: the first-order (local) depends on the

pitch and the second-order (global) depends on the structure of the whole sequence. Using fMRI, ECOG, or recently Utah array, one can show that the primate brain is sensitive to both local and global violations. With fMRI, it was shown that the local effect activates the auditory cortex, thalamus, and striatum as in humans (Wacongne, Changeux, and Dehaene 2012), whereas the global effect activates multiple regions beyond the auditory pathway such as prefrontal, parietal and cingulate cortices (L. Uhrig, Dehaene, and Jarraya 2014). The global effect observed in primates was comparable to the P3b response found in human ERPs (Bekinschtein et al. 2009). As mentioned previously the global response appeared only in the attentive human subjects as demonstrated this effect vanishes when asking subjects to actively divert their attention from the tones (Bekinschtein et al. 2009). The monkey experiments were performed in passively listening animals. However, as the brain activations observed during the global violations formed a large "global workspace" including the prefrontal cortex and the interconnected regions, the authors of the study could speculate, by analogy to the human results, that the monkeys were to some extent attentive to the sound sequence and were somehow conscious of the sounds (L. Uhrig, Dehaene, and Jarraya 2014). To test whether it is the case (Lynn Uhrig et al. 2016) repeated the same experiment under two different anesthetics, propofol (GABAA-agonist) and ketamine (NMDA-antagonist). Anesthesia induces loss of consciousness by disrupting the long-distance corticocortical and corticothalamic network by canceling the feedback connections while the feedforward connections are preserved partially (Velly et al. 2007; U. Lee et al. 2013; Mashour 2014). Propofol and a high dose of ketamine abolished the local effect, but a lower dose of ketamine only reduced and shifted this local effect compared to awake animals. The global effect vanished completely during ketamine anesthesia. This result is in line with the human studies on the loss of consciousness under anesthesia (Nourski et al. 2018) in sleep (Strauss et al. 2015) or

coma and vegetative state (Faugeras et al. 2012). But surprisingly, propofol even at a deep level, disrupted only partially the effect of the global violation. These violation effects "local" and "global" could be restored by deep brain stimulation (DBS) of the central thalamus during propofol anesthetic which activated the same regions as in awake monkeys during the local and global violations. Using the psychophysiological interaction analysis during the stimulation, the authors reported an increase in functional correlations between the auditory cortex and frontal, parietal, cingulate, and a wide range of other cortices which could be in line with the gain of consciousness (Lynn Uhrig et al. 2016). The DBS of the ventral thalamus did not result in an increase in local and global effects during anesthesia even if it activated the frontal cortex but it did not impact the activity of the cingulate cortex. DBS in the central thalamus results in activation of the fronto-parieto-cingular network, which is thought to play an important role in conscious processing (Tasserie et al. 2022). (Chao et al. 2018) proposed a two-level hierarchical model using ECOG where the auditory cortex captures the local violations and then sends these prediction error signals to a higher level processing area (anterior temporal cortex) while the global violations are captured by prefrontal cortex (PFC) and then fed back to the auditory and anterior temporal cortex to update their predictions (Figure 1.11). So there are few pieces of evidence that the global effect is originating in the PFC in both humans and monkeys and then it's fed back to the auditory cortex for updating its predictions (Wacongne et al. 2011; Chao et al. 2018). In human and monkey PFC, the data is encoded in a more abstract manner. For example, the number of items contained in each sequence, the ordinal position of each item, and the local or global violations in the sequence, independent on the identity and timing of each item seem to be separately coded in neuronal populations of the PFC (Bellet et al. 2021; L. Wang et al. 2015). Sequential rules can be encoded in the monkey's prefrontal cortex during the planning to perform a four-movement sequential task

based on a visual guide that was presented to the monkey before movement execution. The sequences contained different combinations of push, pull or turn a handle, presented in three different categories: alternation (for example, push-pull-push-pull), paired alternation (for example, push-push-pull-pull), and repetition of an action (for example, pull-pull-pull-pull). In the action preparation time, distinct and large populations of neurons encoded the abstract representation of each of these three categories independent of the identity of the action or timing (Shima et al. 2007).

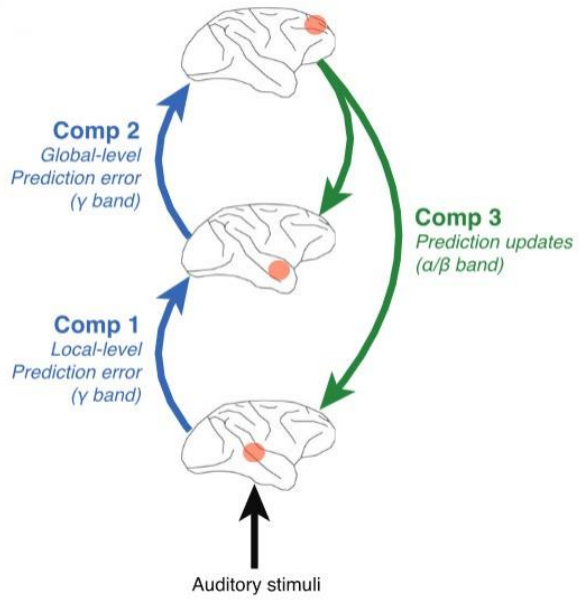


Figure 1.16 : A two hierarchical levels model in the primate brain in the "local-global" paradigm. *This model suggesting that the lower level (local) prediction errors are captured by the auditory cortex and are sent through feedforward connections to the higher level area (Comp 2), and then from Comp 2 to Comp 3. If the predictions created from Comp 3 by a feedback connection are not violated nothing will be transferred from Comp 2 to Comp 3. The higher level (global) predictions error in contrast originated in Comp 3 and updated the Comp 2 and*

Comp 1 predictions. The predictions and prediction errors are characterized by two different frequency band signals in humans and monkeys. This model is consistent with the observations of oscillatory activity in the brain using ECOG on macaque monkeys during the local-global paradigm. Adapted from (Chao et al. 2018).

1.5. Role of inhibitory modulations in violation detection

In the cortex, the three major interneuron types are parvalbumin-positive (PV), somatostatin (SOM), and vasoactive intestinal polypeptide (VIP). The VIP interneurons regulate primarily SOM but also PV interneurons (Pi et al. 2013), while SOM and PV interneurons inhibit directly the excitatory neurons (EXC) by targeting their apical dendrites and soma, respectively. PV and SOM inhibit each other (Figure 1.12A).

The ability of the cortical neurons to detect a violation is not only based on the current stimulus presentation but on a memory of the past stimulations. For example in SSA, the neuronal responses to the standard stimulus are reduced based on their history. This reduction in the firing rate can be explained either with an adaptation in the firing rate at the single neuron level and/or by changes in the excitatory-inhibitory networks, for example through changes in the equilibrium between a bottom-up (feedforward) sensory input and top-down (feedback) contextual modulation. The study of interneuron responses during violations of predictions is therefore potentially informative about the mechanisms underlying predictions in the cortex.

SOM, PV and VIP interneurons express SSA (Natan et al. 2015; Chen, Helmchen, and Lütcke 2015; Yarden, Mizrahi, and Nelken 2022). Several studies suggest that SOM but not PV

interneurons play an important role in deviance detection in vision (Hamm and Yuste 2016), audition (Natan et al. 2015), or in a visuomotor coupling experiment (Attinger, Wang, and Keller 2017). Optogenetic suppression of SOMs results in an increase in the firing rate of excitatory neurons to the standard stimuli and not to the rare stimuli whereas PVs increase the response to both standard and rare stimuli in awake animals (Natan et al. 2015; Natan, Rao, and Geffen 2017). A recent study shows that the suppression of PVs increases the excitatory firing rate, more to the deviant tone than to the standard tone in a SSA protocol (Yarden, Mizrahi, and Nelken 2022). However, this study was performed under anesthesia and could be different from the awake state. As in the anesthetized animals, the inhibitory responses in the auditory cortex are strongly reduced compared to the awake state (Figure 1.11D) (Kato, Gillet, and Isaacson 2015). CNO silencing of SOMs reduced both the difference between the rare and standard stimulus (SSA) and the deviant-detection (DD) response, which is defined as the difference between the response to the rare presentation of a rare stimulus in a SSA protocol and to the presentation of the same stimulus in a "many standard control" block. As mentioned above, DD is an MMN-like signal with longer latencies, which was measured in LFP recordings and in two-photon calcium imaging (Hamm and Yuste 2016). These MMN-like responses have been found both in excitatory and inhibitory neurons of the auditory cortex (Chen, Helmchen, and Lütcke 2015; Yarden, Mizrahi, and Nelken 2022).

Interestingly, interneurons are also involved in adjusting the valence of particular stimuli when they are repeated in different contexts. Excitatory and inhibitory responses show a long-lasting modulation in the L2/3 of the mouse auditory cortex, following passive exposure to an habituation protocol repeating the same tone over several days (200 trials/ day, 5 to 9 s duration) (Kato, Gillet, and Isaacson 2015). After this protocol, L2/3 the excitatory inputs decrease while

the inhibitory inputs increases over 5 days (Figure 1.12C, middle). These modulations are more balanced in L4 thalamo-recipient neurons with a small reduction in the excitatory and inhibitory modulations (Figure 1.12C, bottom). The effect seen in L2/3 neurons is therefore unlikely to be inherited from the subcortical regions. By studying the interneurons' contributions to this effect, the PVs and EXCs responses reduced to the same tone after 5 days of exposure though, the SOMs responses and so their inhibitory effect increased. This suggests that the SOMs play an important role in the increased experiment-dependent inhibition in L2/3 (Kato, Gillet, and Isaacson 2015). After the habituation, the authors involved the same tone in a simple behavioral task where the mice learn to lick at the offset of the same long (5-9 s) tones to get a reward. After this protocol, which changed the valence of the tone for the mouse, the SOMs responses decreased and the ECXs responses increased. This suggests that SOMs adaptation modulations are reversed between habituation to sound-driven behavior such that SOM activation is inversely correlated to the behavioral relevance of the stimulus. This process may be orthogonal to the signaling of violations, e.g. in MMN.

There exist also other signals in interneurons that seem unrelated to deviance detection. For example, under anesthesia, a late response component to standard, deviant and DD stimuli arrive around 30-150 ms in all interneuron types. This component is not an offset because the onset latency of this response is maintained for long tones. This late response was stronger for the standard than for the deviant (Yarden, Mizrahi, and Nelken 2022).

Suppression of VIPs does not significantly affect the firing rate of the excitatory neurons to the standard, rare, or DD even if a reduction in their response has been reported (Yarden, Mizrahi, and Nelken 2022). As VIPs modulate more strongly SOM and then PV directly or more

strongly indirectly via SOM modulation, a reduction in the response of fast-spiking cells has been observed in the standard and DD conditions but not the rare condition. This effect and also the fact that VIPs prefer the DD condition, together suggest that VIPs are the only interneurons that facilitate the true deviant-detection. By probably inhibiting SOMs during the DD condition, which in turn leads to less inhibition in the PVs and more inhibition in the excitatory cells, VIP activation may result in a EXC response that is larger in rare conditions than in DD conditions (Yarden, Mizrahi, and Nelken 2022). But in awake animals, it was also seen that the silencing of SOM reduced the DD effect (Hamm and Yuste 2016). Therefore, it is still unclear if the interplay between interneurons really plays a role in violation detection in the auditory cortex.

VIP interneurons are broadly associated with arousal, attention (Fu et al. 2014; Reimer et al. 2014), and behavioral states (K. V. Kuchibhotla et al. 2017; Garrett et al. 2020). They are believed to modulate the local circuits by receiving stronger long-range cortico-cortical projections from frontal cortices (Zhang et al. 2014; S. Lee et al. 2013; Wall et al. 2016) than the excitatory neurons. They also receive cholinergic inputs (Alitto and Dan 2012). VIP neurons have longer latency responses compared to other interneurons (Mesik et al. 2015). VIPs are controlling the contextual modulations in V1 by regulating the disinhibitory activation of SOMs in a figure-ground perception (A. J. Keller et al. 2020). Their responses to a familiar image set displayed thousands of times reduces as well as the EXCs. Moreover, VIP neurons in the visual cortex show a ramping activity to an omitted image in between the sequences of the repeated images in a change detection task whether the EXCs have a small increase. This ramping activity is stronger during a familiar image set (Figure 1.12E). These data suggest that VIP interneurons are good candidates for bringing some of the predictive information into the local circuits.

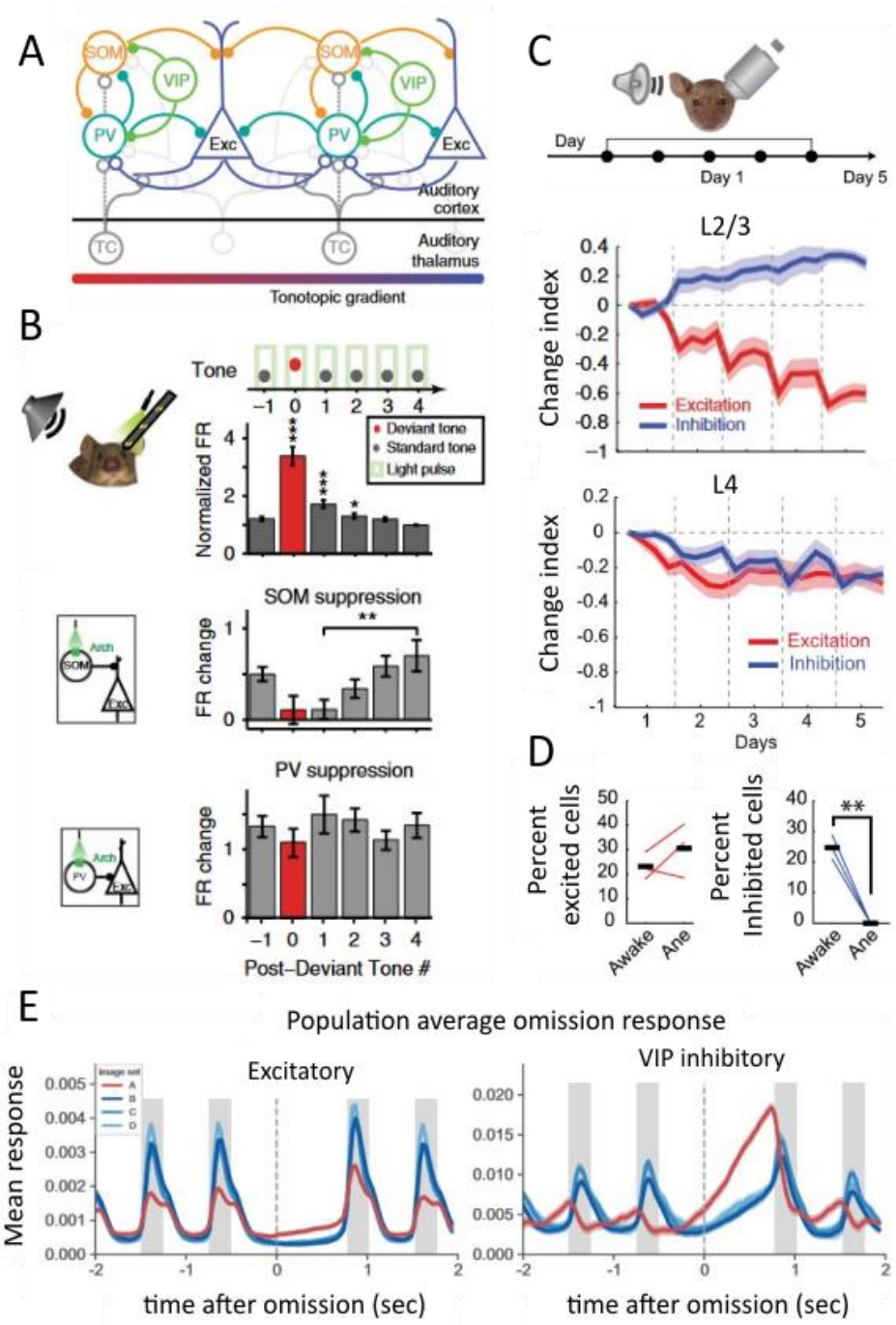


Figure 1.17 : Inhibitory contributions to deviant detection. A) *Schematic of an inhibitory-excitatory network in the auditory cortex.* ***B)*** *Responses of excitatory cells to a standard-deviant paradigm (top), suppression of somatostatin cells increases the firing rate to the standard stimulus with experience (middle), suppression of parvalbumin-positive cells increases the firing rate to all*

stimulus independent of the context (bottom) ***C)*** *Daily passive sound experience over 5 days increases the inhibition and decreases the excitation to this sound in the L2/3 mouse auditory cortex measured with two-photon calcium imaging (top), these modulations are small and balanced in L4 thalamorecipient cells (bottom).* ***D)*** *Fraction of cells with significant excitatory and inhibitory responses during awake and anesthetized conditions.* ***E)*** *A strong ramping response to an omitted stimulus in VIPs in the mouse visual cortex to a familiar image set in red, and to the novels image sets in blue (right), a small change detected in the activity of EXCs (left). Neuronal activity was recorded during a change detection task with a two-photon microscope, each image presented during 250 ms with an inter-stimulus interval of 500 ms. Adapted from (Blackwell and Geffen 2017; Natan, Rao, and Geffen 2017; Kato, Gillet, and Isaacson 2015; Garrett et al. 2020).*

1.6. Absence of evidence for more global predictions in mice

As we reviewed humans and primates, the auditory cortex is able to process auditory sequences over longer timescales and detect violations in predictable sequences. A recent paper, inspired by a change detections task (Barascud et al. 2016) in humans, reported that transitions between rhythmic and random temporal patterns in repeated short noise bursts, can be detected in the spike timing but not the rate of the mouse auditory cortex neurons after ~1s from the onset of the changed pattern. This change affected MGB weakly and it could not be detected in IC, even if the encoding of rapid short noise burst was more robust and with shorter latencies first in IC, then in MGB, and at the end in A1. This change detection became ineffective, even in the auditory cortex when longer and more complex rhythmic sequences were used (Asokan et al. 2021). So far the detection of more complex global regularities at longer time scales in the mouse auditory cortex remains unclear. In this Ph.D. by using the same "local-global" protocol as in humans and

monkeys, we aimed to show that the mouse auditory cortex detects the global regularities and catches violations of regularities that develop over long timescales and that cannot be explained with the adaptation phenomenon.

1.7. Techniques to study sequence representations in the mouse auditory cortex

1.7.1. Mouse model

In my PhD work; I used the mouse model because it is a good model for audition (Linden et al. 2003) which offers a vast number of transgenic lines and viral toolboxes. These tools make it possible to study different cell-types and manipulate them. As mice have a thin dura one can record the neuronal activity in the cortex using two-photon calcium imaging through the dura on the almost intact brain. The auditory cortex is positioned in the mouse such that it is accessible for chronic imaging through a glass window There is a huge amount of literature on deviance-detection in the mouse auditory cortex. The relatively complex behavioral abilities of mice combined with the fact that their auditory cortex receives inputs from both the thalamic and the frontal cortical areas, and the important role of their inhibitory local circuits in the deviant detection make together the mouse a powerful model for studying hierarchical predictive coding.

1.7.2. Two-photon calcium imaging

As violation signals for regularities occurring on longer time scales had not been reported yet in the mouse auditory cortex, we could speculate that these signals are sparse in the brain. It was, therefore, important to use a method to be able to record very large populations of neurons to

achieve sufficient sampling to obtain statistically representative populations of neurons conveying violation signals. With two-photon calcium imaging, we could record almost one thousand neurons in each recording with a neuronal resolution. In two-photon calcium imaging, we are able to measure indirectly the activity of the cells using GCaMP. GCaMP is a calcium-binding fluorescence protein. This fluorescence molecule gets excited by absorbing a photon with particular energy associated to its wavelength and emits light by getting back to its ground state. This molecule binds to the free calcium in the cells which stabilizes its fluorescence. When there is an action potential in the neuron, calcium ions flow into the neuron and we can see using calcium imaging an increase in the fluorescence signal. The two-photon microscopy allows us to image deep around 600 um by using a two-photon excitation instead of one. In two-photon excitation, each photon has roughly twice less energy (twice longer wavelength) than for one-photon, the excitation occurs only if two photons arrive at the same time (within a ~100 femtosecond time window) on the fluorescent molecule. The probability that this happens depends quadratically on light intensity and is therefore much higher at the focal point than elsewhere in the sample. This allows optical sectioning even in tick fluorescent tissue. This method decreases drastically out of focus fluorescence by avoiding the excitation of the molecules which are not in the focal point. The other advantage of two-photon microscopy is that longer wavelengths are less scattered in the tissue and the light, therefore, penetrates more deeply in the tissue. The two-photon microscopy, the mouse transgenic lines, and the viral toolboxes all together facilitate the study of specific cell-types in the mouse auditory cortex. The critical limitation of this method is the high decay time of GCaMP signals which are due to slow calcium dynamics in the neurons and slow dynamics of the indicator itself, even if there is some improvement in this field (Grødem et al. 2021). Two-photon calcium imaging is a scanning method, and therefore also suffers from relatively low frame rates

compared to electrophysiological measurements. In my PhD work, I was using a 31 Hz frame rate with the two-photon with a resonant scanning Femtonics microscope and a 19 Hz frame rate when using multiplane imaging with an acousto-optic microscope (Karthala).

1.7.3. Clustering

The large neuronal populations recorded with two-photon calcium imaging are too big to be visualized. In my PhD work, I used unsupervised clustering which is a dimensionality reduction method to simplify the exploration of large datasets. Clustering groups neurons that respond similarly across conditions. Several methods can be used to calculate the similarity between neurons based on their activity profile across sequences of sounds. Here I used Ward's method. Ward, is an agglomerative hierarchical clustering that instead of calculating a distance between samples, is based on variance analysis. In this method, if we have a sample size (neurons) of n, the algorithm starts with n clusters of size one and then groups together two samples, calculates the "sum of square" between the centroid of the formed cluster and each sample, the two samples with the minimum "sum of square" form the first cluster. Then this procedure is iterated to form either another two-samples cluster or a three-samples cluster and stops when all samples are combined into one big cluster. In the end, we are applying a threshold to the created dendrogram to choose the maximum number of clusters. The threshold is chosen to cover so that increasing the number of clusters would not increase the number of different response types in the data. Many clusters contained neurons with no visible response to sounds and were removed manually.

1.7.4. Population activity classification

In order to ensure that violation signals provide sufficient information to discriminate between rare and common sequences I used classifiers of single-trial population responses to

decode “common” vs “rare” conditions. There exists a large number of classification methods, but we choose linear “support vector machine (SVM)” for their efficiency and simplicity. SVM is a supervised machine learning algorithm that allows finding the maximal margin classifier for two linearly separable classes of data points in an N-dimensional space. A SVM classifier separates two categories with an optimal hyperplane. Optimality is defined as the largest distance e between the plane and the points of each class that are closest to the plane. These points are called support vectors. The hyperplane which can split the two categories according to this optimality criterion is obtained by a convex non-linear optimization procedure that maximizes the margins. Note that if the classes are not linearly separable, a penalty term is added for all misclassified points, which however preserves the convexity (i.e. existence of a single solution) of the optimization problem. A general issue with classifiers is overfitting. If a classifier is defined with a given set of points for two classes, it may not generalize properly to new data points corresponding to these classes and not be used for training because the optimal hyperplane was sensitive to outlier data points.

To avoid introducing biases coming from overfitting when quantifying the performance of a classifier fit is important to use a “cross-validation” method also called “rotation estimation” or “out-of-sample testing”. This consists in resampling the data by splitting it into a test and training set, training the classifier to find the best hyperplane on the training set, and then testing how accurately this hyperplane can classify the test set which was not used to estimate the model. This resampling procedure can be iterated until the whole data get tested.

2. Representations of global sequence violations in the mouse auditory cortex

This chapter is a full research article entitled “Representations of global sequence violations in the mouse auditory cortex” and prepared with S. Dehaene, T. Van Kerkoerle, and B. Bathellier, which describes the main results of my Ph.D. thesis. In this article, we describe evidence for the existence of a sparse population of auditory cortex neurons that encode global sequence violations. We demonstrate that these neurons are specific not only to the presence of a global violation but also to the specific stimulus and context that generates the violation. We also demonstrate that global violation signals are weak in two interneuron populations: VIP and PV interneurons. However, we show that VIP interneurons display interesting sequence termination responses that had not been described yet, at least to the best of our knowledge.

Representations of global sequence violations in the mouse auditory cortex

Sara Jamali[1], Stanislas Deheane[2], Timo Van Kerkoerle[2] and Brice Bathellier[1,*]

[1]Institut Pasteur, INSERM, Université Paris Cité, Institut de l'Audition, 63 rue de Charenton, F-75012 Paris, France.

[2]NeuroSpin, Gif-sur-Yvette, France.

[*]Corresponding author: brice.bathellier@cnrs.fr.

2.1. Abstract

The brain can detect violations of temporal regularities in incoming stimuli, an ability that could reflect the predictions it generates. This ability is often studied based on temporally local changes in the repetition of a stimulus, which can be at least partially explained by stimulus-specific adaptation mechanisms. Instead, violations of more complex global regularities produce specific responses, observed so far only in awake monkeys and humans, and whose detailed neural circuit mechanisms remain elusive. Here, we show that the mouse auditory cortex also represents global sequence regularity violations. The related sparse neuronal responses are sound- and violation-specific, disappear under anesthesia unlike local violation responses, resist long intervals between recurring sequences, and are weak in parvalbumin- and VIP-positive interneurons. VIP neurons were found to mainly implement a sequence termination code independent of sequence identity. These results establish that the mouse brain is able to detect long-term regularities and provide new insights into an underlying circuit logic that does not rely on adaptation.

2.2. Introduction

The ability to extract temporal regularities at different timescales in sensory inputs and detect unexpected deviations from these regularities is a key cognitive ability both in humans and animals (Dehaene et al. 2015). This ability was extensively studied with the classical auditory oddball paradigm that consists in rapidly repeating a single stimulus to generate an expectation. This expectation is violated with an occasional new stimulus whose identity is changed with respect to the repeated stimulus, revealing a range of violation-specific responses, including a neuronal population response to the oddball stimulus which is clearly larger than to the repeated reference stimulus (R. Näätänen, Gaillard, and Mäntysalo 1978) and which is often termed mismatch negativity. This change detection response was early recognized to stem from pre-attentive mechanisms, as it is maintained in non-attentive, sleeping (Nashida et al. 2000; Strauss et al. 2015) or anesthetized (Koelsch et al. 2006; Quaedflieg et al. 2014) subjects. It was also widely studied in animal models using oddball stimuli in repeated sequences (Carbajal and Malmierca 2018; Grimm, Escera, and Nelken 2016; Khouri and Nelken 2015) which established that the strong oddball response reflects stimulus-specific adaptation (Ulanovsky et al. 2004) to the reference stimulus while the processing of features related to the oddball stimulus remains unadapted. As adaptation processes are very common in the brain, stimulus-specific adaptation arises very early in the sensory system (Ayala et al. 2012) but is reinforced in the auditory cortex through circuit computations that involve interneurons (Hamm and Yuste 2016; Natan et al. 2015; Park and Geffen 2020; Yarden, Mizrahi, and Nelken 2022).

A key feature of the classical oddball paradigm is the simplicity of the underlying regularity which implies that violations correspond only to temporally local changes in the structure of the

sequence (Maheu, Dehaene, and Meyniel 2019). Thus this paradigm does not address the ability to build predictions about more elaborate temporal motifs that involve the recurrence of a particular sequence of several stimuli developing over a timescale larger than the interstimulus interval. However, predicting such long-term, global regularities is a requirement to understand natural series of events in the environment and adjust behavior accordingly.

To address this issue, paradigms based on more complex sequence structures were proposed. In particular, the local-global paradigm allows identifying responses to violations that do not involve a temporally local stimulus change (Bekinschtein et al. 2009). The local-global paradigm is based on the repetition of sequences of a fixed number of tones. Occasionally, oddball sequences are presented which break the prediction of the common sequence. If the common sequence has a stimulus change at the end (e.g. XXXXY), while the oddball sequence contains identical tones (e.g. XXXXX) then the violation of the prediction includes no local change and, nevertheless, specific violations response are observed in a widespread cortical network, involving temporal and prefrontal cortex of awake attentive humans (Bekinschtein et al. 2009; El Karoui et al. 2015; Wacongne et al. 2011), but also in monkeys (Chao et al. 2018; Lynn Uhrig et al. 2016). Importantly, unlike mismatch negativity signals, global violation responses vanish if the subject is anesthetized (Nourski et al. 2018; Tasserie et al. 2022; Lynn Uhrig et al. 2016), sleeping (Strauss et al. 2015) or is instructed to avoid focusing attention to the stimuli (Bekinschtein et al. 2009). Hence, global violation responses represent a clear signature of a predictive process that spans multiple events over time and occurs only during wakefulness.

So far however, the neural circuit computations leading to global violation responses remain elusive, potentially because large-scale, single neuron resolution data is missing to address

them. Also, it is not known if they exist in other mammalian species, and in particular in the mouse. Recently, several studies have indicated signatures of specific prediction for complex sequences in the visual system of passively attending (Gavornik and Bear 2014) and virtually navigating (Attinger, Wang, and Keller 2017; G. B. Keller and Mrsic-Flogel 2018; Leinweber et al. 2017) mice. The latter experiments indicated the involvement of inputs from associative cortical areas and of the local interneuron circuitry to generate predictions. However, it is difficult to identify in these studies whether violation responses stem from local or global changes. Moreover, it is unknown if global violation responses can be triggered in mice by sound sequences, despite the importance of temporal regularities in hearing and the existence of representations for sound sequence memory in mouse auditory cortex (Libby and Buschman 2021).

In this study, we recorded large populations of neurons in the mouse auditory cortex during the auditory local-global paradigm. We observed that a small, highly diluted subpopulation of neurons specifically responds to global violations. These responses are both stimulus- and violation-specific indicating that they encode the specificity of the mismatch rather than a generic mismatch signal. In addition, these responses were found to be resistant to long (>20s) inter-sequence intervals, indicating that they reflect robust memorization processes. Global violations disappeared under anesthesia as in monkeys and humans. They were hardly represented in Vaso-Intestinal Peptide (VIP) or parvalbumin (PV) positive interneurons, although VIP interneurons displayed coarser violation responses that signal sequence termination, independent of sequence types. Thus, together, our study shows that the mouse brain learns global regularities in auditory sequences from their passive observation, and uncovers a novel circuit logic in which violation responses depend on the stimulus which generated them.

2.3. Results

Global violation responses in the mouse auditory cortex

In order to investigate if global violation responses exist in the mouse cortex, we adapted the local global paradigm to awake, freely listening mice, which were head-fixed to allow for simultaneous two-photon calcium imaging of their auditory cortex (**Figure 2.1a**). With two distant pures tones, A (4 kHz) and B (12 kHz), we generated four different 5-tones sequences starting with 4 repetitions of the same tone and terminating with an identical (AAAAA, BBBBB) or a different tone (AAAAB, BBBBA). Tones within a sequence were 50 ms long, and their onsets were regularly spaced by a 237.5 ms time-interval. These 4 different sequences were used to construct 4 different blocks of 125 sequence presentations, interspaced by 1.5s of silence. In a block, the common sequence was first presented 25 times to create a global expectation, which was subsequently violated in 25% of the sequence when a rare sequence, differing from the common only for the last tone, occured (e.g. **Figure 2.1a**).

In order to control for sound identity effects, each mouse was presented with all four possible blocks (AAAAA, AAAAB, BBBBB, BBBBA), each repeated twice in a randomized fashion. Two-photon calcium imaging of layer 2/3 neurons in the auditory cortex was performed continuously during each block yielding fluorescence variations produced by sounds across several hundred GCAMP6s expressing neurons across large 1x1mm field-of-views (**Figure 2.1b,c**). In order to estimate firing rate modulations in the neurons, we used a simple linear deconvolution technique which is robust for long decay time constant indicators such as GCAMP6s (Deneux et al. 2019; 2016), and yielded sufficient temporal resolution to partially isolate individual tone responses in a given 5-tone sequence (**Figure 2.1c**). With this method, we collected responses to

the four different sequences in each of the two possible conditions (common or rare) for 8514 neurons across 12 sessions and 4 mice.

For simplicity, we first analyzed the responses obtained in the AAAAA- and AAAAB-dominated blocks, in which rare AAAAB and AAAAA sequences produce global violations with and without an associated local violation respectively. In order to explore these responses in a robust but hypothesis-free manner, we performed clustering of the dataset, using as a similarity metric, the correlation between trial-averaged response time courses for all 4 sequences and condition pairs (**Figure 2.1d**). We observed that the majority of clusters had tone identity-specific responses with pronounced adaptation. Some responded more to A tones and were rapidly adapted by the short-interval repeats of A (clusters #1-17, 19-20). Some responded more to the B tone of AAAAB sequences and displayed averaged responses that were weaker in the AAAAB-dominated block, in which the regular repetition of B also led to adaptation (clusters #19-21, 23, 25, 27-29). Both these responses appeared immediately with sound onset, consistent with the rapid rise of primary sound-evoked responses in the auditory cortex, which have a well-known susceptibility to stimulus-specific adaptation (Nelken 2014).

However, we also observed that 4 out of the 29 clusters shown in **Figure 2.1d** responded to rare sequences in a peculiar fashion. Two clusters of neurons displayed highly-elevated, delayed activity after rare AAAAA sequences (clusters #18 & 22), which was absent after common AAAAA sequences. This behavior is consistent with a signal representing the violation of the global expectation that sequences terminate with B.

For AAAAB sequences, we observed two types of behavior. (i) Many clusters showed responses that were time-locked to the B tone with little latency, and which responded more when AAAAB was rare (clusters #19-21, 23, 25, 27-29). We interpreted these clusters as neurons

responding to the B tone but experiencing adaptation when AAAAB is common. (ii) However, two other clusters responded very little after common AAAAB sequences, and had strong, temporally delayed responses after rare AAAABs (clusters #24 & 26). The large response latency after sequence offset, similar to the global violation responses after AAAAA sequences, suggested to us that they also corresponded to signaling of a global expectation violation.

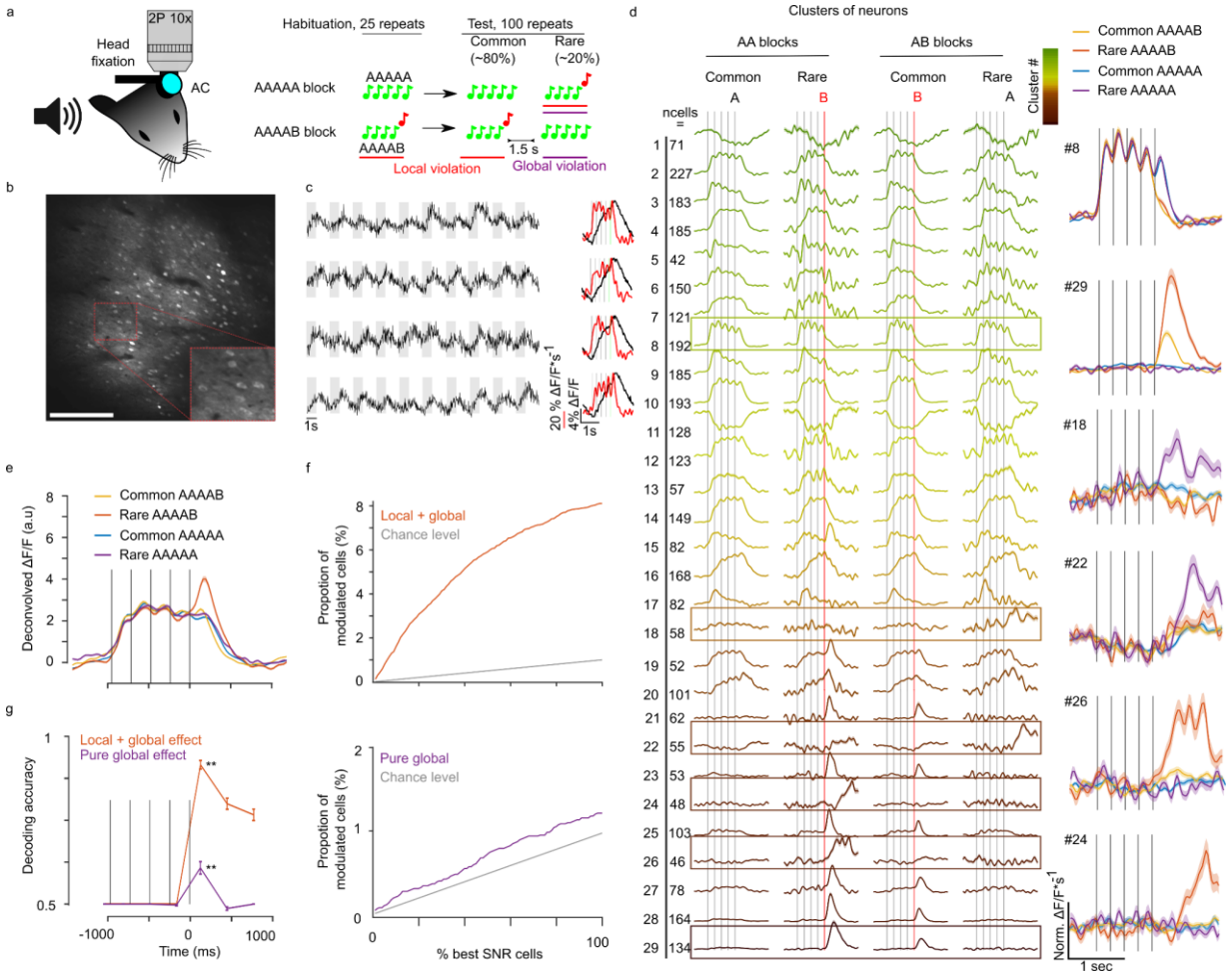


Figure 2.1 : Global sequence violation responses in the mouse auditory cortex. a. *Sketch of the experimental setup and of the two sound sequence blocks.* ***b.*** *Representative 1x1mm two-photon imaging field-of-view. The bottom right inset represents a magnification of the image*

within the red rectangle near the center of the image. ***c.*** *(left) Example raw fluorescence ΔF/F traces from 4 regions-of-interest (ROI = neuron). The gray rectangles represent short sequence presentations. (right) Raw (black) and temporally deconvolved (red) ΔF/F traces averaged across presentations of the same sequence for the 4 ROIs shown on the left.* ***d.*** *Mean population responses (deconvolved calcium signals) to AAAAB and AAAAA sequences depending on whether they are rare or common in the block.* ***e.*** *Mean responses to rare and common AAAAB and AAAAA sequences for 29 clusters of neurons are defined according to their response signature to 5-tones sequences in the AAAAA and AAAAB blocks. Responses of three clusters are aligned and magnified on the right to highlight typical sound responses (top) and those that are specific to global violations (middle) and (bottom).* ***f.*** *(top) Proportion of cells that are significantly responding more to rare AAAAB vs common AAAAB as a function of the fraction of cells selected based on their signal-to-noise ratio (SNR). (bottom) Same for rare AAAAA vs common AAAAA.* ***g.*** *Performance of a fully cross-validated classifier for predicting the rare sequence against the common sequence. The classifier is trained on time-averaged responses from 0 to 480 ms after sequence offset (mean local+global effect: 0.917± 0.012, mean pure global effect: 0.611± 0.017). P values were obtained as the location of the mean accuracy in the distribution of 100 shuffles; **: p=0.01.*

While the observation of these responses in clustered data suggested the existence of responses to global violations in the mouse cortex, the small number of neurons in these clusters raised the question whether these responses could have a significant impact at population level. We first plotted mean population activity for all four conditions (**Figure 2.1e**). Consistent with clustering results, the dominant population effect in rare sequences was the time-locked response

to tone B, which was more adapted when AAAAB is common (**Figure 2.1e**), consistent with the salience of stimulus-specific adaptation effects at population level (Nelken 2014). In contrast, no global violation response could be observed in the mean population firing rate (**Figure 2.1e**), consistent with the low number of cells implicated. Consistently also, the fraction of cells that had a significantly higher response when AAAAA was rare was low and close to the chance level of the statistical test even if run on cells pre-selected for high signal to noise ratio (**Figure 2.1f**). However, even if diluted within the auditory cortex, the few neurons that displayed global violation responses after AAAAA contained sufficient information to identify rare against common AAAAA sequences above chance level, as demonstrated with fully cross-validated classifiers (**Figure 2.1g**). Hence, global violation responses are a sparse but robust signal, which suggests that the mouse brain is able to identify second-order regularities in sound sequences over long time scales.

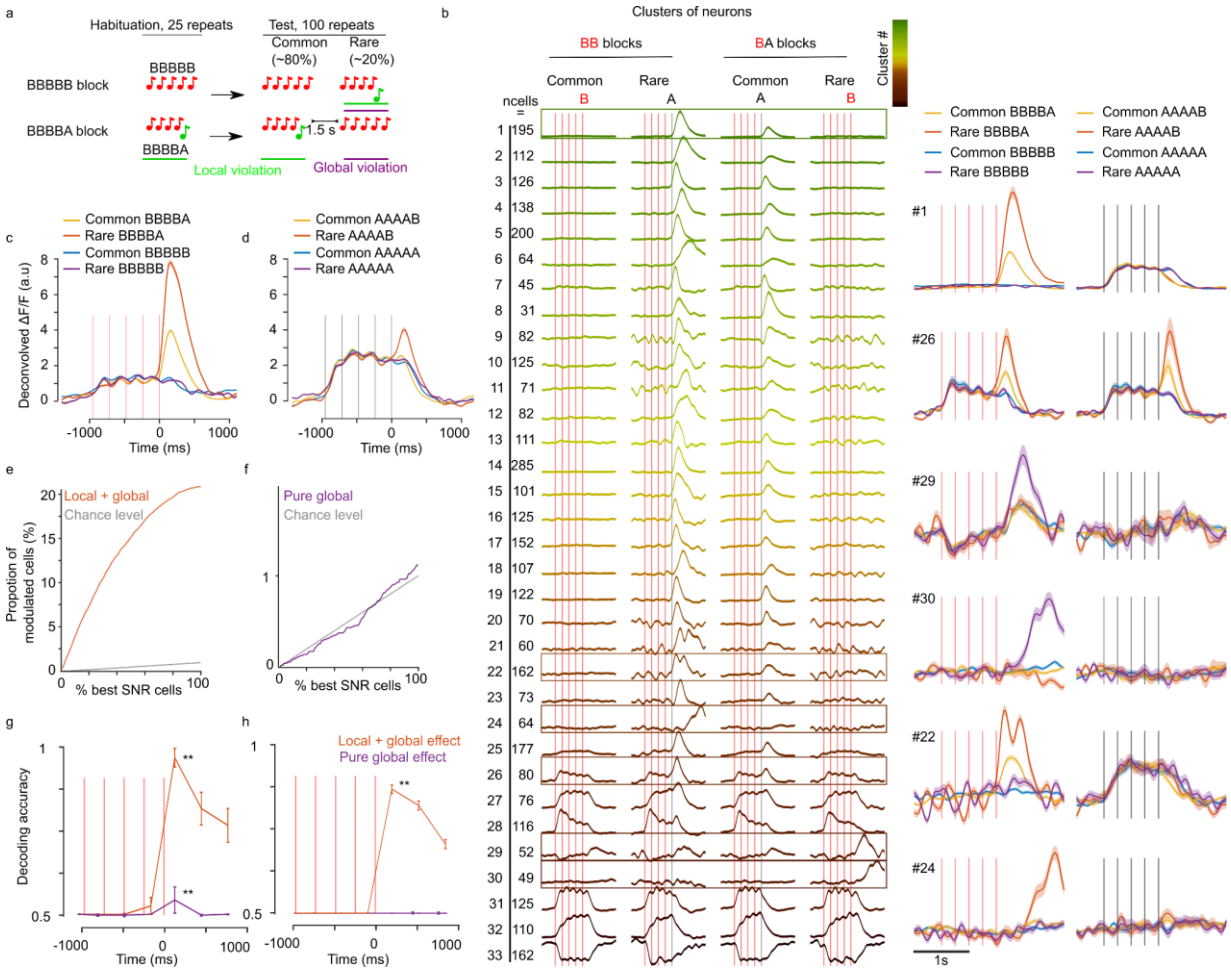


Figure 2.2 : Global sequence violation responses are sound-specific. a. *Sketch of BBBBA and BBBBB sound sequence blocks.* ***b.*** *Mean responses to rare and common BBBBA and BBBBB sequences for 33 clusters of neurons defined according to their response signature to all 5-tone sequences in the BBBBB and BBBBA blocks. Responses of selected clusters are aligned and magnified on the first column of plots to the right. On the second and rightmost column, the responses of the same cluster to the AAAAB and AAAAA blocks are shown.* ***c.*** *Mean population responses (deconvolved calcium signals) to BBBBA and BBBBB sequences depending on whether they are rare or common in the block.* ***d.*** *Same as c. for AAAAB and AAAAA blocks*

(reproduction of Figure 2.1d). ***e.*** *Proportion of cells that are significantly responding more to BBBBA rare vs BBBBA common as a function of the fraction of cells selected based on their signal-to-noise ratio (SNR).* ***f.*** *Same for BBBBB rare vs BBBBB common.* ***g.*** *Performance of a fully cross-validated classifier for predicting the rare BBBBA (respectively BBBBB) sequence against the common BBBBA (respectively BBBBB) sequence (mean local+global effect: 0.914± 0.015, p=0.01; mean pure global effect: 0.529± 0.017, p=0.01). P values were obtained as the location of the mean accuracy in the distribution of 100 shuffles.* ***h.*** *Performance of a fully cross-validated classifier predicting local+global and global-only violations independent of the identity of the sound producing the violation (i.e. both for AAAAA & BBBBB or AAAAB & BBBBA sequences; mean local+global effect: 0.890± 0.011, p=0.01; mean pure global effect: 0.494± 0.002, p=0.61). The classifiers in g-h are trained on time-averaged responses from 0 to 480 ms after sequence offset.*

Global violation responses are violation and context-specific

The existence of responses to rare AAAAB sequences with a delayed time-course similar to the pure global violation response suggested that global responses occurred also when the sequence triggering them also contains a local violation (**Figure 2.1c**). In this case, it was striking to observe that the neurons responding to rare AAAAA, and rare AAAAB with the delayed time course were different. This suggests that global violation responses are not only specific to the occurrence of a violation but also to the exact context in which the violation occurred. To further investigate this point, we analyzed data from the blocks in which the common sequences are BBBBB and BBBBA (**Figure 2.2a**). This analysis first indicated that the population of neurons sampled in this experiment globally preferred the A tone, as indicated by the larger number of

clusters responding to A than to B (**Figures 2.1d & 2.2b**), and the larger population responses to A (**Figures 2.1e & 2.2c-d**). However, irrespective of this asymmetry, clustering for BBBBB- and BBBBA-dominated blocks (**Figure 2.2b**) indicated that small groups of neurons again responded with a delayed time-course to rare BBBBA and rare BBBBB sequences. These neurons did not respond to earlier sounds in the sequence (**Figure 2.2b**) and had no visible impact on the population firing rate (**Figure 2.2c-d**) due to their sparseness within the auditory cortex (**Figure 2.2e-f**). Interestingly, within these groups, neurons responding to rare BBBBB and rare BBBBA were distinct (**Figure 2.2b**). Moreover, these clusters showed no responses to rare AAAAA or AAAAB sequences (**Figure 2.2b**) corroborating the idea that global violation responses are context specific and depend on the sequence that triggers them. In particular, the well-controlled global responses to rare AAAAA and rare BBBBB activated distinct groups of neurons. In line with this, rare BBBBB could be discriminated from common BBBBB sequences based on population classifiers after sequence offset (**Figure 2.2g**). However, one could not discriminate in general rare sequences from common sequences (**Figure 2.2h**) indicating again that the global violation signal is not a generic signal independent of the context in which it is generated. To further investigate this part, we also analyzed responses to omission sequences that were randomly interleaved in each block (15 rare sequences and 10 omissions within the 125 sequences of a block). As an unexpected termination of the sequence (Wacongne et al. 2011), omissions also produced global responses in a sparse set of neurons (**Suppl. Figure 2.1**). Responses to omissions were also specific with some overlap with the responses to other rare sequences in a few cases (**Suppl. Figure 2.1**).

Global violation responses persist over long intersequence intervals.

Having established the existence of global violation responses, we investigated whether they are linked to the temporally dense and periodic structure of the sequence presentation, or whether they represent a longer term, robust appreciation of the recurrence of common sequences. It was shown, for example, that stimulation-specific adaptation acts on multiple time scales, even beyond tens of seconds, but reduces strongly for long time intervals larger than few seconds (Nelken 2014; Ulanovsky et al. 2004). We, therefore, ran another version of our protocol in which we interleaved the sequences with a 25s period of silence which was interrupted by a short white noise burst in order to evaluate the resistance of global violation responses to out of context acoustic perturbations (**Figure 2.3a**).

To limit head-fixation duration, this experiment was done only for the AAAAB and AAAAA-dominated blocks. We observed that stimulus specific adaptation was still present (**Figure 2.3b-c**), but the fraction of cells preferring rare against common AAAAB sequences clearly decreased (compare **Figure 2.3d** and **Figure 2.1f**), consistent with the reduction of adaptation for long time intervals. Yet, we still observed clusters of neurons responding to global sequence violations, in particular after rare AAAAA sounds (**Figure 2.3b&d**). We also observed delayed responses that were fully specific of rare AAAAB sequences, which could be homologous to the global violation responses observed after rare AAAAA sequences (**Figure 2.3b**). Hence global violation responses in the mouse auditory cortex persist despite large changes in the long-term temporal structure of the regularities to be detected.

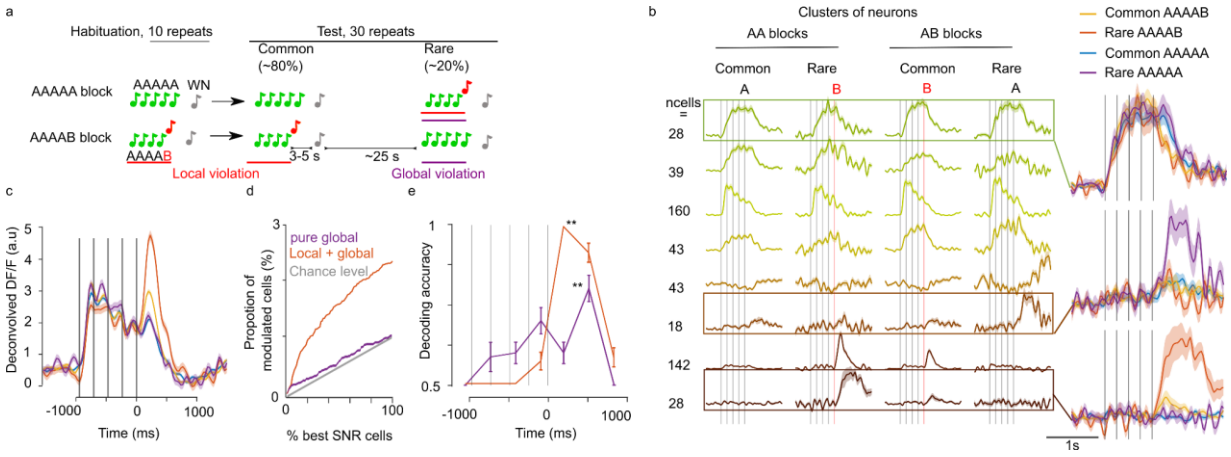


Figure 2.3 : Global sequence violation responses are resistant to long inter-sequence intervals.

a. *Sketch of AAAAB and AAAAA sound sequence blocks highlighting 25s inter-sequence intervals.* ***b.*** *Mean responses to rare and common AAAAB and AAAAA sequences for 8 clusters of neurons are defined according to the similarity of their response signature. Responses of selected clusters are aligned and magnified on the first column of plots to the right.* ***c.*** *Mean population responses (deconvolved calcium signals) to AAAAB and AAAAA sequences depending on whether they are rare or common in the block.* ***d.*** *Proportion of cells that are significantly responding more to AAAAB rare vs AAAAB common (orange) as a function of the fraction of cells selected for the based on their signal-to-noise ratio (SNR). Same for AAAAA rare vs AAAAA common (purple).* ***e.*** *Performance of a fully cross-validated classifier for predicting on single trials the rare AAAAB (respectively AAAAA) sequence against the common AAAAB (respectively AAAAA) sequence (mean local+global effect: 1.00 ± 0 ; mean pure global effect: 0.8± 0.04). P values were obtained as the location of the mean accuracy in the distribution of 100 shuffles; **: p=0.01). The classifier is trained on time-averaged responses from 0 to 480 ms after sequence offset.*

Global violation responses are sensitive to anesthesia

An important property of global violation responses in humans and monkeys is that they were selectively affected by anesthesia (Nourski et al. 2018; Lynn Uhrig et al. 2016). We therefore repeated our protocols, using a short time interval of 1.5s between sequences, but now following the same neurons across wakefulness and anesthesia (Filipchuk et al. 2022) (**Figure 2.4**). Mice were first presented with the sequences during awake passive listening (**Figure 2.4a**) and then anesthetized in the head fixation apparatus by approaching a nozzle for delivery of isoflurane (**Figure 2.4b**). After anesthesia induction, the isoflurane level was lowered to reach the minimal dose to obtain motionless narcosis (~1.3%) and the full sequence protocol was repeated. Thanks to the stability of the head fixation, neurons could be straightforwardly followed across the two states for which the same regions of interest were used. Anesthesia resulted in an overall decrease of population responses to sounds (**Figure 2.4c-e**). However, the increased population firing rate at the end of common and rare BBBBA sequences, and the larger response to rare BBBBA sequences, were preserved during anesthesia. This is in line with previous reports indicating that local violations or stimulus specific adaptation is preserved under anesthesia (Nourski et al. 2018; Lynn Uhrig et al. 2016). We quantified the presence of global responses through the accuracy of a classifier to discriminate between rare and common BBBBB sequences (**Figure 2.4d-f**). While in the awake state the classifier reached a performance of 0.648 ± 0.022, significantly above chance level, during anesthesia, performance dropped to chance level (0.5 ± 0.0). Therefore, unlike local violation responses, global violations responses vanish during anesthesia, in line with observations made in humans and primates.

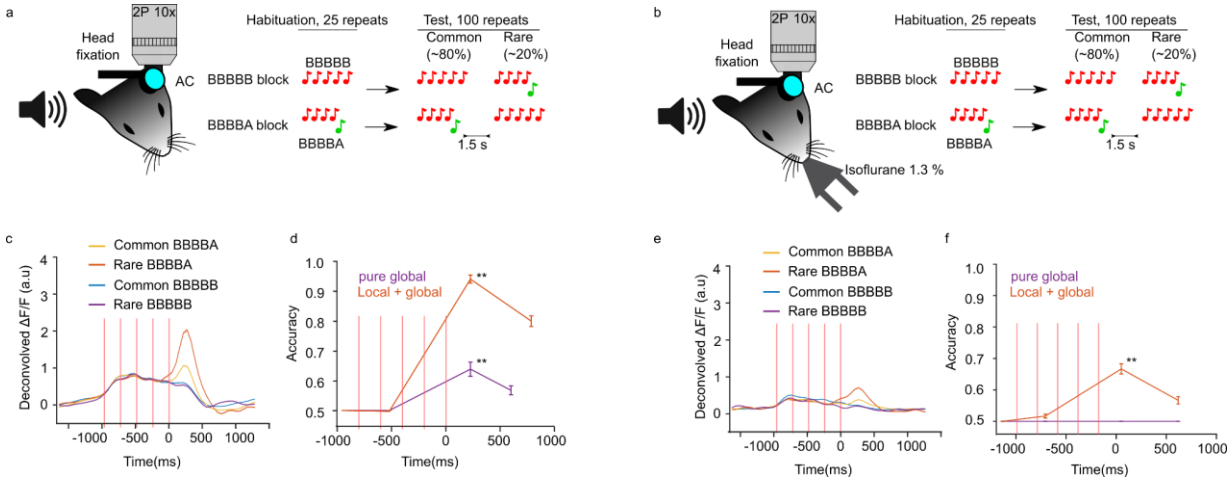


Figure 2.4 : Global sequence violation responses vanish under anesthesia. a. *Sketch awake two-photon imaging during BBBBA and BBBBB blocks.* ***b.*** *Same as a, during isoflurane anesthesia.* ***c.*** *Mean population responses (deconvolved calcium signals) to BBBBA and BBBBB sequences depend on whether they are rare or common in the block.* ***c.*** *Performance of a fully cross-validated classifier for predicting on single trials the rare BBBBA (respectively BBBBB) sequence against the common BBBBA (respectively BBBBB) sequence based on population activity (mean local+global effect: 0.933 ± 0.011, p=0.01; mean pure global effect: 0.648± 0.022, p=0.01). P values were obtained as the location of the mean accuracy in the distribution of 100 shuffles.* ***e-f*** *Same as c-d, for the same neurons imaged during isoflurane anesthesia (mean local+global effect: 0.680± 0.019, p=0.01; mean pure global effect: 0.5± 0, p=0.03). The classifier is trained on time-averaged responses from 0 to 0.5s after sequence offset.*

Global violation responses are weak in PV and VIP neurons

In order to better understand the circuit underpinnings of global violation responses, we investigated whether they occur in specific neuronal types in the auditory cortex. Earlier studies in

the mouse visual system have suggested that VIP interneurons specifically signal stimulus predictions, while PV- or somatostatin-positive (SOM) neurons are less involved (Garrett et al. 2020) but see also (Hamm and Yuste 2016). VIP interneurons are also considered major recipients of top-down inputs (X.-J. Wang and Yang 2018; Zhang et al. 2014). We, therefore, investigated representations of global violation responses in VIP and PV interneurons of the mouse auditory cortex. We injected floxed GCAMP6s AAV viruses in VIP-Cre or PV-Cre mice (**Figure 2.5a-b**) yielding specific expression in these interneuron subtypes (**Figure 2.5c-d**). Using clustering to explore VIP interneurons data, we observed broadly-tuned excitatory responses, broadly tuned inhibitory responses (some of them with excitatory rebounds) and specific responses to the B and A tones with clear adaptation patterns (**Figure 2.5e**). We also observed a few sequence termination responses, characterized by a systematic firing rate elevation after the end of all sequences (**Figure 2.5e**). However, we observed only very few responses to rare AAAAA sequences or BBBBB sequences (**Figure 2.5e**). Consistent with these observations, classification of rare against common AAAAB sequences yielded accuracy levels above chance in both awake and anesthetized animals (0.643 ± 0.021 for awake and 0.564 ± 0.014 under anesthesia, **Figure 2.5f & h**). However, classification of rare against common AAAAA sequences yielded chance level accuracies (0.486 ± 0.006 for awake and 0.506 ± 0.011 under anesthesia, **Figure 2.5f & h**) indicating that VIP-interneurons do not robustly represent global violation responses. We performed the same classifier analysis for PV interneurons population activity and also found that they have distinct representations of rare and common sequences when those have a local violation at the end (0.676 ± 0.022 for awake and 0.525 ± 0.012 under anesthesia, **Figure 2.5g & i**), a property that likely reflects stimulus specific adaptation. Yet, PV population activity did not contain sufficient information for classification of rare and common sequences without local violations (0.503 ±

0.013 for awake and 0.484 ± 0.005 under anesthesia, **Figure 2.5g & i**). Hence, PV interneurons like VIP positive interneurons do not robustly represent global violation responses.

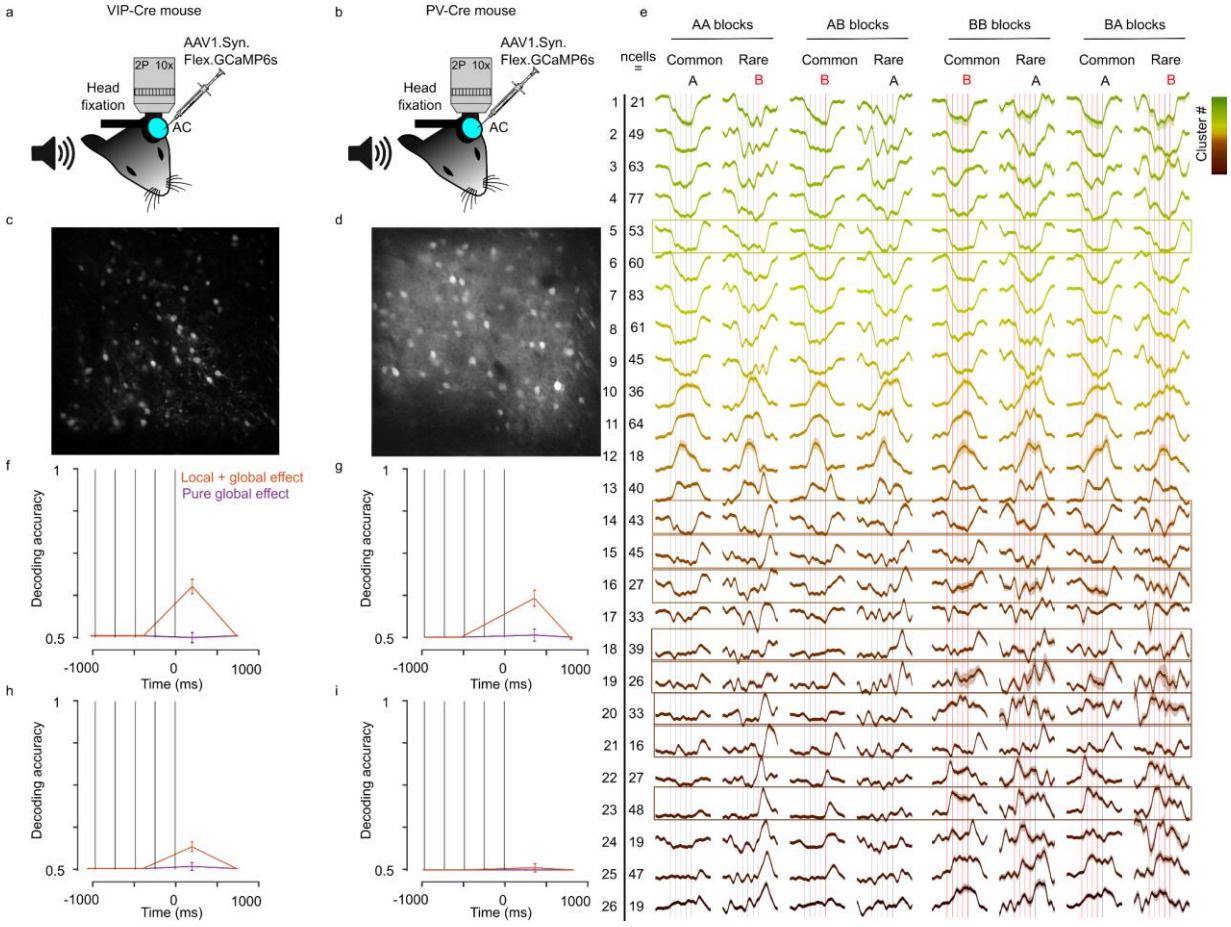


Figure 2.5 : Global sequence violation responses are weak in PV and VIP neurons. ***a-b.*** *Sketch of imaging experiments in mice expressing GCAMP6s in VIP or PV interneurons.* ***c.*** *Typical field-of-view showing GCAMP6s-labeled VIP positive neurons.* ***d.*** *Same as c. for parvalbumin-positive neurons in PV-Cre mice.* ***e.*** *Mean responses to rare and common AAAAB and AAAAA sequences for the clusters of VIP positive neurons are defined according to the similarity of their response signature. Black and red vertical lines indicate the timing of A and*

B sounds respectively. ***f.*** *Performance of a fully cross-validated classifier for predicting on single trials the rare AAAAB (respectively AAAAA) sequence against the common AAAAB (respectively AAAAA) sequence based on the activity of the recorded population of VIP positive neurons. The classifier is trained on time-averaged responses from 0 to 0.5s after sequence offset. Vertical lines indicate the timing of the sounds in the sequence (mean local+global effect: 0.676 ± 0.023, p=0.01; mean pure global effect: 0.486 ± 0.006; p=0.74). P values were obtained as the location of the mean accuracy in the distribution of 100 shuffles.* ***g.*** *Same as f. but for PV positive neurons (mean local+global effect: 0.643± 0.022, p=0.01; mean pure global effect:0.503 ± 0.013, p=0.35).* ***h-i.*** *Same as f-g during isoflurane anesthesia (mean local+global effect: 0.564 ± 0.014; mean pure global effect: 0.506 ± 0.011 for VIP interneurons and mean local+global effect: 0.525 ± 0.012; mean pure global effect: 0.484 ± 0.005 for PV interneurons).*

Subpopulations of VIP neurons signal sequence terminations

Clustering of VIP neurons activity however revealed interesting response motifs (**Figure 2.6**). VIP activity was recorded during the awake state but also during anesthesia. We found only one cluster of responses which weakly but broadly responded to all sounds (**Figure 2.6**). In wakefulness, this cluster was mostly silent during sequence presentations (**Figure 2.6j**). In contrast, other clusters were active in the awake state and inactive under isoflurane anesthesia (**Figure 2.6a-i**). Two of these clusters responded selectively to the B tone and displayed clear stimulus-specific adaptation and therefore stronger responses at the end of rare AAAAB sequences (**Figure 2.6f & h**). The rest of clusters were more or less strongly inhibited during sequence presentation and displayed a strong non-specific, or weakly specific excitatory response at the end of all sequences (**Figure 2.6a-e, g, i**). This excitatory response, which was seen much more

sporadically in VIP-interneurons (**Figure 2.5e**) was characterized by a long latency from the offset of the last sound. This latency corresponded roughly to the expected temporal position of the next tone in the sequence, as if the sequence would have an additional tone.

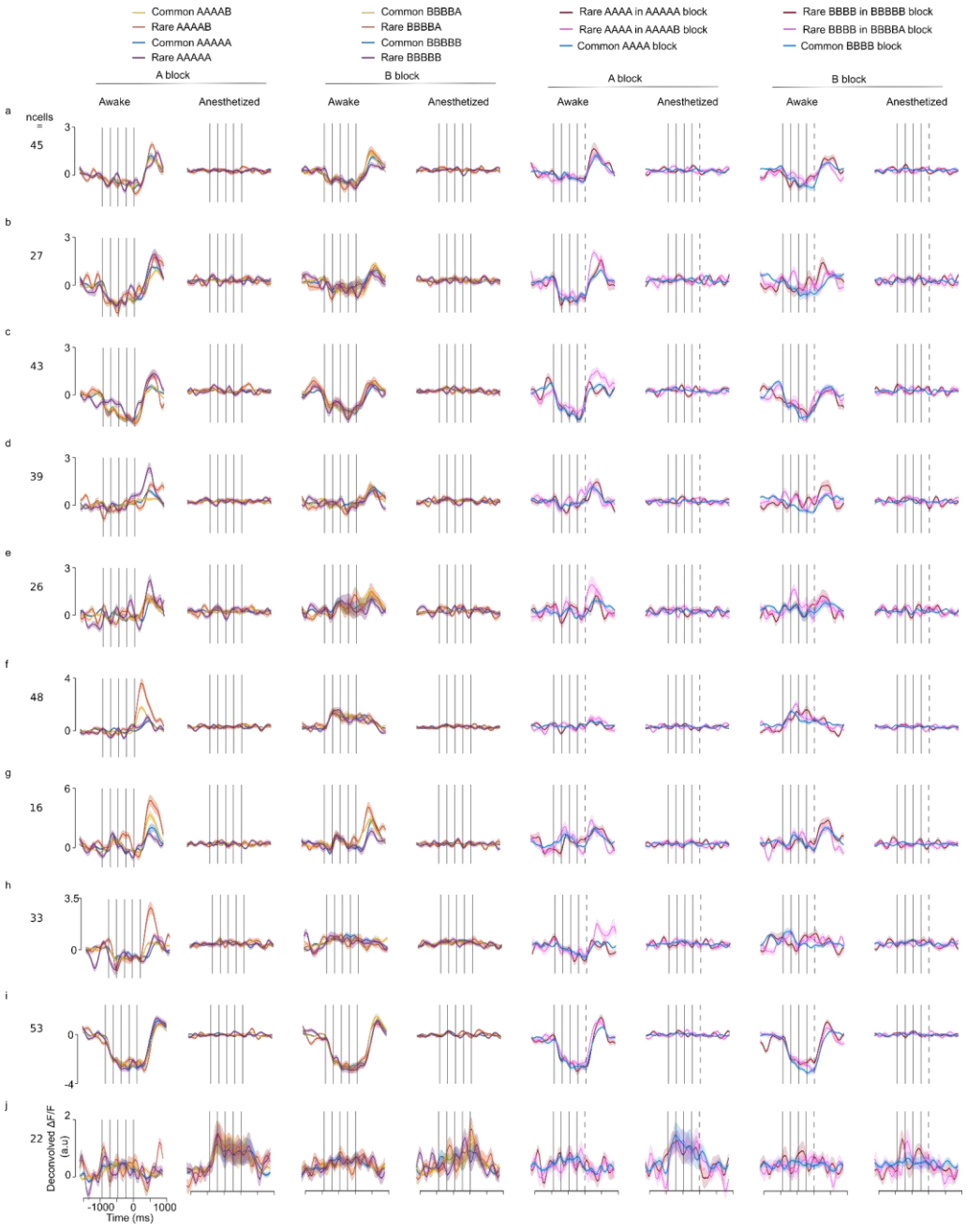


Figure 2.6 : Sequence termination coding in VIP neurons. *(**First 4 columns**) Mean responses to rare and common AAAAB and AAAAA sequences for 10 clusters of VIP neurons were defined according to the similarity of their response signature. Responses are shown both in the awake state and under anesthesia. (**Columns 3 and 4**) The responses of the same clusters to shorter AAAA and BBBB sequences during blocks in which they are expected (Common AAAA or BBBB blocks) and blocks in which they are not (AAAAB, BBBBA, AAAAA and BBBBB blocks).*

This long latency is consistent with the idea that these responses signal the termination of the sequence as a whole, and therefore represent a violation of the expectation that the sequence continues, eventually based on inter-tones interval regularity. In order to further test this idea, we plotted responses of the same cluster but now for omission sequences with only four tones (AAAA or BBBB) in three different contexts. In the first two contexts, the omission sequences were oddball sequences within a block dominated by 5-tones sequences that included the same first four tones. In the third context, the 4-tones sequences were presented in a block containing exclusively one 4-tone sequence (either AAAA or BBBB). In all three contexts, the identified response cluster displayed a sequence termination response, which was shifted about 250 ms earlier than when the response occurred after a 5-tones sequence, equal to the duration between two stimuli. This observation corroborated the idea that VIP interneurons signal a violation of the expectation that the sequence continues, mostly independently of the next tone expected in the sequence. This observation is consistent with observations that VIP interneurons are involved in the processing of some local predictions (Attinger, Wang, and Keller 2017; Garrett et al. 2020). However, we did not find evidence that VIP neurons directly participate in the global violation signals observed when imaging neurons expressing GCAMP6s under the synapsin promoter (**Figures 2.1-4**).

2.4. Discussion

In this study, we used two-photon calcium imaging to provide evidence that the mouse auditory cortex represents global violation responses within a sparse set of neurons which very specifically responded with long latencies to the global violation and to the conditions in which this violation occurred. The groups of neurons that responded to global violations usually did not respond to sound onsets and were slightly responsive to sequence offsets (**Figure 2.1-4**). These global violation responses indicate that the mouse brain is able to build expectations about the structure of acoustic sequences that develop across ~ 1s, if these sequences repeat often enough. Although global responses in humans and macaques were mostly evidenced with protocols in which the sequences are presented at short intervals, we have shown that predictions do not require short intervals and persist even if the acoustic sequences are interleaved with several tens of seconds of silence and resist even to perturbation sounds occurring during these long intervals. Hence, global violation responses reveal predictions built on the repetition of particular events rather than on a crystalline temporal regularity (L. Wang et al. 2015), contrary to sequence completion phenomena observed in mouse auditory or visual cortex which both required a specific range of time intervals and an exquisite regularity (Gavornik and Bear 2014; Li et al. 2017).

Global violation responses also clearly differ from the phenomenon of stimulus-specific adaptation. Stimulus-specific adaptation, which explains part of the mismatch negativity response, corresponds to a simple, yet potent, prediction mechanism: responses to a repeated stimulus are decreased by stimulus-specific neuronal and/or synaptic ‘fatigue’ mechanisms. Therefore, a less frequent stimulus, which has not recruited these mechanisms produces a much larger response than the frequent stimulus. This mechanism is a potent way to increase the saliency of short-term

stimulus changes and it is at the heart of the large local violation responses that we robustly observed both in anesthetized and awake animals after the AAAAB and BBBBA sequences (**Figures 1-4**). For example, in the AAAAB sequence, the last A has undergone more stimulus specific adaptation than B. Therefore, the response to B appears much larger than the response to the last A. This effect is magnified if AAAAB is a rare deviant in a AAAAA-dominated block, as A-specific adaptation is even larger and B is much less frequently presented than in a AAAAB-dominated block, therefore responses to B are even larger when AAAAB is a rare deviant sequence (**Figures 2.1-4**). In contrast, global violation responses to rare AAAAA cannot be explained by adaptation as they represent a large response to an additional A tone compared to the predicted AAAAB sequence. Adaptation would on the contrary predict a decreased response to the last A, as it actually occurs in most A responding neurons, except those that specifically respond to the rare AAAAA sequence. Another distinction between stimulus specific adaptation and global violation responses is the sensitivity of the latter to anesthesia (**Figure 2.4**) (Lynn Uhrig et al. 2016). It was shown in human subjects that major modifications of the alertness state, as occurs in sleep or in a wandering mind condition, impairs global responses (Bekinschtein et al. 2009; Strauss et al. 2015). This indicates that global violation responses require mechanisms that are the hallmark of the awake state. For example, the mechanisms enabling sequence specific short-term memory are likely to be key for establishing predictions at the time scales engaged in our protocol. Recent observations in mice have shown that cortical dynamics is profoundly modified under anesthesia (Bharioke et al. 2022; Filipchuk et al. 2022; Suzuki and Larkum 2020). Some of these dynamical changes could explain the loss of global predictions.

This raises the question of the mechanisms at play to generate global violation responses in the mouse auditory cortex. Human and monkey data suggests that they would be first generated

in frontal areas and propagated back to the auditory cortex (Chao et al. 2018; El Karoui et al. 2015; Wacongne et al. 2011). The late onset of global violation responses in our data is compatible with this view. Our clustering analysis for broad GCAMP6s expression revealed up to two clusters signaling global violation responses with markedly different response latencies (**Figures 2.1-2**). This observation is potentially of high interest to understand the mechanisms of global predictions and of the associated violation responses, as it suggests that the process underlying these responses is dynamical and that the information propagates in a sequence of activation of different neurons. Our study also brings new evidence about the cell types involved in global violation responses. Imaging of PV and VIP-positive interneurons revealed that they are very weakly represented in these cell types. This result was expected for PV cells that were so far little involved in the channeling of prediction-related signals, although their modulation leads to changes in stimulus specific adaptation (Natan et al. 2015). VIP interneurons in contrast were involved in the visual cortex in the signaling of some local violations, such as the absence of a repeated image (Garrett et al. 2020), or in the computation of mismatch responses between visual flow and motor behavior (Attinger, Wang, and Keller 2017), which may be regarded as local predictions.

In the auditory cortex, we did not find evident global violation responses in VIP interneurons. However, we found strong signaling of some local violations in VIP cells. In particular, we observed that a large subset of VIP interneurons was involved in an excitatory response following the termination of sequences independent of their acoustic content (**Figure 2.6**). In particular, as observed in the visual cortex during the repetition of identical images (Garrett et al. 2020), the activity of several functional clusters of VIP interneurons was inhibited by tone presentations and ramped up in between tones, with a large excitatory boost when the sequence of tones stops (**Figure 2.6**). This may correspond to a similar prediction mechanism across the auditory and

visual cortex. These broad omission responses share with global violation responses a large latency to the last sound of the sequence (**Figure 2.6**), indicating that the computation of these responses requires long integration times. It is unclear whether the detection of sequence terminations in VIP interneurons relies on the regular interval between sounds and thus reflects a precisely timed expectation or if it is a temporally less accurate process. In monkeys and humans, global violation responses seem to rely more on the identity of stimuli in the sequence rather than on the temporal regularity (L. Wang et al. 2015). Whether this would transfer to sequence termination responses in VIP neurons is an open question, as well as whether sequence termination responses participate in the computations leading to global violation responses.

A very important aspect of our findings is the high sparseness of global violation responses, likely less than 1% of the neurons displaying these responses in our protocol. This sparseness is surprising as global violation responses are detected in population activity measures in humans and monkeys, however this is not necessarily a contradiction between mouse and human observations. A similar sparseness of global signals at the single cell level was recently observed in the monkey prefrontal cortex (Bellet et al. 2021). It is conceivable that population-level global violation signals (EEG, MEG, or fMRI) reflect to a large extent subthreshold signals that have only moderate effects on the actual neuronal firing rates. The specificity and sparseness of global violation responses may explain why global violation responses had not been found earlier at single cell resolution. Interestingly, the stimulus specificity of global violation responses may also be related to their sparseness. Indeed, if violation responses are, to some extent, stimulus specific, a large set of neurons will be required to capture the variety of deviations from expectations that a mouse may encounter throughout its lifespan, and that would be encoded in the prediction system that our data uncovers. Sensory responses in the auditory cortex are known to be extremely sparse

(Bathellier, Ushakova, and Rumpel 2012; Hromadka, Deweese, and Zador 2008), raising the question of the function of neurons that are not (or weakly) activated by sounds. Already, a large set of data indicates that auditory cortex neurons encode non-auditory information such as behavioral, heteromodal, or reward-related information (K. Kuchibhotla and Bathellier 2018). One may also hypothesize that a large part of the auditory cortex encoding space is dedicated to prediction processing, based on a sparse code.

The observation that global violation responses are stimulus-specific (**Figure 2.2**) is striking as it is not fully in line with hierarchical predictive coding theory, which rather postulates that detection of global violations is a generic signal independent of the stimulus. Also, recent data in the monkey prefrontal cortex suggest the existence of global violation signals independent of stimulus identity (Bellet et al. 2021). Yet, our observation is not the first example of a non-generic mismatch signal. Recent data suggests that the detection of mismatches between visual and motor inputs in the mouse visual cortex is specific to the stimuli that lead to the mismatch (Attinger, Wang, and Keller 2017; G. B. Keller, Bonhoeffer, and Hübener 2012; Leinweber et al. 2017). It is possible that stimulus-independent violation responses exist in higher-order, associative cortical areas. In this case, one possibility is that these generic global violation responses would be fed back to the auditory cortex and interfere with stimulus representation to yield specific responses. Alternatively, one could also hypothesize that stimulus-specific violation responses in early sensory areas are generated, in part, locally and contribute to the emergence of stimulus independent violation signals in downstream structures. The evidence that global prediction processing exists in the mouse cortex is the first step to start addressing these new important mechanistic questions.

2.5. Methods

Cranial window implantation and viral injections.

All procedures were conducted in accordance with protocols approved by the French Ethical Committees #59 and #89 (authorizations APAFIS#9714-2018011108392486 v2 and APAFIS#27040-2020090316536717 v1). We used 8 to 12-weeks-old C57BL/6J, VIP-Cre and PV-Cre male and female mice housed 1-7 per cage, in normal light/dark cycle (12h/12h). Cranial window implantation and viral injections were performed under ketamine medetomidine or under isoflurane anesthesia (1.3-1.7%) with body temperature maintained constant at 37°C using a thermal blanket. Part of the right masseter was surgically remove to expose the temporal bone. A craniotomy of 5 mm in diameter was drilled over the auditory cortex on the right hemisphere. Three injections of 150 nl of AAV1.Syn.GCaMP6s or AAV1.Syn.Flex.GCaMP6s (~$1x10^{-12}$ $vg.ml^{-1}$), obtained from Addgene and Vector Core (Philadelphia, PA, USA), were performed with glass micropipettes and a programmable oil-based injector (Nanoliter 2000 & Micro 4; World Precision Instruments) at 30 $nl.min^{-1}$. The craniotomy was sealed with a glass window comprising a circular coverslip (5 mm diameter, pre-sealed with cyano-acrylate glue) and a metal post for head-fixation was implanted using two dental cements: Super-Bond C&B (Sun Medical Co. Ltd.) directly on the bone and Orthojet (Lang Dental, Wheeling, Illinois) for final sealing of the cranial window and the fixation post. Mice were given one week to recover from the surgery. Imaging was performed between 4 to 7 weeks after virus injection.

Sounds and stimulation protocols

We generated two pure tones at 4 and 12 kHz with an intensity of 70dB SPL over a duration of 50ms including 10ms linear intensity up- and down-ramp to avoid onset and offset artifacts. The 4 and 12kHz tones were labeled as A and B respectively, and combined in sequences of 4 to 5 tones with 237.5 ms time-intervals in between tone onsets. We used two sound stimulation protocols, one with short inter-sequence intervals (**Figures 2.1-2 and 2.4-6**), and one with long inter-sequence intervals (**Figure 2.3**). The short interval protocol included 10 blocks of 125 sequences separated by a fixed 1.5s silence period. 2 blocks each contained 50 sequences of AAAA, 50 sequences of BBBB and 25 blank sequences. 8 others blocks consisted of one common 5-tones sequence repeated 25 times alone and then 75 times randomly interleaved with a rare 5-tones sequence, repeated 15 times and with a 4-tones omission sequence repeated 10 times. There were 4 different block types repeated each twice:

block type 1 (common AAAAB, rare AAAAA, omission AAAA), block type 2 (common AAAAA, rare AAAAB, omission AAAA), block type 3 (common BBBBA, rare BBBBB, omission BBBB), block type 4 (common BBBBB, rare BBBBA, omission BBBB). The long interval protocol included two blocks of 45 sequences each separated by a random 3-5s silence period preceding a white noise burst (70dB SPL, 50ms duration including 10ms intensity ramps) and a 25s period of silence. The 45-sequences blocks consisted of one common 5-tones sequence repeated 10 times alone and then 25 times randomly interleaved with rare 5-tones and rare omission sequences, repeated each 5 times. There were 2 different block types repeated only once: block type 1 (common AAAAB, rare AAAAA), block type 2 (common AAAAA, rare AAAAB).

Two-photon calcium imaging in awake mice

One week before imaging, mice were trained to stand still, head-fixed under the microscope for five consecutive days for 15 min to 1 h per day. Then mice were imaged for 1h long sessions with up to four vertical depths imaged per mouse on different days. Imaging was performed using a two-photon microscope (Femtonics, Budapest, Hungary) equipped with an 8 kHz resonant scanner combined with a pulsed laser (MaiTai-DS, SpectraPhysics, Santa Clara, CA, USA) tuned at 920 nm. The objective was a 10x Olympus (XLPLN10XSVMP), obtaining a field of view of 1000 x 1000 µm. Images were acquired at 31.5 Hz during trials of 315.5 sec. For the anesthesia, VIP and PV interneurons experiments, 2-photon imaging was performed with an acousto-optic microscope (Karthala) combined with a pulsed laser (Insight, Spectra Physics). The objective was a 16x (N16XLWD-PF, Nikon). Images were acquired from four planes at 19.1 Hz per plane interleaved by 50 µm with fields of view of 478 x 478 µm. In the short interval protocol, calcium activity was acquired continuously during an entire block with the Femtonics microscope and during half a block with the karthala microscope (the interruption between two half blocks was below 3s). In the long interval protocol, calcium activity was recorded during the sequence and until 1s after the white noise presentations (between -1 s to -2s and 4.5 s to 7.5 s from sequence onset).

Calcium imaging data analysis

Motion artifacts, regions of interest selection, and the signal extraction were carried out using the Python-based version of Suite2p (Pachitariu et al. 2017). Then, data analysis was performed using custom Matlab scripts. Neuropil contamination was subtracted by applying the following equation: $F_{cor}(t) = F(t) – 0.7\ F_n(t)$. Then the change in fluorescence $\Delta F/F_0$ was computed

as $(F_{cor}(t) - F_0) / F_0$, where F_0 is estimated as the minimum of gaussian filtered calcium trace, for each block. $\Delta F/F_0$ was then temporally deconvolved to yield a more accurate estimate of neuronal firing rate changes, using a linear algorithm using the following formula: $r(t) = \Delta F/F_0'(t) + \Delta F/F_0(t) / \tau$ in which $\Delta F/F_0$' is the first temporal derivative of $\Delta F/F_0$ and τ the calcium decay time constant which we set to 2 seconds for GCaMP6s. After deconvolution a Gaussian smoothing filter (σ = 1.5 or 2 frames) was applied to the data.

Clustering analysis

Single trial responses to each 5- or 4-tone sequence were extracted from the raw deconvolved traces including a 0.5s baseline and a 1 s post-sequence period for each neuron. We averaged all trials (including trials from repeated blocks) for each condition, separating for each sequence the context in which it is common and in which it is rare. Clustering was performed by using the average response signatures of each cell to selected sets of sequence and conditions as described in the Figure 2. legend. We used agglomerative hierarchical clustering based on the Ward method to group together neurons with similar response signatures. The similarity metric used was the Pearson correlation between response profiles. A threshold was applied on the resulting dendrogram to obtain 100 clusters. These clusters were then manually sorted to remove all obvious groups of non-responsive cells as indicated by an absence of activity above or below the typical baseline noise level for the cluster.

Fraction of neurons selective to a condition

To evaluate the fraction of neurons selective to the rare or common presentation of a sequence in the time bin spanning 0 to 500ms after the last tone onset, we first evaluated the signal

to noise ratio (SNR = mean activity divided by its standard deviation) for this time bin. Cells were then ranked according to their SNR. For each cell, we also computed the p-value of a Wilcoxon ranksum test evaluating the probability that responses in the considered time bin are identical for the common and the rare condition. Using an alpha-value of 0.01, we then computed the fraction of cells with a p-value below the alpha-value for different values of an SNR threshold above which cells were retained. This procedure allowed us to evaluate whether or not there was a larger fraction of cells selective to the common or rare condition in among the higher SNR cells.

Cross-validated classifiers

In order to evaluate if violation responses provide sufficient information to differentiate between a rare and a common sequence, we used a population decoder applied to all neurons of a given dataset pooled across mice and recording sessions, and cross-validated using a leave-one-out procedure. The activity of the test trial was left out and was not used in any step of the following procedure. (i) To reduce dimensionality and improve classifier training, we select neurons whose average activity from 0 to 500ms after the last tone onset is significantly different between the common and the rare condition for the considered sequence using a Wilcoxon ranksum test with alpha value 0.01. (ii) We construct population vectors with the selected neurons and then trained a linear SVM classifier to discriminate between population vectors of the common and rare condition. (iii) We test the classifier with the left out trial for all time bins describing the response to the sequence, including the training time bin. Then another trial is left out and the procedure from (i) to (iii) is performed again. This is repeated until all rare and common trials are tested. The performance of the classifier in each time bin is given as the average of the classifier output (0 = wrong, 1 = correct classification) for all test trials balanced by the number of trials per condition.

Data availability

The data that supports the findings of this study will be available on Xenodo.

Code availability

The code used for the analysis is available from the corresponding author upon request.

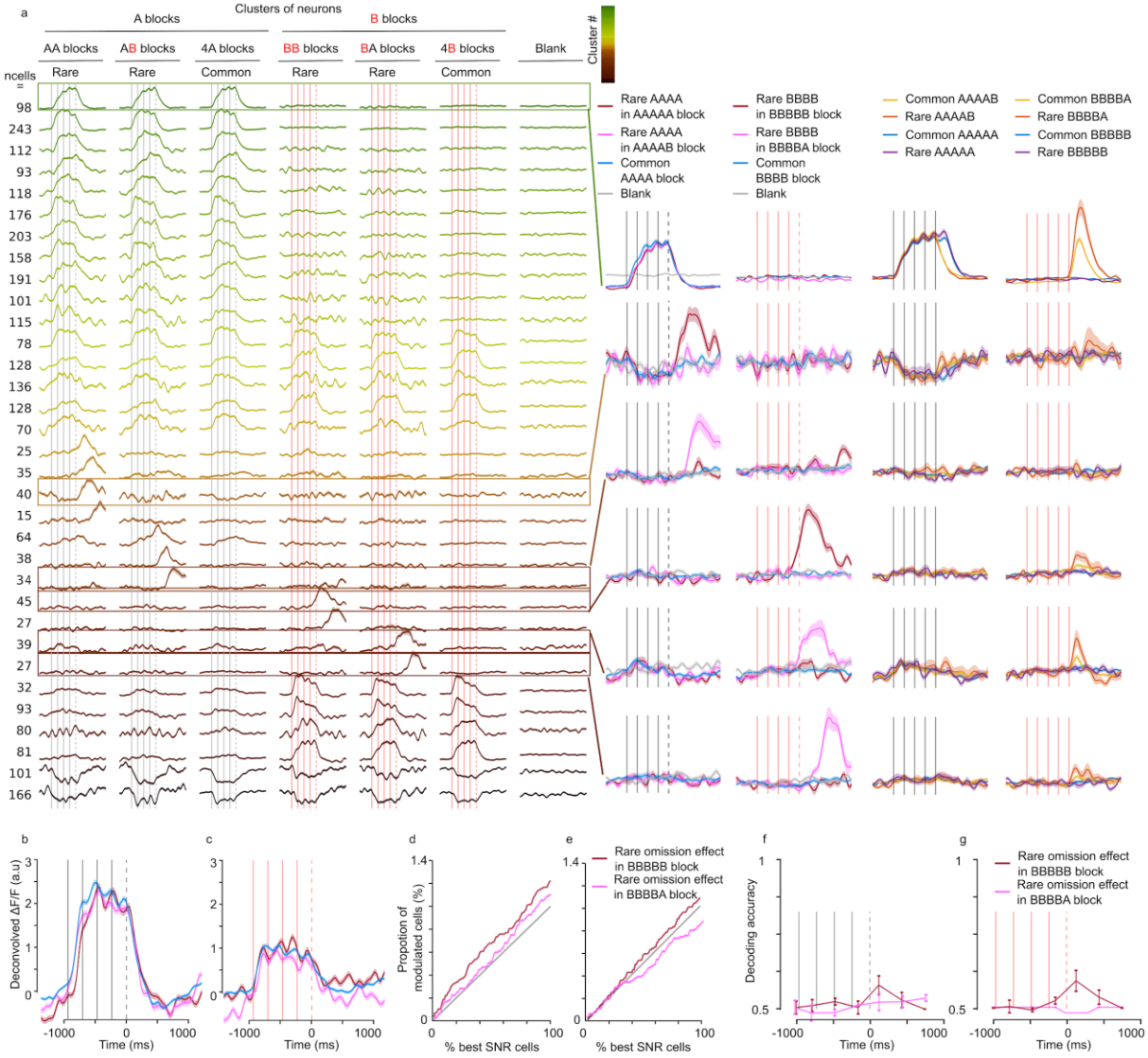
Acknowledgements

This work was supported by the European Research Council (ERC CoG 770841) and Fondation pour l'Audition (FPA IDA02). We acknowledge the support of the Fondation pour l'Audition to the Institut de l'Audition.

Author contributions

S.J. performed the experiments and analyzed the data, S.J., S.D., TvK and B.B. designed the study and B.B. wrote the manuscript with the help of all authors.

2.6. Supplementary figure



Supplementary Figure 2.1: Omission responses in mouse auditory cortex. a. *(left) Responses of all neuronal clusters and (right) zoom in for particular clusters which either display a typical response (1st cluster) or omission-specific responses (2nd to 6th cluster). The two columns on the left show responses to omission sequences (4-tones) and the two columns on the right show responses to 5-tone sequences for the same cluster. The clustering was performed on responses to*

omission sounds and blank stimulus only. ***b.*** *Population responses to AAAA omission sequences in the three different conditions (cyan: block with only AAAA sequences, brown: block with common AAAAA, magenta: block with common AAAAB sequences).* ***c.*** *Population responses to BBBB omission sequences in the three different conditions (cyan: block with only BBBB sequences, brown: block with common BBBB, magenta: block with common BBBBA sequences).* ***d.*** *Fraction of cells responding specifically to rare AAAA sequences for different values of a threshold on the response SNR of the cells, in the AAAAA dominated blocks (brown) and in the AAAAB dominated blocks (magenta) as compared to the reference block where only AAAA sequences are played.* ***e.*** *Fraction of cells responding specifically to rare BBBB sequences for different values of a threshold on the response SNR of the cells, in the BBBBB dominated blocks (brown) and in the BBBBA dominated blocks (magenta) as compared to the reference block where only BBBB sequences are played.* ***f****. Performance of a cross-validated classifier to discriminate common AAAA sequences from rare AAAA sequences in AAAAB (magenta) and AAAAA (brown) blocks.* ***g****. Performance of a cross-validated classifier to discriminate common BBBBB sequences from rare BBBB sequences in BBBBA (magenta) and BBBBB (brown) blocks.*

3. Global violations responses during an auditory detection task

3.1. Goal of the behavioral experiments

As presented in the introductory chapter, human studies suggested the detection of global effects in the local-global paradigm to be a conscious process involving attention. This paradigm includes short-term (local) and long-term (global) regularities. The violation of the local regularity elicits MMN-like responses and the violation of the global regularity elicits a P300-like response on human ERPs. In contrast to the local violation, the detection of the global violation does not occur in an automated, pre-attentive manner. It requires the perception of the rule on longer time scales and active capturing of the global deviant. The global effect, the P300 ERPs, was found in the subjects that have been asked to count the global violations (Bekinschtein et al. 2009) or only to be attentive to the sound without counting (Wacongne et al. 2011). The effect disappeared in mind-wandering subjects or when they were performing another distracting visual task at the same time. After the experiment, none of these subjects were able to report the existence of the global rule, except one subject in the mind-wandering group. This effect also reduces or disappears in the non-communicating minimally conscious (MCS) or vegetative (VS) patients asked to be attentive to the sound, respectively, and vanishes also during sleep although MMN persists.

In macaque monkeys, the global effect has been detected in prefrontal, parietal and cingulate cortices in passively listening to sound animals. The activated circuit during the detection of global violation formed a “global workspace” by making the information available through the workplace which corresponds to a conscious state in humans (Dehaene et al. 2006; Sergent et al.

2021; Bekinschtein et al. 2009; Sergent, Baillet, and Dehaene 2005) and other monkeys' studies (Panagiotaropoulos, Dwarakanath, and Kapoor 2020; Kapoor et al. 2022). The authors hypothesized that the animals were aware of the global violations.

Interestingly, as described in the last chapter, we found a sparse population of neurons that respond robustly to the global violations in mice passively listening to sounds. Because of the observations in humans suggesting the involvement of attention in global violation detection, we have aimed to study if this is also the case in mice. For this purpose, we developed a behavioral task to make the mouse attentive to the sound sequences, while preventing the confusion of the global effect responses with arousal, reward-triggered activity (Petreanu et al. 2012; Pi et al. 2013), motor activities during licking (Schneider, Nelson, and Mooney 2014; McGinley et al. 2015; Zhou et al. 2014) and movement artifacts under the two-photon microscope. In other words, my goal was to avoid changing the meaning of each sequence for the animals because of an immediately associated reward. Ideally, we wanted to have mice pay attention almost equally to all sequences while being able to demonstrate by some indirect behavioral measure that they differently perceive the different sequences. It was also critical that mice do not move or do not receive a reward between 0.5s before and 2s after the onset of the sequence.

The design of the task was influenced by a visual categorization task in monkeys (Minamimoto, Saunders, and Richmond 2010). During this task, monkeys learned to hold a lever and release it when a small red target in the middle of an image becomes green. The amount of the reward was changed based on the visual categorical cue which was the image behind the target. For example, monkeys received more reward if the visual cue was a dog and less delayed reward if the cue was a cat. The error rate of the monkeys depended on the predicted reward based on the

categorical information and was significantly higher for the less rewarded category (Minamimoto, Saunders, and Richmond 2010). This led us to believe that changing the amount of reward would give us the possibility to measure stimulus perception without direct change in the response to the stimulus, which could produce a confound with global violation signals.

Our strategy was to associate the reward with a third tone which was a white noise (WN) click of 500 ms to prevent the mice from licking on the sequence and to modulate the amount of reward with respect to a sequence played a few seconds before. One of the sequences was rewarded twice more. The head-fixed water-restricted mice were trained to lick after the WN to receive a reward 1s after the onset of the WN. To measure how the neuronal responses are modulated in a local-global paradigm during this task, I recorded the activity of 9471 neurons in the auditory cortex of 2 mice during 13 sessions, with a two-photon microscope during the behavioral task.

3.2. Experimental protocol

Behavioral setup

Behavioral experiments were done using homemade software (Elphy, G. Sadoc, UNIC, France) coupled to a National Instrument card (PCIe-6351). We could monitor licks when the head-fixed mice sitting on an aluminum foil in a tube, were closing a 5V electrical circuit by licking on a conductive lick port. The licking voltage was estimated by measuring the voltage through a series resistor in this circuit. Sounds were amplified (SA1 Stereo power amp, Tucker-Davis Technologies) and delivered through high-frequency loudspeakers (MF1-S, Tucker-Davis Technologies). Water delivery was controlled with a solenoid valve (LVM10R1-6B-1-Q, SMC) around 5 to 6µl for less rewarded and around 10 to 12µl for more rewarded sequences. Before

starting the training procedure, mice were water restricted for two consecutive days. Training started by training mice to lick for receiving water without sound for approximately 2 days depending on the mouse. Then they were trained to lick after a white noise click (WN) for 1 day. Finally, they were trained to receive a reward when licking after the WN and to avoid the licking during the sound sequences. Each behavioral session contained 50 high incentive and 50 low incentive rewarded trials which allowed mice to obtain their daily water supply of ~800 µl. Mice were trained from Monday to Friday usually and received water ad libitum from Friday evening to Saturday morning before resuming deprivation. Two-photon imaging was started approximately three days after completion of the training (i.e. when licking on the WN occurred in more than 90% of the trials and licking on the sequence was less than 50%).

Sounds and task structure

I used two sequences of 5 short tones, AAAAA and AAAAB as in the passive experiments. The A tone was a pure tone of 4kHz and the B tone, 12kHz with an intensity of 70dB SPL and a duration of 50 ms including 10ms linear intensity up- and down-ramp to avoid onset and offset artifacts. The interval between tones onset in each sequence was 237.5 ms and the interval between two sequences of tones was a random interval of 32.5 to 33.5 s. These sequences were played in two different blocks. In each block, one of the sequences was played more commonly and the other one more rarely (Figure 3.1b). Each trial had the following structure: (i) a waiting period of 25s to force mice to stop licking before the appearance of the sound sequence AAAAA or AAAAB (ii) a random stimulation delay of 0.5 to 1.5 s, (iii) the sequence presentation, (iv) a random delay of 2 to 5s, to avoid the prediction of the WN appearance that signals the reward and to prevent licking before the WN, (v) the WN presentation, and (vi) a fixed response window of 1s (Figure 3.1c).

Licking during the response window triggered delivery of the water reward. Reward volume was twice larger for a AAAAB sequence. The reward was delivered at the end of the reward window which permits estimating the licking behavior in function of the predicted reward based on the preceded sequence rather than the reward itself. All the analyses are obtained after removing the trials where the mice were licking on a period of 0.5s before and 2s after the onset of the sequence. The sessions with more than 70% of licking during this period and less than 60% of performance were removed.

Two-photon calcium imaging

Imaging was performed during the behavior using a two-photon microscope (Femtonics, Budapest, Hungary) equipped with an 8 kHz resonant scanner combined with a pulsed laser (MaiTai-DS, SpectraPhysics, Santa Clara, CA, USA) tuned at 920 nm. The objective was a 10x Olympus (XLPLN10XSVMP), obtaining a field of view of 1000 x 1000 µm. Images were acquired at 31.5 Hz during trials of 9s every 25s.

3.3. Results

Behavior

In the local-global paradigm a violation of global regularities of the sound has been only detected in the brain activities of subjects attentive to the sounds (Bekinschtein et al. 2009). The purpose of our task was to repeat these experiments. Three mice were trained to perform the above-mentioned sound WN detection task during which AAAAA or AAAAB sequences were presented

a few seconds before the WN and indicated the reward value (10-12 µl for the AAAAB sequence and 5-6 µl for the AAAAA sequence, Figure 3.1a-b). The sequences were played in two different blocks, one block in which AAAAB presented commonly and large rewards dominates, and another block in which AAAAA presented commonly and low rewards dominates. Mice performed this basic task well from almost the first session. The performance is calculated only based on the WN detection and shown in Figure 3.1c for 10 selected days (see methods). For avoiding motor responses during licking and the motion artifacts under the two-photon imaging, I removed all the trials in which mice were licking on the sequence (see methods). In the selected days, mice on average licked during the sequence in less than 40% of trials per session (Figure 3.1c). As others did in the past, I observed that it is difficult for mice to learn to wait and not lick on non-rewarded sounds (Francis et al. 2018). In the AAAAB block, the AAAAB sequence was played commonly, and as the AAAAB sequence is the more rewarded stimulus this block was more rewarded than the AAAAA block.

In order to identify behavioral correlates of the high reward anticipation, I first measured reaction times. I observed that when pooling data from all three animals, reaction time was significantly lower for the AAAAB block in which rewards are higher (Figure 3.1d). This indicates that the animals perceived the change of incentive and were slightly more rapid when the overall reward probability was high. Note, however, that this difference was significant in only one out of three animals (Figure 3.1d). Moreover, I did not find significant reaction time differences between low and high reward trials within a block (not shown). Therefore, contrary to the visual categorization task which motivated our task design, the prediction of reward incentive through the priming sequence had probably only a weak influence on this response latency.

I also observed that the amplitude of the licking signals after the white noise was very similar between more and less rewarded trials (Figure 3.1e). However, peaks of the licking signal were larger for rare AAAAB (high reward) trials than for the common (low reward) AAAAA trials in the same block (Figure 3.1e). In the AAAAB block in which high rewards are common, this effect is not seen (Figure 3.1e). We interpreted the increase at peaks of the signal as a higher synchronization of the rhythmic licking activity across trials after the rare AAAAB sequence in anticipation of the high incentive reward (Figure 3.1f-g). We, therefore, quantified synchronization by computing the correlation coefficient between the average licking signal and the single trial licking signal (Figure 3.1f) or alternatively by computing the correlation coefficient across all pairs of trials (Figure 3.1g). This is confirmed in both cases, for data pooled across animals, the observation made on the averaged traces (Figure 3.1e) that increased synchronization occurs when high rewards are rare.

Although robustness of lick timing is hard to interpret, together, this observation suggests that mice slightly changed their anticipatory licking after the WN burst based on the predicted reward value given by the priming sequence. This change is hardly detectable in individual behavioral data. This was actually one of the purposes of our task design to obtain a slight change in the behavior. However, this design made it difficult to conclude that the mice clearly learned the difference of predicted reward between the two sequences.

Two-photon calcium imaging

To investigate how the neuronal population's responses to these sequences change with attention, or at least with task engagement, two-photon imaging was performed during the behavior

of two mice. Then I performed the same analyses as for passive sound presentations. Note that in the time course of my thesis, this behavioral experiment was performed before the passive sequence presentation with long intervals presented in Chapter 2 (Figure 2.3), which was originally conceived as a control for demonstrating the impact of behavioral engagement on global violation responses and copy all details of the behavioral task except for the reward schedule. Therefore, comparing the results presented in Figure 2.3 and in Figure 3.2 allow us to identify response specificities resulting from behavioral engagement.

A first surprising result was that the deconvolved calcium traces averaged across the entire population of recorded neurons showed no difference between the end responses to the rare and the common AAAAB sequence. This is not the case in the passive condition (Figure 2.3), in which the response at the end of the rare AAAAB is clearly larger than the response to the end of the common AAAAB sequence. This could be interpreted as an impact of behavior on stimulus-specific adaptation as suggested in a previous study (Yaron et al. 2020). In order to verify this idea we looked at the responses to functionally defined clusters of neurons as defined in chapter 2 (Methods). We observed that some clusters, specific to the B tone (e;g. clusters #9 & 10), still displayed stimulus-specific adaptation despite the apparent absence of this phenomenon at the population. Such clusters were also seen in the passive presentation protocol (Figure 2.3). Hence there is no complete disappearance of adaptation in single-cell responses. In fact, the specificity of the behaviorally engaged mice is the presence of clusters that responded more strongly to the common local violation (common AAAAB) than to the rare local violation. This is very clear in clusters #4 & 5. Such response profiles were not observed in passive animals. One may speculate that these responses are predictions of the high reward, which contrary to violations are reinforced when the stimulus is more predictable, i.e. when the AAAAB is common. Interestingly,

symmetrically we observed also cells which respond to the end of AAAAA sequences but more when they are common than when they are rare (cluster #7). Possibly these could be low reward cells. Therefore, a possible alternative or complementary explanation for the absence of stimulus-specific adaptation at the population level is the compensation of the adaptation phenomenon by new, engagement-specific responses that display an opposite response profile and may be signaling behavior-related predictions.

Aside from this striking effect, we did not observe a strong change in the global violation responses during behavior as compared to passive stimulation. First, we observed a similarly low number of cells with global violation responses (only cluster #6 for rare AAAAA and putatively cluster 8 for a delayed response to rare AAAAB sequences). Second, when using classifiers to predict rare against common AAAAA or AAAAB sequences the scores obtained were not better than in the passive situation (Figure 3.2c-d vs Figure 2.3).

In summary, in these 2 mice, we did not observe any encouraging evidence of a boosting of global responses during the behavioral context. The global effect seemed even reduced compared to passive animals (Figure 3.2 vs Figure 2.3). This can arise from the choice of the task. Maybe using a third tone as the cue for the reward distracted the attention from the sequences. It is also possible that associating a reward to each sequence changed the meaning of the sound for animals and this weakened regularity estimation mechanisms Also associating a reward to each sequence and doubling the reward for one of the sequences limited the number of trials that mice could perform each day (by achieving their daily water supply), making the data noisier and harder to explain.

Therefore our task design did not allow us to increase the magnitude of global effects. It also did not unambiguously demonstrate modulation of attention. Hence, we believe that new experimental designs need to be developed to address this question. This will be discussed in the next section.

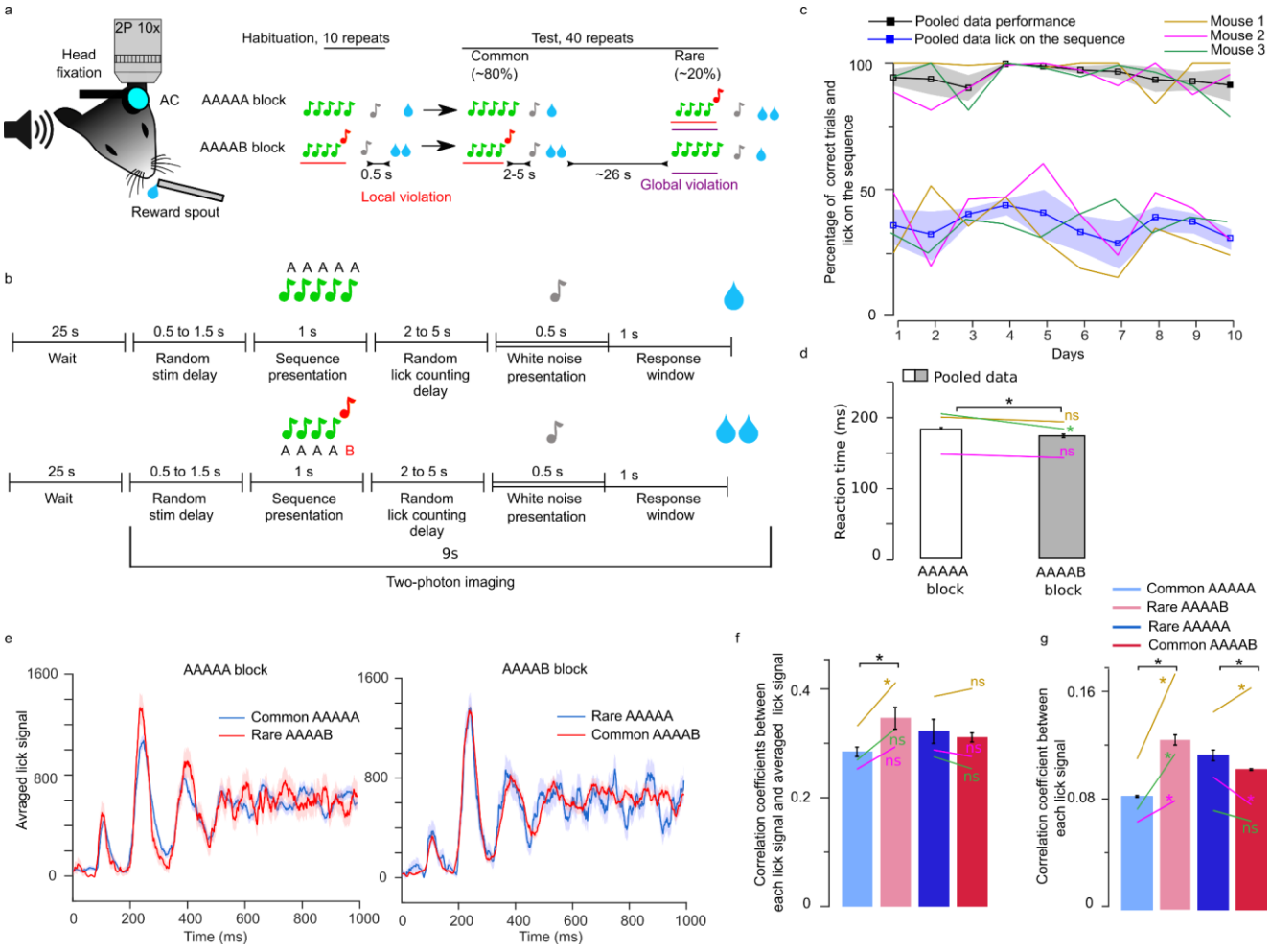


Figure 3.1: Slight modulation of licking behavior for the high incentive trials. a. *Sketch of the experimental setup* ***b.*** *Sketch of the two sound sequence blocks presented during each session and the sound detection task.* ***c.*** *Performance and the percentage of removed trials where the mice licked at*

least once between 0.5s before and 2s after the onset of the sequence. The plot displays average data of three mice and individual curves for each mouse. ***d.*** *Reaction time for pooled data in the two blocks of sequences. The reaction time in block AAAAB, the more rewarded block, is significant when pooling data across mice (Wilcoxon rank-sum test, $p<0.05$) lower than for block AAAAA. However, this difference is only significant for one of the three mice and significance across the three mice cannot be assessed due to the low number of animals.* ***e.*** *Trial-averaged licking signal traces (voltage across the lick detection circuit) for 3 mice in the high intensive trials (red) and low incentive trials (blue) in two AAAAA blocks (left) and AAAAB blocks (right). Calculated on 1s time window from the onset of the WN before delivering the reward.* ***f.*** *Correlation coefficient between each lick signal and the averaged lick signal was calculated for 4 different conditions (Wilcoxon rank-sum test, *: $p<0.05$). High incentive conditions in red and low incentive conditions in blue* ***g.*** *Mean correlation coefficient between each pair of single-trial lick signals calculated for 4 different conditions.*

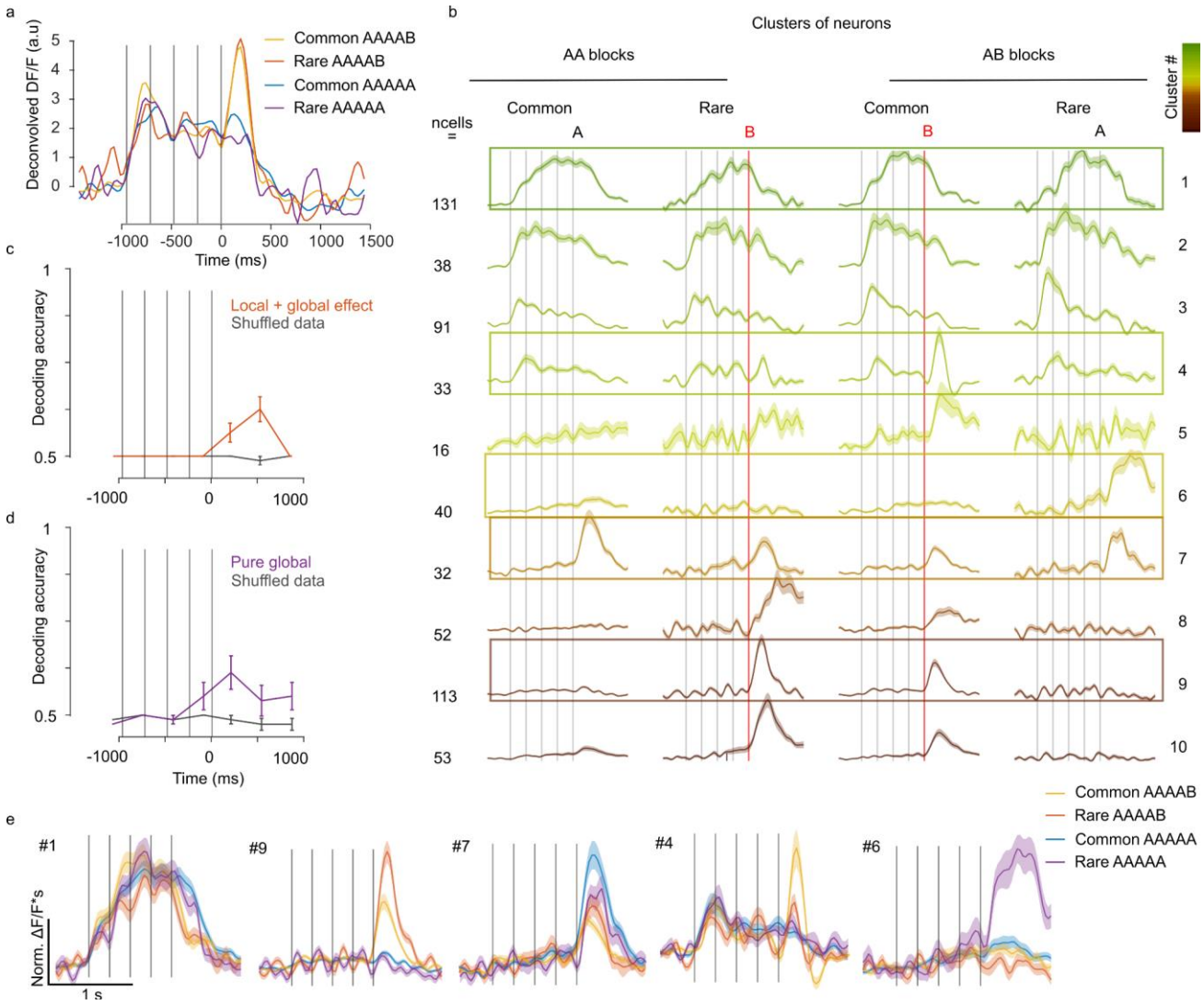


Figure 3.2: Difference between a common and a rare sequence is reduced in the auditory cortex in behaving mice. a. *Mean population responses (deconvolved calcium signals) to AAAAB and AAAAA sequences depending on whether they are rare or common in the block.* ***b.*** *10 clusters of different response types to these 4 conditions.* ***c.*** *Performance of a fully cross-validated classifier for predicting the rare AAAAB sequence against the common AAAAB sequence (orange) and the same but on shuffled conditions (gray).* ***d.*** *Same as c for the rare AAAAA sequence against the common AAAAA sequence.* ***e.*** *Zoom of interesting clusters from panel b. The main difference with the passive data presented in*

chapter 2 (Figure 2.3) is the existence of clusters that respond more to the common AAAAB sequences compared to the rare ones.

4. Discussion

Throughout this thesis, I have studied the mouse auditory cortex responses to a paradigm called “local-global” using sequences of tones that could be regular at two different hierarchical levels. By introducing changes in the sequences, the regularity of the presentation of sequences is modulated: when there is a change of stimulus within the sequence it is called a local violation, and when there is a change across sequences, it is called a “global” violation. Studies in humans and monkeys suggest that these two types of violations arise from two different mechanisms. While the detection of the local violations happens in an automatic, preattentive manner, detection of the global violations only occurs in conscious, attentive human subjects. So far the latter mechanism has been shown to be available only in the human and monkey brain.

The results presented in chapters 2 and 3 indicate that sparse populations of auditory cortex neurons encode a global regularity violation signal, which is specific to the stimulus and context that produces the violation. These signals do not seem to be widely expressed in VIP and PV interneurons and would rather come from pyramidal neurons, although we cannot exclude a role of somatostatin-positive interneurons so far. VIP interneurons displayed non-specific sequence termination responses which had not been previously described to my knowledge. Moreover, we have shown that global violation signals are not boosted within the context of a behavioral task that we had originally developed to foster auditory attention. However, behavioral engagement generated new response types in a population of the neurons in the auditory cortex, which may be related to well-predicted rewards and thereby generated a convergence of the population activity level for rare and common local violations. These results are further discussed below.

4.1. Effect of behavioral engagement on violation signals

Although we could not demonstrate that the involvement in the behavioral task described in Chapter 3 led to a modulation of attention and/or an increase of global violation signal, imaging the auditory cortex during the behavioral task led to some interesting and unexpected results.

First, we observed that rare and common AAAAB sequences led to the same population level activity, which was not observed for the control passive listening task with identical stimulus presentation intervals (Figures 2.3 vs 3.2). Based on this population-level observation, one can speculate that by associating sequences with a reward, the importance of the regularities and their violations decreases. The response to the rewarded sound may be enhanced and therefore stay robust to inform the rest of the brain. As a consequence, the difference between common and rare stimuli is reduced.

Auditory cortex plasticity associated with learning is a well-established phenomenon. This plasticity may reflect in part an auditory memory (Weinberger 2007; Froemke et al. 2013). In the context of SSA protocols, a previous study has demonstrated that when a common stimulus is associated with fear conditioning, the auditory responses are enhanced to the common sound after conditioning and the difference between the same stimulus presented rarely or commonly is reduced (Yaron et al. 2020). In this study, it was even found that the responses to the repeated common condition were on average larger than responses to the rare condition. In our case, both rare AAAAB and common AAAAB are rewarded equally, which may cancel out the importance of how regularly these sequences are played in the block. In other words, the auditory cortex may use its resources for local reward prediction rather than for regularity predictions that are less crucial for passing robust information to the rest of the brain about the reward value of the sequence

(no matter if it is presented rarely or commonly). Our data actually supports this view by showing that B-tone-specific neurons still display adaptation for common AAAAB. Crucially, what actually changes during task engagement is the appearance of neurons that respond more strongly to the end of the common AAAAB sequence than to the end of the rare AAAAB sequence. These neurons may thus be related to reward prediction. This observation, although preliminary (n=2 mice), is novel and interesting, because, so far, the effect of behavior on local violation response was only studied with local field potential (Yaron et al. 2020). Hence the single cell resolution description of the phenomenon could not be performed in this previous study. Our data suggest that the interaction between behavioral significance and regularity violation coding is not due to a modulation of stimulus adaptation but to the recruitment of new cells potentially dedicated to behavioral predictions.

An alternative hypothesis to explain the increased response to common AAAAB during behavior, maybe a block effect on attention. Since the block AAAAB is the more-rewarded block, the increased response to the common AAAAB sequence could be explained by a global change in arousal compared to the AAAAA block. However, two observations are not in agreement with this explanation. First, only slight changes in reaction time were detected in behavior (Figure 3.1). Second, some clusters of neurons responded more to the common than to the rare sequence (Figure 3.2e, #6 and 7) while others responded in an opposite manner, suggesting a more detailed change in cortical activity rather than a global modulation. Furthermore, a fully cross-validated classifier could decode the rare AAAAB and common AAAAB sequences above chance but with an important reduction in the accuracy compared to the passive-condition data (Figure 3.2c). This is not in line with a global attentional up-modulation, from which one would expect better stimulus encoding. Decoding accuracy for the rare and the common AAAAA sequences was much less

affected (Figure 3.2d). However, it is necessary to confirm these results in the future with more experimental data and more controls. For example, more insights could also be gained by monitoring pupil modulations as a proxy of arousal.

Another unwanted effect of behavioral engagement in our task was eventually a reduction of attention to the sequences or at least an absence of attention modulation with respect to the passive condition. We used a third tone as a cue for reward in order to generally increase auditory attention. However, if the mice did not learn or only weakly learned the rule that the amount of reward is modulated based on the sequence, they may not have cared much about the sequences. In this case, our task may have diverted partially the attention from the sequences instead of reinforcing it. To disambiguate these two options, it would be interesting to record the activities of the same neurons to these sequences in a control condition with many standard sequences in passive animals, in a block where AAAAB and AAAAA sequences are surrounded by other rare sequences. In this case, if the auditory cortex response changes to the full presentation of these two sequences by associating them with rewards, we should see similar response amplitudes during behavior and during passive listening. Moreover, measuring the response to a CCCCB sequence in this context would tell us if mice do associate B with the higher reward and A with the lower reward without taking into account the full sequence, or if the full sequence eventually matters for their prediction. Passive blocks can be presented to passive animals after the behavior, when they are less aroused for example if one manages to keep them head-fixed long enough.

4.2. Global violation signals versus known prediction signals in rodents

The most studied paradigm to detect novelty responses in the rodent auditory cortex consists of a deviant introduction in between the repetitions of a common tone in the so-called stimulus-specific adaptation (SSA) and mismatch negativity (MMN) protocols. In both cases, one observes a higher response to the same stimulus when it is presented rarely than commonly. The main issue of these protocols is that the resulting effect can be well explained by adaptation mechanisms. However, a part of the enhanced response also likely reflects an active predictive system rather than only passive adaptation. This is demonstrated when one controls for adaptation by comparing the response to the rare stimulus when the standard is repeated and predictable to the response observed in an unpredictable "many-standard" context for the same frequency of the rare stimulus. Yet, cross-stimuli adaptation is more difficult to control in the "many standard" context. Therefore protocols in which the repetition of a single stimulus is supposed to generate predictions, do not offer optimal conditions to disambiguate active predictions from passive adaptation through mechanisms such as activity-dependent decrease of neuronal excitability or short-term plasticity in synapses.

Short-term plasticity refers to changes in synaptic strength. An action potential at the presynaptic terminal causes the release of a proportion of vesicles containing neurotransmitters. If another action potential happens before the recovery of the vesicles' stock, the number of released neurotransmitters will be smaller and result in a weaker postsynaptic potential. This dynamic can be explained by two parameters, the rate of vesicles used per presynaptic action potentials and the

time constant of vesicle recovery (David and Shamma 2013). Both the decay and recovery time constant of corticocortical synapses are better explained by two-time scales (compared to a single time scale), a stronger and more rapid time scale of about several hundred milliseconds, and a weaker and more persistent time scale of several seconds (10 sec) (Varela et al. 1997; Benda 2021). Neither the recovery nor the decay is affected by the blockage of N-methyl-D-aspartate (NMDA) receptors (Markram, Wang, and Tsodyks 1998; Varela et al. 1997). The two-time constants reported in these studies are consistent with the adaptations observed in cat A1 neurons during a SSA paradigm (Ulanovsky et al. 2004). Moreover, SSA is not dependent on NMDA receptors (Polterovich, Jankowski, and Nelken 2018). Another potential contributor to adaptation is the activity-dependent decrease in neuronal excitability which is commonly observed in cortical neurons. This phenomenon, also termed spike-frequency adaptation, is mainly due to calcium-dependent potassium channels (Engel, Schultens, and Schild 1999).

Essentially, any prediction-like effect explained with short-term plasticity is expected to have the following features. First, the response is present both when the stimulus is rare and when it is common. But it is attenuated when it is common. Second, the attenuation depends on the repetition of a particular stimulus within the time scale of adaptation phenomena (seconds to tens of seconds). Third, for the same reason, it should decay rapidly and disappear without stimulation after several seconds to tens of seconds. Single neuron studies of the SSA paradigm in the rodent auditory cortex do not permit full disambiguation of the adaptation and the prediction signals, in particular, because it always overlaps with the above-mentioned adaptation features. Hence it is not clear whether this rodent SSA is coming from a mechanism that relies only on the passive synaptic or neuronal adaptation of repeated stimuli or if it is a mechanism that actively predicts the next stimulus either based on a simple recurrence rule.

In contrast, the local-global paradigm uses two levels of regularities (i.e. two recurrence rules) and allows to generate the conditions in which the repetition of a stimulus becomes a surprising event. Five tones (AAAAA) can elicit a positive novelty response when they represent a violation of global regularity. A positive response to an additional repeated stimulus cannot be explained by known adaptation mechanisms which all feature increasing adaptation and thus decreasing response when there is a supplementary repeat. Therefore the local-global paradigm can convincingly distinguish between prediction violations and adaptation at the single neuron level.

In the local-global paradigm instead of playing tones in a continuous stream, they are presented as sequences of tones. In addition to creating a local violation in the sequence for animals, it is offering the opportunity to generate a global violation in between these sequences. Detection of global violations on the longer times scales cannot be formed only with an inter-stimulus interval (ISI), as it is the case for SSA. It requires a memory of the whole sequence which develops over 1s considering the global context, in which perceiving a violation relies on a memory of several sequences in a block. In humans and monkeys, the detection of global violations is suggested to arise from a clearly different mechanism. They are slower than the local responses. They have been found only in attentive human subjects and they are engaging the higher brain areas. In the mouse auditory cortex, we found different populations of neurons encoding for the local and global violations. In the critical condition, animals are listening to the repetition of AAAAB sequences, and the AAAAA sequences are presented rarely in between these sequences. We see a population of neurons that respond only at the end of the rare AAAAA sequences without responding to the tones. This shows that the mouse auditory cortex can detect the global regularities at longer time scales and this cannot be explained by adaptation as the repetition of

five identical tones elicit a novelty response. Crucially, the neuronal populations which encode the violation generated by the AAAAA sequence do not respond to the A tones during the sequence or other non-predictive contexts.

In my Ph.D. thesis (Chapter 2), I show that the global violations are resistant to the long interval of ~30s between sequences. In humans, MMN/P3a ERP components are sensitive to the inter-stimulus interval (ISI) and vanish for ISI more than a few seconds (Mäntysalo and Näätänen 1987; Pegado et al. 2010). But P3b is insensitive to the ISIs exceeding tens of seconds and has been detected using ISIs as long as 10 mins (Wetter, Polich, and Murphy 2004; Rohaut and Naccache 2017). The late responses to the global violations detected in the auditory and precentral cortex have the same features as P3b ERPs. This is consistent with our passive (Figure 2.3b, clusters 6,8) and also active (Figure 3.2b, clusters 6,8) long ISI data, where most of the global violation clusters respond only to the global violations and no other stimuli, and have quite longer latencies similar to some clusters observed in a short ISI paradigm (Figure 2.1d, clusters 18,22,26,24). There is no experimental proof showing the persistence of adaptation for intervals longer than ~1.5 s in the SSA context (Ulanovsky, Las, and Nelken 2003) and 4s in natural sounds (Asari and Zador 2009) under anesthesia. One study has shown the absence of adaptation to the rare stimuli played every ~7s (Ulanovsky et al. 2004). Yet, the slow and long time constant of adaptation (Ulanovsky et al. 2004; Varela et al. 1997) and also some response clusters in our data that look like adaptation, do not allow us to neglect adaptation even with a ~30s long trial interval.

Another important specificity of global violation response is its sensitivity to anesthesia. In both humans and monkeys under anesthesia, global violation responses disappeared. In the mouse auditory cortex, both global violations whether to the rare AAAAB or rare AAAAA showed

an important reduction during anesthesia, and differences between rare and common AAAAA disappeared. In contrast, stimulus-specific adaptation is routinely observed under anesthesia.

The local-global paradigm enabled us to study new capacities of the mouse auditory cortex to detect a violation in global regularities on longer time scales. An ability that (i) requires memorization of the repetition of several sequences of ~1s over long intervals, (ii) cannot be explained with adaptation, (iii) occurs with longer latencies, (iv) is not affected by the long trial interval and (v) is abolished under anesthesia.

4.3. Dynamics of global violation signals and prediction models

Unlike other well-known oddball paradigms, the design of the local-global paradigm allows us to evaluate the hierarchical organization of prediction errors. In both human and monkey studies, they observed activities in an early and late time window in the auditory cortex as well as higher processing areas such as frontal, parietal, and cingulate cortices. The late responses in the auditory cortex are suggested to come from higher processing areas. In this Ph.D. work, I recorded only the mouse auditory cortex and I found very similar results to those of the human and monkey auditory cortex. This allows us to speculate that a part of these violations is feedback from higher processing areas to the mouse auditory cortex. In the following, I will compare these results in more detail. In the auditory cortex of humans and monkeys, studies have detected:

i) An early local violation's effect (XXXXY sequence). Which should be a bottom-up signal due to a local transition probability violation. In the mouse auditory cortex, there is also a population of neurons with a short latency that respond to the

violation of B in AAAAB sequence (e.g. Figure 2.1, cluster #29). The same for the BBBBA sequence.

ii) An early global violation effect, an increased response to the rare XXXXY compared to the common XXXXY sequences, which should be also a bottom-up signal. This could be explained by lower transition probabilities of X→Y in a XXXXX block. In the mouse auditory cortex, we also observed this effect in populations of neurons that respond to both rare and common AAAAB sequence and more when it is rare (e.g. Figure 2.1, cluster #29). The larger response when AAAAB is rare results probably from a lower transition probability of A→B in the AAAAA blocks or equivalently from weaker adaptation to B.

iii) A late global violation's effect on the rare AAAAB and AAAAA sequence which is suggested to be a top-down signal from higher processing areas for updating the predictions. In mice, we could observe cluster that responds to the rare sequences but with a very long latency (e.g. Figure 2.1, cluster #24) for AAAAB and (e.g. Figure 2.1, cluster #22) for AAAAA. Some clusters have (Figure 2.1, #26 and #18) have slightly shorter latencies but only respond to rare sequences. The temporal precision of temporally deconvolved two-photon calcium imaging is only about 100-200ms. Therefore precise latency measurements are difficult and I did not perform such measurements yet. It is possible that early-firing global violation-specific neurons are also partially sensitive to transition probabilities. But also in general, it is possible that global violation signals generate temporal sequences of activity.

In VIP and PV interneurons, we did not observe a strong effect on the global violations.

We observed some sequence termination clusters in VIPs with quite long latencies, but which are not violation specific. These responses can be seen as a violation of the expectation tone that the sequence will continue. Alternatively, it can be explained as a sequence chunking response. A "chunk" is a group of items that occur together. For example, one can suppose each sequence to be a chunk and the brain encodes them as a single group (Dehaene et al. 2015). The sequence chunking has been studied broadly in the language (Saffran, Aslin, and Newport 1996) but also in other domains (Bor et al. 2003; Minier, Fagot, and Rey 2016) but there are few studies at the single-cell level (Fujii and Graybiel 2003; Jin, Tecuapetla, and Costa 2014) and it's detailed neuronal mechanisms remain unknown.

Based on these results and following the hierarchical predictive coding framework, we can try to construct a simplistic circuit for explaining the global violation effects that we observed and the absence of global effects in the VIP or PV interneurons can look like Figure 4.1. One can imagine a model in which the prediction error neurons receive inhibitory inputs from the prediction neurons and excitatory input from the sensory stimuli (G. B. Keller and Mrsic-Flogel 2018). The predictive neurons encode the global regularity of a given block such neurons may be present locally in the auditory cortex and/or located in higher associative cortical areas. If the sensory input is not the same as the predicted input, a prediction error signal will be generated in the prediction error neurons. This will be fed back to prediction neurons. The fact that we observe different timings in prediction error signals indicates that there may be several sources of prediction errors (Figure 4.1). The early violation effect observed in the auditory cortex could be an early prediction error generated locally and the latter responses are excitatory feedbacks from higher-order areas. However, despite their timing difference both early and late prediction error neurons have the same stimulus-specificity. Hence if they come from distinct circuits, these circuits may perform the same

computations. The rare omission responses in human studies last over long timescales with the first peak around 100 ms followed by a late response in the auditory cortex and also in the precentral cortex coinciding with the P3b signals. In the mouse auditory cortex, I also find two distinct populations of neurons with an enhanced response to the omitted tone with short and long latencies (Supplementary Figure 2.1). Omission responses, therefore, share very similar properties to global responses as predicted by the simple input-prediction comparison model of Figure 4.1.

While this simplistic model can explain some of our observations, some aspects of the response are not totally compatible with it. It is difficult to explain for example that errors generated by a different stimulus and by an omission activate different neurons (but we observed a weak overlap between the two responses, Supplementary Figure 2.1). Also, the global prediction error lasts several 100ms, much longer than sensory-driven responses in the same circuit. The input-prediction comparison model of Figure 4.1 would thus imply that the sensory inputs received by prediction error neurons have long durations and latencies, and thus would markedly differ from typical input-driven representations. Also, this simplistic model does not explain how specific predictions are actually generated in prediction neurons, particularly if the prediction must account for the position of the stimulus in the sequence. Overall, more exquisite models should be developed to explain our results.

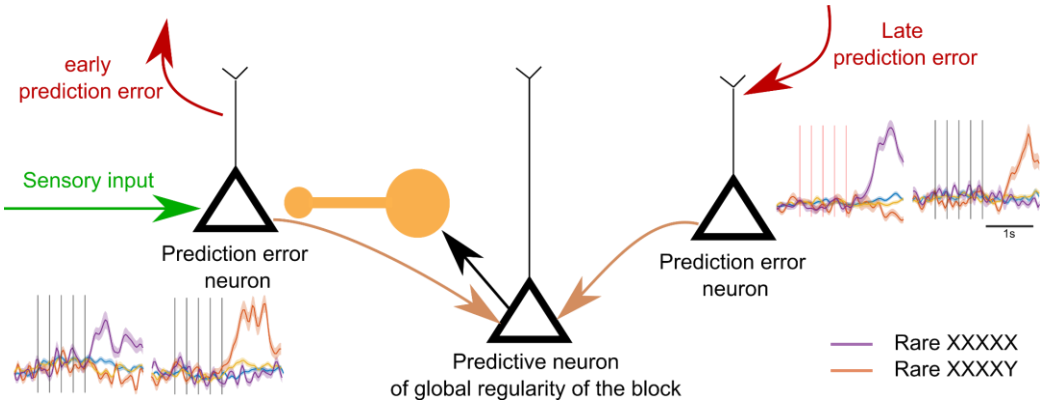


Figure 4.1: A simplistic schematic of a hypothesis for explaining global sequence violation.

Triangles are excitatory neurons and the circle is an inhibitory neuron.

Interestingly also, the VIP sequence termination responses share similar latencies with the global violation responses. This could suggest that VIPs contribute to global violation detection. Possibly, the VIP sequence termination responses are feedback from higher-order areas, as they have long latency and the VIPs have strong long-range connections. However, it is difficult to construct a model that can transform the non-specific VIPs termination response into a specific response as seen in the global violation-specific responses of figures 2.1 and 2.2.

4.4. Sparseness and specificity of local violation signals and the predictive coding hypothesis

The predictive coding framework suggests that the brain develops an internal model of the world that is used to predict sensory inputs. This prediction in the sensory areas of the brain is compared to the sensory input. Only if this prediction differs from the external input the sensory (lower order) areas will send a prediction error signal to higher level areas. When the input of the lower level areas is the sensory input, the input of higher level areas is the prediction error signals coming from lower level areas. It has been already speculated that, in the sensory cortex, the prediction and prediction error neurons are likely to be stimulus-specific (G. B. Keller and Mrsic-Flogel 2018). By contrast, in higher-order areas pooling prediction errors from many downstream neurons, more generic prediction error signals could be generated. Several experimental data actually demonstrated that the predictions and prediction errors in the sensory areas are stimulus-specific (Attinger, Wang, and Keller 2017; Fiser et al. 2016; G. B. Keller, Bonhoeffer, and Hübener 2012) while the higher processing areas such as PFC are encoding information in a more abstract manner and are not stimulus-specific (L. Wang et al. 2015; Bellet et al. 2021; Shima et al. 2007). In this case, the more abstract feedback signals sent from higher-order areas may interact with the sensory information in the sensory areas to generate stimulus-specific error signals.

The stimulus specificity of prediction errors in the auditory cortex demands a large population of neurons to encode the prediction and violation of many stimuli independently. This is in line with the sparseness of the global violations seen in our data.

In the predictive coding hypothesis, the existence of prediction and prediction error neurons in a given area is sufficient to explain a predictive system. In the mouse auditory cortex, we also see many neurons that encode for the presence of specific stimuli and do not change their response according to the predictability of the stimulus (same response for rare and common sequences). This feature of the auditory cortex is not in line with efficient coding and so with a predictive system. Hence, the auditory cortex is not just a predictive coding system. However, it is possible that the large encoding space in the auditory cortex allows for coexistence. At higher level cortical areas as the transfer of information over long distances cost energy, efficient coding theories would predict an encoding of only the predictions and prediction errors that are transferred from the lower level areas. For example, in the local-global paradigm, the precentral cortex in humans and PFC in monkeys are only encoding the global violations. Efficient coding in the subcortical areas of the auditory pathway seems to be less present, harder to detect, or less studied than in the vision. In the visual system, this compression of information starts already at the retina level. In the auditory system, a recent study shows a repetition suppression effect in the inferior colliculus (Lesicko et al. 2022). Nevertheless, in the visual cortex, there are also evident indications of neurons that are encoding for the stimulus presentation.

What is not clearly shown in our data is the presence of pure prediction signals or predictive neurons. We can see an omission response in the absence of the stimulus which indicates the presence of an expectation or a prediction signal. We can also suppose that the short latency responses to the onset of the sequence are predictive neurons. But we cannot really distinguish neurons encoding sounds from the predictive neurons. There are few studies that propose the existence of these neurons but they are also considering the neurons which are responding after the onset of the stimuli for example Fiser and al. (Fiser et al. 2016) are considering a time window

of 133 ms before to 333 ms after grating onset. The problem with identifying these neurons is that in most cases their responses are similar to the neurons which are responding to sounds. They are suggested to have roughly the same latencies and be stimulus-specific. These neurons have few properties that can help to recognize them. (i) They are supposed to be a different class of neurons than the prediction error neurons, so they are not encoding the violations. (ii) They provide inputs to the local prediction error neurons or also to other cortical areas. (iii) They would receive feedbacks directly or indirectly from other cortical areas. (iv) They should have shorter latencies than the sensory-driven neurons. For understanding the predictive coding framework it is important to clearly demonstrate the existence of these neurons which does not seem to be done (G. B. Keller and Mrsic-Flogel 2018).

Certainly, the predictive processing we described in this work is not providing a complete explanation of cortical functioning. Hence, we tried to explain its power and identify its limits to enable steps toward the formulation of a more complete theory. What is certain that is if we want to make conceptual progress in this field it is important to move away from a pure button-up representation of the sensory inputs in the cortex.

4.5. Future perspectives

In this section, I outline the experiments that may improve our understanding of the neuronal mechanisms underlying global violation detection.

First, I think what is missing in the local-global paradigm is a “many-standard” control. A block where we play the sequences with the same rarity but in between many other rare sequences.

It will be informative to see how the same neurons respond to the same rare sequence when it is presented in an expected vs unexpected context.

It will be also interesting to study if these global responses depend on the regular interval between each tone and also each sequence of tones and how these predictions are modified or interrupted by using random intervals. These interval modulations can be also studied with the VIPs. The sequence termination responses in VIPs to all the sequences can be explained by a violation of a prediction that expects the sequence continues. These responses arise at roughly the same timing as the next tone if it existed, in complete sequences (e.g. XXXXX) as well as omitted sequences (e.g. XXXX). This idea can be verified by changing the interval between the tones and seeing how the latency of the sequence termination responses varies. Or by making them random to reduce expectations and see if they persist.

As described in section 4.2 in humans MMN/P3a ERP components vanish for ISI for more than a few seconds but P3b ERP is insensitive to the ISI and has been detected with ISI of as long as 10 min (Rohaut and Naccache 2017). This suggests the existence of different mechanisms underlying these two novelty responses. In our data, we could detect local and global violation responses with both short 1.5s and long ~30s intersequence intervals. To evaluate if the same population of neurons is responding to the violations with both ISI and how the responses to the local and global violations are changing with a short and long ISI in the same neurons it would be useful to record the same neurons during a short and long ISI paradigm, maybe by reducing the number of repetitions in short ISI to avoid holding mice head- fixed for a long time. I think this comparison should be also interesting in a simple and well-controlled SSA paradigm.

In this Ph.D. work in contrast to the human results, I could not show an enhancement of the global effect with attention. As I said in section 4.1 this can come from the choice of the task. It would be more straightforward to reward directly the sequence but would generate confusion between prediction signals and response signals. The relationship between attention and the global violation detections can be also studied indirectly by measuring the pupil size which is a proxy of the attentional states. It should be also important to verify if the mice can detect behaviorally the global violations in a block.

The last thing which is important to be studied, is the existence of global responses in the higher-order areas, as suggested in human and monkey studies. In rodents, the prediction and prediction errors are mostly studied in sensory areas. One of the reasons can be explained by the fact that the experimentalists have more control over the inputs to these regions while the inputs of the higher-order areas are suggested to be the prediction errors coming from lower brain areas. In a hierarchical predictive system, we expect to detect more abstract prediction errors in higher-order regions. There is strong evidence of the presence of motor-related signals in the rodent visual and auditory cortex (Saleem et al. 2013; G. B. Keller, Bonhoeffer, and Hübener 2012; Leinweber et al. 2017; Attinger, Wang, and Keller 2017; Audette, Zhou, and Schneider 2021) and some evidence suggesting that, in rodents, the anterior cingulate cortex sends predictions to the visual cortex (Leinweber et al. 2017; Hamm et al. 2021). The higher-order areas which have strong connections with the auditory cortex in rodents are the posterior parietal cortex and anterior cingulate cortex (Zingg et al. 2014) and would be extremely interesting areas to study in the context of our protocol.

5. Appendix A: three-photon microscopy

Human and monkey studies suggested that in the local-global paradigm the higher order novelty responses to the global violation of sound regularities activate precentral areas (Wacongne et al. 2011; El Karoui et al. 2015; Chao et al. 2018; L. Uhrig, Dehaene, and Jarraya 2014). This suggests that the late responses observed in the auditory cortex feedback from these areas to update their predictions (Chao et al. 2018; El Karoui et al. 2015). We saw in chapter 2 that these global violations in the mouse auditory cortex tend to appear with longer latencies. In this Ph.D., I aimed to record also the mouse cingulate cortex in addition to the auditory cortex. The cingulate cortex is one of the higher processing regions, equivalent to the precentral areas in humans and PFC in monkeys that have strong feedback and feedforward connections with the auditory cortex (Zingg et al. 2014). But as the cingulate cortex is deep around 0.8 to 1 mm, it is not possible to image this depth with a two-photon microscope in scattering biological tissues. The two-photon microscopy in the best cases can image around 600 um deep. The new developments in optical imaging propose a three-photon microscope that allows calcium imaging in deep scattering tissues. Chris Xu and colleagues were able to image hippocampal pyramidal neurons labeled with GCaMP6s at around 1mm in an intact mouse brain with a three-photon microscope (Ouzounov et al. 2017; T. Wang et al. 2018; Horton et al. 2013). Thus, I was helping in the optimization and testing of a new three-photon microscope in NeuroSpin with Timo Van Kerkoerle during my first years of Ph.D.

The three-photon microscopy is based on the same concept as the two-photon microscopy described in section 1.7 but uses three photons. Here the photons have roughly three times less energy (three times longer wavelength) than one photon microscopy. So a fluorescent molecule

should absorb three simultaneous photons (within a ~100 femtosecond time window) to get excited and emit light. For realizing such low probability events we are using femtosecond lasers where the laser is mostly off and emitting extremely bright and short pulses. This method decreases drastically out of focus fluorescence by avoiding the excitation of the molecules which are not in the focal point. Three-photon microscopy uses longer wavelengths than two-photon microscopy so they are less scattered in the tissue and the light, therefore, penetrates more deeply.

My two major involvements in the development of the three-photon microscope were controlling and monitoring the laser power under objective automatically with a computer and helping in the measurement of the pulse width. I participated actively in the tests made to monitor the microscope improvement by imaging the mouse auditory cortex. I also imaged the mouse vasculatures marked with Dextran and the mouse cingulate cortex.

For the power control, we mounted a half-wave plate which is controlled with a computer, and a Glan-Taylor polarizer that allows us to change the power automatically. A beamsplitter reflects approximately 1% of the beam and sends it to a power meter connected to the computer to monitor the power continuously.

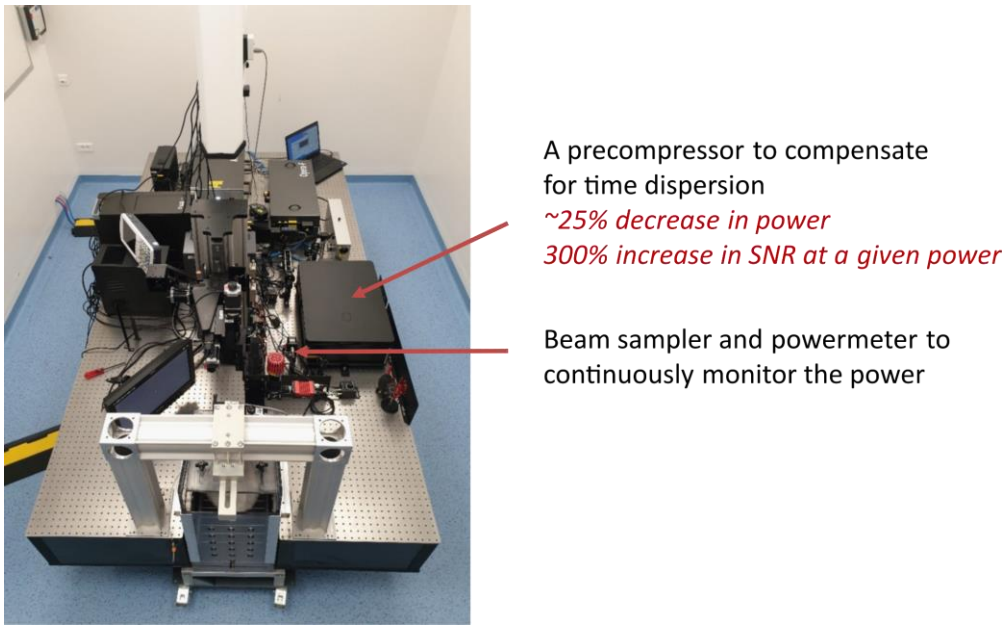


Figure 5.1: Three-photon microscope at NeuroSpin and the technical improvements I was engaged in.

As described above in three-photon microscopy, we need that the 3 photons arrive nearly simultaneously at the focal point. This requires small high peak intensity pulses which explains the interest of femtosecond lasers. If these pulses pass through the optical system of the microscope without compensation, we will have dispersed low-intensity pulses at the sample. To avoid this effect, we need a dispersion precompensator (Figure 5.2) that disperses the pulse negatively compensating for the later impact of the optics which results in a small high peak intensity pulse at the sample. After the installation, we obtained a pulse width smaller than 60 fs, and this decreased 4 times the power needed on the sample to excite the fluorescent beads. By this, we could image down to around 1mm and have good resolution images at around -750 um (Figure

5.3). This depth could be improved by injecting GCaMP6s deep and reducing the background noise due to the superficial layers.

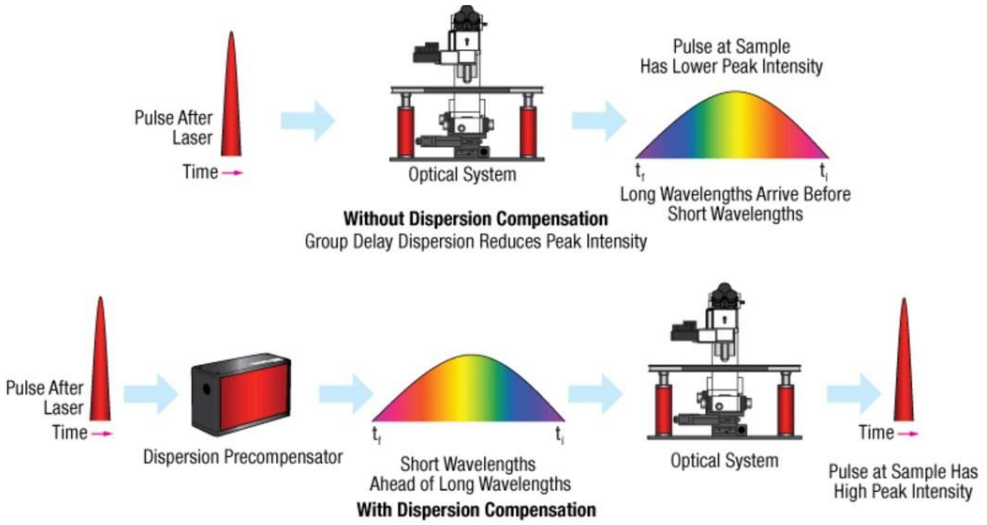


Figure 5.2: Schematic of precompensation.

Scattering decreases, the excitation of photons exponentially at the focal volume. In figure 5.3c where I imaged the mouse brain vasculatures with the three-photon microscope, one can see that brightness is reduced drastically with depth.

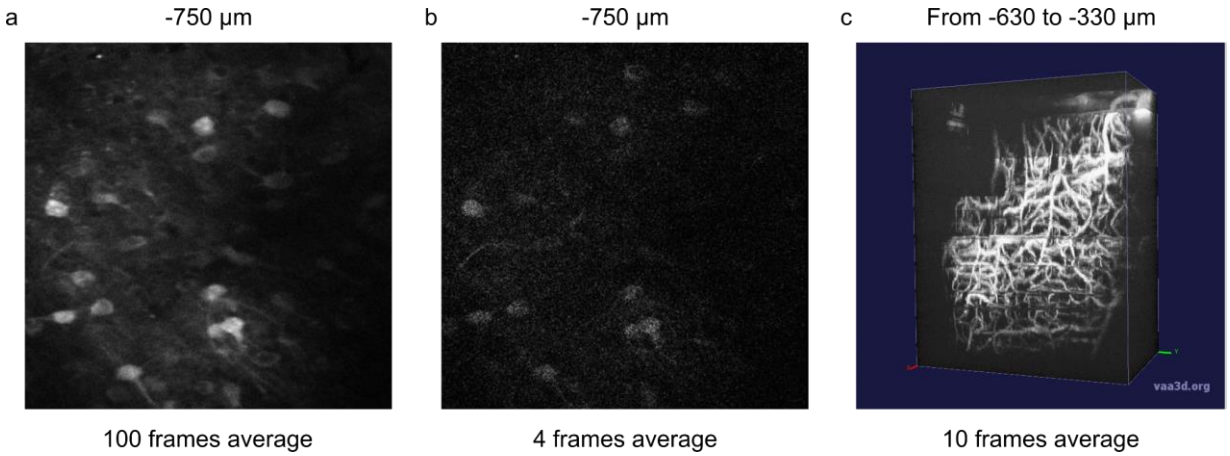


Figure 5.3: Three-photon images captured in the mouse auditory cortex. *a,b: GCaMP6s, c: Dextran. With a field of view of 480 um x 480 um and a frame rate of 25 Hz.*

6. Bibliography

Alain, Claude, David L. Woods, and Keith H. Ogawa. 1994. “Brain Indices of Automatic Pattern Processing.” *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience* 6 (1): 140–44. https://doi.org/10.1097/00001756-199412300-00036.

Alitto, Henry J., and Yang Dan. 2012. “Cell-Type-Specific Modulation of Neocortical Activity by Basal Forebrain Input.” *Frontiers in Systems Neuroscience* 6: 79. https://doi.org/10.3389/fnsys.2012.00079.

Anderson, Lucy A., G. Björn Christianson, and Jennifer F. Linden. 2009. “Stimulus-Specific Adaptation Occurs in the Auditory Thalamus.” *Journal of Neuroscience* 29 (22): 7359–63. https://doi.org/10.1523/JNEUROSCI.0793-09.2009.

Antunes, Flora M., Israel Nelken, Ellen Covey, and Manuel S. Malmierca. 2010. “Stimulus-Specific Adaptation in the Auditory Thalamus of the Anesthetized Rat.” *PLOS ONE* 5 (11): e14071. https://doi.org/10.1371/journal.pone.0014071.

Asari, Hiroki, and Anthony M. Zador. 2009. “Long-Lasting Context Dependence Constrains Neural Encoding Models in Rodent Auditory Cortex.” *Journal of Neurophysiology* 102 (5): 2638–56. https://doi.org/10.1152/jn.00577.2009.

Asokan, Meenakshi M., Ross S. Williamson, Kenneth E. Hancock, and Daniel B. Polley. 2021. “Inverted Central Auditory Hierarchies for Encoding Local Intervals and Global Temporal Patterns.” *Current Biology* 31 (8): 1762-1770.e4. https://doi.org/10.1016/j.cub.2021.01.076.

Attinger, Alexander, Bo Wang, and Georg B. Keller. 2017. “Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex.” *Cell* 169 (7): 1291-1302.e14. https://doi.org/10.1016/j.cell.2017.05.023.

Audette, Nicholas J., WenXi Zhou, and David M. Schneider. 2021. “Temporally Precise Movement-Based Predictions in the Mouse Auditory Cortex.” Preprint. Neuroscience. https://doi.org/10.1101/2021.12.13.472457.

Awwad, Bshara, Maciej M Jankowski, and Israel Nelken. 2020. “Synaptic Recruitment Enhances Gap Termination Responses in Auditory Cortex.” *Cerebral Cortex* 30 (8): 4465–80. https://doi.org/10.1093/cercor/bhaa044.

Ayala, Yaneri A., and Manuel S. Malmierca. 2012. “Stimulus-Specific Adaptation and Deviance Detection in the Inferior Colliculus.” *Frontiers in Neural Circuits* 6: 89. https://doi.org/10.3389/fncir.2012.00089.

Ayala, Yaneri A., David Pérez-González, Daniel Duque, Israel Nelken, and Manuel S. Malmierca. 2012. “Frequency Discrimination and Stimulus Deviance in the Inferior Colliculus and Cochlear Nucleus.” *Frontiers in Neural Circuits* 6: 119. https://doi.org/10.3389/fncir.2012.00119.

Barascud, Nicolas, Marcus T. Pearce, Timothy D. Griffiths, Karl J. Friston, and Maria Chait. 2016. “Brain Responses in Humans Reveal Ideal Observer-like Sensitivity to Complex Acoustic Patterns.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (5): E616-625. https://doi.org/10.1073/pnas.1508523113.

Basirat, Anahita, Stanislas Dehaene, and Ghislaine Dehaene-Lambertz. 2014. “A Hierarchy of Cortical Responses to Sequence Violations in Three-Month-Old Infants.” *Cognition* 132 (2): 137–50. https://doi.org/10.1016/j.cognition.2014.03.013.

Bastos, Andre M., W. Martin Usrey, Rick A. Adams, George R. Mangun, Pascal Fries, and Karl J. Friston. 2012. "Canonical Microcircuits for Predictive Coding." *Neuron* 76 (4): 695–711. https://doi.org/10.1016/j.neuron.2012.10.038.

Bathellier, B., L. Ushakova, and S. Rumpel. 2012. "Discrete Neocortical Dynamics Predict Behavioral Categorization of Sounds." *Neuron* 76 (2): 435–49. https://doi.org/10.1016/j.neuron.2012.07.008.

Bekinschtein, Tristan A., Stanislas Dehaene, Benjamin Rohaut, François Tadel, Laurent Cohen, and Lionel Naccache. 2009. "Neural Signature of the Conscious Processing of Auditory Regularities." *Proceedings of the National Academy of Sciences* 106 (5): 1672–77. https://doi.org/10.1073/pnas.0809667106.

Bellet, Marie E., Marion Gay, Joachim Bellet, Bechir Jarraya, Stanislas Dehaene, Timo van Kerkoerle, and Theofanis I. Panagiotaropoulos. 2021. "Prefrontal Neural Ensembles Encode an Internal Model of Visual Sequences and Their Violations." bioRxiv. https://doi.org/10.1101/2021.10.04.463064.

Benda, Jan. 2021. "Neural Adaptation." *Current Biology* 31 (3): R110–16. https://doi.org/10.1016/j.cub.2020.11.054.

Bharioke, Arjun, Martin Munz, Alexandra Brignall, Georg Kosche, Max Ferdinand Eizinger, Nicole Ledergerber, Daniel Hillier, et al. 2022. "General Anesthesia Globally Synchronizes Activity Selectively in Layer 5 Cortical Pyramidal Neurons." *Neuron* 110 (12): 2024-2040.e10. https://doi.org/10.1016/j.neuron.2022.03.032.

Blackwell, Jennifer M., and Maria N. Geffen. 2017. "Progress and Challenges for Understanding the Function of Cortical Microcircuits in Auditory Processing." *Nature Communications* 8 (1): 2165. https://doi.org/10.1038/s41467-017-01755-2.

Bor, Daniel, John Duncan, Richard J. Wiseman, and Adrian M. Owen. 2003. “Encoding Strategies Dissociate Prefrontal Activity from Working Memory Demand.” *Neuron* 37 (2): 361–67. https://doi.org/10.1016/s0896-6273(02)01171-6.

Carbajal, Guillermo V., and Manuel S. Malmierca. 2018. “The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection.” *Trends in Hearing* 22 (December): 2331216518784822. https://doi.org/10.1177/2331216518784822.

Chao, Zenas C., Kana Takaura, Liping Wang, Naotaka Fujii, and Stanislas Dehaene. 2018. “Large-Scale Cortical Networks for Hierarchical Prediction and Prediction Error in the Primate Brain.” *Neuron* 100 (5): 1252-1266.e3. https://doi.org/10.1016/j.neuron.2018.10.004.

Chen, I.-Wen, Fritjof Helmchen, and Henry Lütcke. 2015. “Specific Early and Late Oddball-Evoked Responses in Excitatory and Inhibitory Neurons of Mouse Auditory Cortex.” *Journal of Neuroscience* 35 (36): 12560–73. https://doi.org/10.1523/JNEUROSCI.2240-15.2015.

Chennu, Srivas, Valdas Noreika, David Gueorguiev, Yury Shtyrov, Tristan A. Bekinschtein, and Richard Henson. 2016. “Silent Expectations: Dynamic Causal Modeling of Cortical Prediction and Attention to Sounds That Weren’t.” *Journal of Neuroscience* 36 (32): 8305–16. https://doi.org/10.1523/JNEUROSCI.1125-16.2016.

David, Stephen V., and Shihab A. Shamma. 2013. “Integration over Multiple Timescales in Primary Auditory Cortex.” *Journal of Neuroscience* 33 (49): 19154–66. https://doi.org/10.1523/JNEUROSCI.2270-13.2013.

Dehaene, Stanislas, and Jean-Pierre Changeux. 2011. "Experimental and Theoretical Approaches to Conscious Processing." *Neuron* 70 (2): 200–227. https://doi.org/10.1016/j.neuron.2011.03.018.

Dehaene, Stanislas, Jean-Pierre Changeux, Lionel Naccache, Jérôme Sackur, and Claire Sergent. 2006. "Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy." *Trends in Cognitive Sciences* 10 (June): 204–11. https://doi.org/10.1016/j.tics.2006.03.007.

Dehaene, Stanislas, Florent Meyniel, Catherine Wacongne, Liping Wang, and Christophe Pallier. 2015. "The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees." *Neuron* 88 (1): 2–19. https://doi.org/10.1016/j.neuron.2015.09.019.

Deneux, T., E. R. Harrell, A. Kempf, S. Ceballo, A. Filipchuk, and B. Bathellier. 2019. "Context-Dependent Signaling of Coincident Auditory and Visual Events in Primary Visual Cortex." *Elife* 8 (May). https://doi.org/10.7554/eLife.44006.

Deneux, T., A. Kempf, A. Daret, E. Ponsot, and B. Bathellier. 2016. "Temporal Asymmetries in Auditory Coding and Perception Reflect Multi-Layered Nonlinearities." *Nat Commun* 7 (September): 12682. https://doi.org/10.1038/ncomms12682.

Descartes, René. 1664. *L'Homme*. Paris: Charles Angot.

Dhruv, Neel T., and Matteo Carandini. 2014. "Cascaded Effects of Spatial Adaptation in the Early Visual System." *Neuron* 81 (3): 529–35. https://doi.org/10.1016/j.neuron.2013.11.025.

Doeller, Christian F., Bertram Opitz, Axel Mecklinger, Christoph Krick, Wolfgang Reith, and Erich Schröger. 2003. "Prefrontal Cortex Involvement in Preattentive Auditory Deviance

Detection: Neuroimaging and Electrophysiological Evidence." *NeuroImage* 20 (2): 1270–82. https://doi.org/10.1016/S1053-8119(03)00389-6.

Duque, Daniel, and Manuel Malmierca. 2014. "Stimulus-Specific Adaptation in the Inferior Colliculus of the Mouse: Anesthesia and Spontaneous Activity Effects." *Brain Structure & Function* 220 (August). https://doi.org/10.1007/s00429-014-0862-1.

El Karoui, Imen, Jean-Remi King, Jacobo Sitt, Florent Meyniel, Simon Van Gaal, Dominique Hasboun, Claude Adam, et al. 2015. "Event-Related Potential, Time-Frequency, and Functional Connectivity Facets of Local and Global Auditory Novelty Processing: An Intracranial Study in Humans." *Cerebral Cortex (New York, N.Y.: 1991)* 25 (11): 4203–12. https://doi.org/10.1093/cercor/bhu143.

Engel, J., H. A. Schultens, and D. Schild. 1999. "Small Conductance Potassium Channels Cause an Activity-Dependent Spike Frequency Adaptation and Make the Transfer Function of Neurons Logarithmic." *Biophysical Journal* 76 (3): 1310–19. https://doi.org/10.1016/S0006-3495(99)77293-0.

Ernst, Marc O., and Heinrich H. Bülthoff. 2004. "Merging the Senses into a Robust Percept." *Trends in Cognitive Sciences* 8 (4): 162–69. https://doi.org/10.1016/j.tics.2004.02.002.

Faugeras, Frédéric, Benjamin Rohaut, Nicolas Weiss, Tristan Bekinschtein, Damien Galanaud, Louis Puybasset, Francis Bolgert, et al. 2012. "Event Related Potentials Elicited by Violations of Auditory Regularities in Patients with Impaired Consciousness." *Neuropsychologia* 50 (3): 403–18. https://doi.org/10.1016/j.neuropsychologia.2011.12.015.

Filipchuk, Anton, Joanna Schwenkgrub, Alain Destexhe, and Brice Bathellier. 2022. "Awake Perception Is Associated with Dedicated Neuronal Assemblies in Cerebral Cortex." bioRxiv. https://doi.org/10.1101/2021.08.31.458322.

Fiorillo, Christopher D. 2008. “Towards a General Theory of Neural Computation Based on Prediction by Single Neurons.” *PloS One* 3 (10): e3298. https://doi.org/10.1371/journal.pone.0003298.

Fiser, Aris, David Mahringer, Hassana K. Oyibo, Anders V. Petersen, Marcus Leinweber, and Georg B. Keller. 2016. “Experience-Dependent Spatial Expectations in Mouse Visual Cortex.” *Nature Neuroscience* 19 (12): 1658–64. https://doi.org/10.1038/nn.4385.

Foster, David H. 2011. “Color Constancy.” *Vision Research* 51 (7): 674–700. https://doi.org/10.1016/j.visres.2010.09.006.

Francis, Nikolas A., Daniel E. Winkowski, Alireza Sheikhattar, Kevin Armengol, Behtash Babadi, and Patrick O. Kanold. 2018. “Small Networks Encode Decision-Making in Primary Auditory Cortex.” *Neuron* 97 (4): 885-897.e6. https://doi.org/10.1016/j.neuron.2018.01.019.

Friston, Karl. 2005. “A Theory of Cortical Responses.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360 (1456): 815–36. https://doi.org/10.1098/rstb.2005.1622.

Froemke, Robert C., Ioana Carcea, Alison J. Barker, Kexin Yuan, Bryan A. Seybold, Ana Raquel O. Martins, Natalya Zaika, et al. 2013. “Long-Term Modification of Cortical Synapses Improves Sensory Perception.” *Nature Neuroscience* 16 (1): 79–88. https://doi.org/10.1038/nn.3274.

Fu, Yu, Jason M. Tucciarone, J. Sebastian Espinosa, Nengyin Sheng, Daniel P. Darcy, Roger A. Nicoll, Z. Josh Huang, and Michael P. Stryker. 2014. “A Cortical Circuit for Gain Control by Behavioral State.” *Cell* 156 (6): 1139–52. https://doi.org/10.1016/j.cell.2014.01.050.

Fujii, Naotaka, and Ann M. Graybiel. 2003. “Representation of Action Sequence Boundaries by Macaque Prefrontal Cortical Neurons.” *Science* 301 (5637): 1246–49. https://doi.org/10.1126/science.1086872.

Garrett, Marina, Sahar Manavi, Kate Roll, Douglas R. Ollerenshaw, Peter A. Groblewski, Nicholas D. Ponvert, Justin T. Kiggins, et al. 2020. “Experience Shapes Activity Dynamics and Stimulus Coding of VIP Inhibitory Cells.” *ELife* 9 (February): e50340. https://doi.org/10.7554/eLife.50340.

Garrido, Marta I., James M. Kilner, Klaas E. Stephan, and Karl J. Friston. 2009. “The Mismatch Negativity: A Review of Underlying Mechanisms.” *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 120 (3): 453–63. https://doi.org/10.1016/j.clinph.2008.11.029.

Gavornik, Jeffrey P., and Mark F. Bear. 2014. “Learned Spatiotemporal Sequence Recognition and Prediction in Primary Visual Cortex.” *Nature Neuroscience* 17 (5): 732–37. https://doi.org/10.1038/nn.3683.

Gregory, R. L. 1980. “Perceptions as Hypotheses.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 290 (1038): 181–97. https://doi.org/10.1098/rstb.1980.0090.

Grimm, Sabine, Carles Escera, and Israel Nelken. 2016. “Early Indices of Deviance Detection in Humans and Animal Models.” *Biological Psychology* 116 (April): 23–27. https://doi.org/10.1016/j.biopsycho.2015.11.017.

Grødem, Sverre, Ingeborg Nymoen, Guro Helén Vatne, Valgerdur Björnsdottir, Kristian Kinden Lensjø, and Marianne Fyhn. 2021. “An Updated Suite of Viral Vectors for in Vivo Calcium Imaging Using Local and Retro-Orbital Injections.” bioRxiv. https://doi.org/10.1101/2021.05.14.443815.

Hamm, Jordan P., Yuriy Shymkiv, Shuting Han, Weijian Yang, and Rafael Yuste. 2021. "Cortical Ensembles Selective for Context." *Proceedings of the National Academy of Sciences* 118 (14): e2026179118. https://doi.org/10.1073/pnas.2026179118.

Hamm, Jordan P., and Rafael Yuste. 2016. "Somatostatin Interneurons Control a Key Component of Mismatch Negativity in Mouse Visual Cortex." *Cell Reports* 16 (3): 597–604. https://doi.org/10.1016/j.celrep.2016.06.037.

Harpaz, Mor, Maciej M. Jankowski, Leila Khouri, and Israel Nelken. 2021. "Emergence of Abstract Sound Representations in the Ascending Auditory System." *Progress in Neurobiology* 202 (July): 102049. https://doi.org/10.1016/j.pneurobio.2021.102049.

Harrison, C. W. 1952. "Experiments with Linear Prediction in Television." *Bell System Technical Journal* 31 (4): 764–83. https://doi.org/10.1002/j.1538-7305.1952.tb01405.x.

Helmholtz, Hermann von. 1867. *Handbuch der physiologischen Optik,*. Leipzig: Voss.

Hershenhoren, Itai, Nevo Taaseh, Flora M. Antunes, and Israel Nelken. 2014. "Intracellular Correlates of Stimulus-Specific Adaptation." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (9): 3303–19. https://doi.org/10.1523/JNEUROSCI.2166-13.2014.

Horton, Nicholas G., Ke Wang, Demirhan Kobat, Catharine G. Clark, Frank W. Wise, Chris B. Schaffer, and Chris Xu. 2013. "In Vivo Three-Photon Microscopy of Subcortical Structures within an Intact Mouse Brain." *Nature Photonics* 7 (3): 205–9. https://doi.org/10.1038/nphoton.2012.336.

Hromadka, T., M. R. Deweese, and A. M. Zador. 2008. "Sparse Representation of Sounds in the Unanesthetized Auditory Cortex." *PLoS Biol* 6 (1): e16. https://doi.org/07-PLBI-RA-1814 [pii] 10.1371/journal.pbio.0060016.

Jin, Xin, Fatuel Tecuapetla, and Rui M. Costa. 2014. “Basal Ganglia Subcircuits Distinctively Encode the Parsing and Concatenation of Action Sequences.” *Nature Neuroscience* 17 (3): 423–30. https://doi.org/10.1038/nn.3632.

Jordan, Rebecca, and Georg B. Keller. 2020. “Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex.” *Neuron* 108 (6): 1194-1206.e5. https://doi.org/10.1016/j.neuron.2020.09.024.

Kapoor, Vishal, Abhilash Dwarakanath, Shervin Safavi, Joachim Werner, Michel Besserve, Theofanis I. Panagiotaropoulos, and Nikos K. Logothetis. 2022. “Decoding Internally Generated Transitions of Conscious Contents in the Prefrontal Cortex without Subjective Reports.” *Nature Communications* 13 (1): 1535. https://doi.org/10.1038/s41467-022-28897-2.

Kato, Hiroyuki K., Shea N. Gillet, and Jeffry S. Isaacson. 2015. “Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance.” *Neuron* 88 (5): 1027–39. https://doi.org/10.1016/j.neuron.2015.10.024.

Kekoni, Jouni, Heikki Hämäläinen, Mikael Saarinen, Jutta Gröhn, Kalevi Reinikainen, Anne Lehtokoski, and Risto Näätänen. 1997. “Rate Effect and Mismatch Responses in the Somatosensory System: ERP-Recordings in Humans.” *Biological Psychology* 46 (2): 125–42. https://doi.org/10.1016/S0301-0511(97)05249-6.

Keller, Andreas J., Mario Dipoppa, Morgane M. Roth, Matthew S. Caudill, Alessandro Ingrosso, Kenneth D. Miller, and Massimo Scanziani. 2020. “A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex.” *Neuron* 108 (6): 1181-1193.e8. https://doi.org/10.1016/j.neuron.2020.11.013.

Keller, Andreas J, Morgane M Roth, and Massimo Scanziani. 2020. “Feedback Generates a Second Receptive Field in Neurons of Visual Cortex.” *Nature* 582 (7813): 545–49. https://doi.org/10.1038/s41586-020-2319-4.

Keller, Georg B., Tobias Bonhoeffer, and Mark Hübener. 2012. “Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse.” *Neuron* 74 (5): 809–15. https://doi.org/10.1016/j.neuron.2012.03.040.

Keller, Georg B., and Thomas D. Mrsic-Flogel. 2018. “Predictive Processing: A Canonical Cortical Computation.” *Neuron* 100 (2): 424–35. https://doi.org/10.1016/j.neuron.2018.10.003.

Khouri, Leila, and Israel Nelken. 2015. “Detecting the Unexpected.” *Current Opinion in Neurobiology* 35 (December): 142–47. https://doi.org/10.1016/j.conb.2015.08.003.

Kimura, Motohiro, Jun’ichi Katayama, Hideki Ohira, and Erich Schröger. 2009. “Visual Mismatch Negativity: New Evidence from the Equiprobable Paradigm.” *Psychophysiology* 46 (2): 402–9. https://doi.org/10.1111/j.1469-8986.2008.00767.x.

Koelsch, Stefan, Wolfgang Heinke, Daniela Sammler, and Derk Olthoff. 2006. “Auditory Processing during Deep Propofol Sedation and Recovery from Unconsciousness.” *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 117 (8): 1746–59. https://doi.org/10.1016/j.clinph.2006.05.009.

Kuchibhotla, K., and B. Bathellier. 2018. “Neural Encoding of Sensory and Behavioral Complexity in the Auditory Cortex.” *Curr Opin Neurobiol* 52 (October): 65–71. https://doi.org/10.1016/j.conb.2018.04.002.

Kuchibhotla, Kishore V., Jonathan V. Gill, Grace W. Lindsay, Eleni S. Papadoyannis, Rachel E. Field, Tom A. Hindmarsh Sten, Kenneth D. Miller, and Robert C. Froemke. 2017.

"Parallel Processing by Cortical Inhibition Enables Context-Dependent Behavior." *Nature Neuroscience* 20 (1): 62–71. https://doi.org/10.1038/nn.4436.

Lee, Soohyun, Illya Kruglikov, Z Josh Huang, Gord Fishell, and Bernardo Rudy. 2013. "A Disinhibitory Circuit Mediates Motor Integration in the Somatosensory Cortex." *Nature Neuroscience* 16 (11): 1662–70. https://doi.org/10.1038/nn.3544.

Lee, UnCheol, SeungWoo Ku, GyuJeong Noh, SeungHye Baek, ByungMoon Choi, and George A. Mashour. 2013. "Disruption of Frontal-Parietal Communication by Ketamine, Propofol, and Sevoflurane." *Anesthesiology* 118 (6): 1264–75. https://doi.org/10.1097/ALN.0b013e31829103f5.

Leinweber, Marcus, Daniel R. Ward, Jan M. Sobczak, Alexander Attinger, and Georg B. Keller. 2017. "A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions." *Neuron* 95 (6): 1420-1432.e5. https://doi.org/10.1016/j.neuron.2017.08.036.

Lesicko, Alexandria MH, Christopher F Angeloni, Jennifer M Blackwell, Mariella De Biasi, and Maria N Geffen. 2022. "Corticofugal Regulation of Predictive Coding." *ELife* 11 (March): e73289. https://doi.org/10.7554/eLife.73289.

Li, Jingcheng, Xiang Liao, Jianxiong Zhang, Meng Wang, Nian Yang, Jun Zhang, Guanghui Lv, et al. 2017. "Primary Auditory Cortex Is Required for Anticipatory Motor Response." *Cerebral Cortex* 27 (6): 3254–71. https://doi.org/10.1093/cercor/bhx079.

Libby, Alexandra, and Timothy J. Buschman. 2021. "Rotational Dynamics Reduce Interference between Sensory and Memory Representations." *Nature Neuroscience* 24 (5): 715–26. https://doi.org/10.1038/s41593-021-00821-9.

Linden, Jennifer F., Robert C. Liu, Maneesh Sahani, Christoph E. Schreiner, and Michael M. Merzenich. 2003. "Spectrotemporal Structure of Receptive Fields in Areas AI and AAF

of Mouse Auditory Cortex." *Journal of Neurophysiology* 90 (4): 2660–75. https://doi.org/10.1152/jn.00751.2002.

Maheu, Maxime, Stanislas Dehaene, and Florent Meyniel. 2019. "Brain Signatures of a Multiscale Process of Sequence Learning in Humans." *ELife* 8 (February): e41541. https://doi.org/10.7554/eLife.41541.

Malmierca, Manuel S., Salvatore Cristaudo, David Pérez-González, and Ellen Covey. 2009. "Stimulus-Specific Adaptation in the Inferior Colliculus of the Anesthetized Rat." *Journal of Neuroscience* 29 (17): 5483–93. https://doi.org/10.1523/JNEUROSCI.4153-08.2009.

Mäntysalo, S., and R. Näätänen. 1987. "The Duration of a Neuronal Trace of an Auditory Stimulus as Indicated by Event-Related Potentials." *Biological Psychology* 24 (3): 183–95. https://doi.org/10.1016/0301-0511(87)90001-9.

Markram, Henry, Yun Wang, and Misha Tsodyks. 1998. "Differential Signaling via the Same Axon of Neocortical Pyramidal Neurons." *Proceedings of the National Academy of Sciences* 95 (9): 5323–28. https://doi.org/10.1073/pnas.95.9.5323.

Mashour, George A. 2014. "Top-down Mechanisms of Anesthetic-Induced Unconsciousness." *Frontiers in Systems Neuroscience* 8. https://www.frontiersin.org/articles/10.3389/fnsys.2014.00115.

May, Patrick J. C., and Hannu Tiitinen. 2010. "Mismatch Negativity (MMN), the Deviance-Elicited Auditory Deflection, Explained." *Psychophysiology* 47 (1): 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x.

McGinley, Matthew J., Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R. Cadwell, Andreas S. Tolias, Jessica A. Cardin, and David A. McCormick.

2015. “Waking State: Rapid Variations Modulate Neural and Behavioral Responses.” *Neuron* 87 (6): 1143–61. https://doi.org/10.1016/j.neuron.2015.09.012.

Mesik, Lukas, Wen-pei Ma, Ling-yun Li, Leena A. Ibrahim, Z. J. Huang, Li I. Zhang, and Huizhong W. Tao. 2015. “Functional Response Properties of VIP-Expressing Inhibitory Neurons in Mouse Visual and Auditory Cortex.” *Frontiers in Neural Circuits* 9. https://www.frontiersin.org/articles/10.3389/fncir.2015.00022.

Minamimoto, Takafumi, Richard C. Saunders, and Barry J. Richmond. 2010. “Monkeys Quickly Learn and Generalize Visual Categories without Lateral Prefrontal Cortex.” *Neuron* 66 (4): 501–7. https://doi.org/10.1016/j.neuron.2010.04.010.

Minier, Laure, Joël Fagot, and Arnaud Rey. 2016. “The Temporal Dynamics of Regularity Extraction in Non-Human Primates.” *Cognitive Science* 40 (4): 1019–30. https://doi.org/10.1111/cogs.12279.

Muller-Gass, Alexandra, Robert M. Stelmack, and Kenneth B. Campbell. 2005. “"...And Were Instructed to Read a Self-Selected Book While Ignoring the Auditory Stimuli": The Effects of Task Demands on the Mismatch Negativity.” *Clinical Neurophysiology* 116 (9): 2142–52. https://doi.org/10.1016/j.clinph.2005.05.012.

Näätänen, R., and K. Alho. 1997. “Mismatch Negativity--the Measure for Central Sound Representation Accuracy.” *Audiology & Neuro-Otology* 2 (5): 341–53. https://doi.org/10.1159/000259255.

Näätänen, R., A. W. K. Gaillard, and S. Mäntysalo. 1978. “Early Selective-Attention Effect on Evoked Potential Reinterpreted.” *Acta Psychologica* 42 (4): 313–29. https://doi.org/10.1016/0001-6918(78)90006-9.

Näätänen, R., P. Paavilainen, K. Alho, K. Reinikainen, and M. Sams. 1989. “Do Event-Related Potentials Reveal the Mechanism of the Auditory Sensory Memory in the Human Brain?” *Neuroscience Letters* 98 (2): 217–21. https://doi.org/10.1016/0304-3940(89)90513-2.

Näätänen, Risto. 1990. “The Role of Attention in Auditory Information Processing as Revealed by Event-Related Potentials and Other Brain Measures of Cognitive Function.” *Behavioral and Brain Sciences* 13 (2): 201–33. https://doi.org/10.1017/S0140525X00078407.

Näätänen, Risto. 2003. “Mismatch Negativity: Clinical Research and Possible Applications.” *International Journal of Psychophysiology*, The Progress of Psychophysiology at the Beginning of the Twenty-First Century. 11th World Congress of Psychophysiology. The Olympics of the Brain. Celebrating the 20th Anniversary of the International Organization of Psychophysiology., 48 (2): 179–88. https://doi.org/10.1016/S0167-8760(03)00053-9.

Nashida, T., H. Yabe, Y. Sato, T. Hiruma, T. Sutoh, N. Shinozaki, and S. Kaneko. 2000. “Automatic Auditory Information Processing in Sleep.” *Sleep* 23 (6): 821–28.

Natan, Ryan G, John J Briguglio, Laetitia Mwilambwe-Tshilobo, Sara I Jones, Mark Aizenberg, Ethan M Goldberg, and Maria Neimark Geffen. 2015. “Complementary Control of Sensory Adaptation by Two Types of Cortical Interneurons.” Edited by Andrew J King. *ELife* 4 (October): e09868. https://doi.org/10.7554/eLife.09868.

Natan, Ryan G., Winnie Rao, and Maria N. Geffen. 2017. “Cortical Interneurons Differentially Shape Frequency Tuning Following Adaptation.” *Cell Reports* 21 (4): 878–90. https://doi.org/10.1016/j.celrep.2017.10.012.

Neisser, Ulric. 1978. “Perceiving, Anticipating, and Imagining.” http://conservancy.umn.edu/handle/11299/185331.

Nelken, Israel. 2014. “Stimulus-Specific Adaptation and Deviance Detection in the Auditory System: Experiments and Models.” *Biological Cybernetics* 108 (5): 655–63. https://doi.org/10.1007/s00422-014-0585-7.

Nourski, Kirill V., Mitchell Steinschneider, Ariane E. Rhone, Hiroto Kawasaki, Matthew A. Howard, and Matthew I. Banks. 2018. “Auditory Predictive Coding across Awareness States under Anesthesia: An Intracranial Electrophysiology Study.” *Journal of Neuroscience* 38 (39): 8441–52. https://doi.org/10.1523/JNEUROSCI.0967-18.2018.

O’Toole, Sean M., Hassana K. Oyibo, and Georg B. Keller. 2022. “Prediction Error Neurons in Mouse Cortex Are Molecularly Targetable Cell Types.” bioRxiv. https://doi.org/10.1101/2022.07.20.500837.

Ouzounov, Dimitre G., Tianyu Wang, Mengran Wang, Danielle D. Feng, Nicholas G. Horton, Jean C. Cruz-Hernández, Yu-Ting Cheng, et al. 2017. “In Vivo Three-Photon Imaging of Activity of GCaMP6-Labeled Neurons Deep in Intact Mouse Brain.” *Nature Methods* 14 (4): 388–90. https://doi.org/10.1038/nmeth.4183.

Pachitariu, Marius, Carsen Stringer, Mario Dipoppa, Sylvia Schröder, L. Federico Rossi, Henry Dalgleish, Matteo Carandini, and Kenneth D. Harris. 2017. “Suite2p: Beyond 10,000 Neurons with Standard Two-Photon Microscopy.” *BioRxiv*, July, 061507. https://doi.org/10.1101/061507.

Panagiotaropoulos, Theofanis I., Abhilash Dwarakanath, and Vishal Kapoor. 2020. “Prefrontal Cortex and Consciousness: Beware of the Signals.” *Trends in Cognitive Sciences* 24 (5): 343–44. https://doi.org/10.1016/j.tics.2020.02.005.

Park, Youngmin, and Maria N. Geffen. 2020. “A Circuit Model of Auditory Cortex.” *PLoS Computational Biology* 16 (7): e1008016. https://doi.org/10.1371/journal.pcbi.1008016.

Parras, Gloria G., Javier Nieto-Diego, Guillermo V. Carbajal, Catalina Valdés-Baizabal, Carles Escera, and Manuel S. Malmierca. 2017. "Neurons along the Auditory Pathway Exhibit a Hierarchical Organization of Prediction Error." *Nature Communications* 8 (1): 2148. https://doi.org/10.1038/s41467-017-02038-6.

Pegado, Felipe, Tristan Bekinschtein, Nicolas Chausson, Stanislas Dehaene, Laurent Cohen, and Lionel Naccache. 2010. "Probing the Lifetimes of Auditory Novelty Detection Processes." *Neuropsychologia* 48 (10): 3145–54. https://doi.org/10.1016/j.neuropsychologia.2010.06.030.

Petreanu, Leopoldo, Diego A. Gutnisky, Daniel Huber, Ning-long Xu, Dan H. O'Connor, Lin Tian, Loren Looger, and Karel Svoboda. 2012. "Activity in Motor–Sensory Projections Reveals Distributed Coding in Somatosensation." *Nature* 489 (7415): 299–303. https://doi.org/10.1038/nature11321.

Pi, Hyun-Jae, Balázs Hangya, Duda Kvitsiani, Joshua I. Sanders, Z. Josh Huang, and Adam Kepecs. 2013. "Cortical Interneurons That Specialize in Disinhibitory Control." *Nature* 503 (7477): 521–24. https://doi.org/10.1038/nature12676.

Polich, John. 2007. "Updating P300: An Integrative Theory of P3a and P3b." *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology* 118 (10): 2128–48. https://doi.org/10.1016/j.clinph.2007.04.019.

Polterovich, Ana, Maciej M. Jankowski, and Israel Nelken. 2018. "Deviance Sensitivity in the Auditory Cortex of Freely Moving Rats." *PloS One* 13 (6): e0197678. https://doi.org/10.1371/journal.pone.0197678.

Quaedflieg, Conny W. E. M., Sinikka Münte, Eija Kalso, and Anke Sambeth. 2014. "Effects of Remifentanil on Processing of Auditory Stimuli: A Combined MEG/EEG Study."

Journal of Psychopharmacology (Oxford, England) 28 (1): 39–48. https://doi.org/10.1177/0269881113512036.

Recasens, Marc, and Peter J. Uhlhaas. 2017. "Test-Retest Reliability of the Magnetic Mismatch Negativity Response to Sound Duration and Omission Deviants." *NeuroImage* 157 (August): 184–95. https://doi.org/10.1016/j.neuroimage.2017.05.064.

Reimer, Jacob, Emmanouil Froudarakis, Cathryn R. Cadwell, Dimitri Yatsenko, George H. Denfield, and Andreas S. Tolias. 2014. "Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness." *Neuron* 84 (2): 355–62. https://doi.org/10.1016/j.neuron.2014.09.033.

Rohaut, B., and L. Naccache. 2017. "Disentangling Conscious from Unconscious Cognitive Processing with Event-Related EEG Potentials." *Revue Neurologique*, Neuropsychology: from pathology to cognitive functioning, 173 (7): 521–28. https://doi.org/10.1016/j.neurol.2017.08.001.

Saffran, J. R., R. N. Aslin, and E. L. Newport. 1996. "Statistical Learning by 8-Month-Old Infants." *Science (New York, N.Y.)* 274 (5294): 1926–28. https://doi.org/10.1126/science.274.5294.1926.

Saleem, Aman B., Aslı Ayaz, Kathryn J. Jeffery, Kenneth D. Harris, and Matteo Carandini. 2013. "Integration of Visual Motion and Locomotion in Mouse Visual Cortex." *Nature Neuroscience* 16 (12): 1864–69. https://doi.org/10.1038/nn.3567.

Sams, M., P. Paavilainen, K. Alho, and R. Näätänen. 1985. "Auditory Frequency Discrimination and Event-Related Potentials." *Electroencephalography and Clinical Neurophysiology* 62 (6): 437–48. https://doi.org/10.1016/0168-5597(85)90054-1.

Schneider, David M., Anders Nelson, and Richard Mooney. 2014. “A Synaptic and Circuit Basis for Corollary Discharge in the Auditory Cortex.” *Nature* 513 (7517): 189–94. https://doi.org/10.1038/nature13724.

Schultz, Wolfram. 2016. “Dopamine Reward Prediction Error Coding.” *Dialogues in Clinical Neuroscience* 18 (1): 23–32.

Schwartz, Greg, and Michael J. Berry. 2008. “Sophisticated Temporal Pattern Recognition in Retinal Ganglion Cells.” *Journal of Neurophysiology* 99 (4): 1787–98. https://doi.org/10.1152/jn.01025.2007.

Schwartz, Greg, Rob Harris, David Shrom, and Michael J. Berry. 2007. “Detection and Prediction of Periodic Patterns by the Retina.” *Nature Neuroscience* 10 (5): 552–54. https://doi.org/10.1038/nn1887.

Sergent, Claire, Sylvain Baillet, and Stanislas Dehaene. 2005. “Timing of the Brain Events Underlying Access to Consciousness during the Attentional Blink.” *Nature Neuroscience* 8 (10): 1391–1400. https://doi.org/10.1038/nn1549.

Sergent, Claire, Martina Corazzol, Ghislaine Labouret, François Stockart, Mark Wexler, Jean-Rémi King, Florent Meyniel, and Daniel Pressnitzer. 2021. “Bifurcation in Brain Dynamics Reveals a Signature of Conscious Processing Independent of Report.” *Nature Communications* 12 (1): 1149. https://doi.org/10.1038/s41467-021-21393-z.

Shen, Li, Lingyun Zhao, and Bo Hong. 2015. “Frequency-Specific Adaptation and Its Underlying Circuit Model in the Auditory Midbrain.” *Frontiers in Neural Circuits* 9: 55. https://doi.org/10.3389/fncir.2015.00055.

Shima, Keisetsu, Masaki Isoda, Hajime Mushiake, and Jun Tanji. 2007. “Categorization of Behavioural Sequences in the Prefrontal Cortex.” *Nature* 445 (7125): 315–18. https://doi.org/10.1038/nature05470.

Squires, Kenneth C., Christopher Wickens, Nancy K. Squires, and Emanuel Donchin. 1976. “The Effect of Stimulus Sequence on the Waveform of the Cortical Event-Related Potential.” *Science* 193 (4258): 1142–46. https://doi.org/10.1126/science.959831.

Squires, N. K., K. C. Squires, and S. A. Hillyard. 1975. “Two Varieties of Long-Latency Positive Waves Evoked by Unpredictable Auditory Stimuli in Man.” *Electroencephalography and Clinical Neurophysiology* 38 (4): 387–401. https://doi.org/10.1016/0013-4694(75)90263-1.

Srinivasan, M. V., S. B. Laughlin, and A. Dubs. 1982. “Predictive Coding: A Fresh View of Inhibition in the Retina.” *Proceedings of the Royal Society of London. Series B, Biological Sciences* 216 (1205): 427–59. https://doi.org/10.1098/rspb.1982.0085.

Strauss, Melanie, Jacobo D. Sitt, Jean-Remi King, Maxime Elbaz, Leila Azizi, Marco Buiatti, Lionel Naccache, Virginie van Wassenhove, and Stanislas Dehaene. 2015. “Disruption of Hierarchical Predictive Coding during Sleep.” *Proceedings of the National Academy of Sciences* 112 (11): E1353–62. https://doi.org/10.1073/pnas.1501026112.

Suzuki, Mototaka, and Matthew E. Larkum. 2020. “General Anesthesia Decouples Cortical Pyramidal Neurons.” *Cell* 180 (4): 666-676.e13. https://doi.org/10.1016/j.cell.2020.01.024.

Taaseh, Nevo, Amit Yaron, and Israel Nelken. 2011. “Stimulus-Specific Adaptation and Deviance Detection in the Rat Auditory Cortex.” *PLOS ONE* 6 (8): e23369. https://doi.org/10.1371/journal.pone.0023369.

Tanaka, Saori C., Nicolas Schweighofer, Shuji Asahi, Kazuhiro Shishida, Yasumasa Okamoto, Shigeto Yamawaki, and Kenji Doya. 2007. “Serotonin Differentially Regulates Short- and Long-Term Prediction of Rewards in the Ventral and Dorsal Striatum.” *PLOS ONE* 2 (12): e1333. https://doi.org/10.1371/journal.pone.0001333.

Tasserie, Jordy, Lynn Uhrig, Jacobo D. Sitt, Dragana Manasova, Morgan Dupont, Stanislas Dehaene, and Béchir Jarraya. 2022. “Deep Brain Stimulation of the Thalamus Restores Signatures of Consciousness in a Nonhuman Primate Model.” *Science Advances* 8 (11): eabl5547. https://doi.org/10.1126/sciadv.abl5547.

Todorovic, Ana, Freek van Ede, Eric Maris, and Floris P. de Lange. 2011. “Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study.” *The Journal of Neuroscience* 31 (25): 9118–23. https://doi.org/10.1523/JNEUROSCI.1425-11.2011.

Uhrig, L., S. Dehaene, and B. Jarraya. 2014. “A Hierarchy of Responses to Auditory Regularities in the Macaque Brain.” *Journal of Neuroscience* 34 (4): 1127–32. https://doi.org/10.1523/JNEUROSCI.3165-13.2014.

Uhrig, Lynn, David Janssen, Stanislas Dehaene, and Béchir Jarraya. 2016. “Cerebral Responses to Local and Global Auditory Novelty under General Anesthesia.” *NeuroImage* 141 (November): 326–40. https://doi.org/10.1016/j.neuroimage.2016.08.004.

Ulanovsky, Nachum, Liora Las, Dina Farkas, and Israel Nelken. 2004. “Multiple Time Scales of Adaptation in Auditory Cortex Neurons.” *Journal of Neuroscience* 24 (46): 10440–53. https://doi.org/10.1523/JNEUROSCI.1905-04.2004.

Ulanovsky, Nachum, Liora Las, and Israel Nelken. 2003. “Processing of Low-Probability Sounds by Cortical Neurons.” *Nature Neuroscience* 6 (4): 391–98. https://doi.org/10.1038/nn1032.

Varela, Juan A., Kamal Sen, Jay Gibson, Joshua Fost, L. F. Abbott, and Sacha B. Nelson. 1997. "A Quantitative Description of Short-Term Plasticity at Excitatory Synapses in Layer 2/3 of Rat Primary Visual Cortex." *The Journal of Neuroscience* 17 (20): 7926–40. https://doi.org/10.1523/JNEUROSCI.17-20-07926.1997.

Velly, Lionel J., Marc F. Rey, Nicolas J. Bruder, François A. Gouvitsos, Tatiana Witjas, Jean Marie Regis, Jean Claude Peragut, and François M. Gouin. 2007. "Differential Dynamic of Action on Cortical and Subcortical Structures of Anesthetic Agents during Induction of Anesthesia." *Anesthesiology* 107 (2): 202–12. https://doi.org/10.1097/01.anes.0000270734.99298.b4.

Wacongne, Catherine, Jean-Pierre Changeux, and Stanislas Dehaene. 2012. "A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity." *Journal of Neuroscience* 32 (11): 3665–78. https://doi.org/10.1523/JNEUROSCI.5003-11.2012.

Wacongne, Catherine, Etienne Labyt, Virginie van Wassenhove, Tristan Bekinschtein, Lionel Naccache, and Stanislas Dehaene. 2011. "Evidence for a Hierarchy of Predictions and Prediction Errors in Human Cortex." *Proceedings of the National Academy of Sciences of the United States of America* 108 (51): 20754–59. https://doi.org/10.1073/pnas.1117807108.

Wall, Nicholas R., Mauricio De La Parra, Jordan M. Sorokin, Hiroki Taniguchi, Z. Josh Huang, and Edward M. Callaway. 2016. "Brain-Wide Maps of Synaptic Input to Cortical Interneurons." *Journal of Neuroscience* 36 (14): 4000–4009. https://doi.org/10.1523/JNEUROSCI.3967-15.2016.

Wang, Liping, Lynn Uhrig, Bechir Jarraya, and Stanislas Dehaene. 2015. "Representation of Numerical and Sequential Patterns in Macaque and Human Brains." *Current Biology: CB* 25 (15): 1966–74. https://doi.org/10.1016/j.cub.2015.06.035.

Wang, Tianyu, Dimitre G. Ouzounov, Chunyan Wu, Nicholas G. Horton, Bin Zhang, Cheng-Hsun Wu, Yanping Zhang, Mark J. Schnitzer, and Chris Xu. 2018. "Three-Photon Imaging of Mouse Brain Structure and Function through the Intact Skull." *Nature Methods* 15 (10): 789–92. https://doi.org/10.1038/s41592-018-0115-y.

Wang, Xiao-Jing, and Guangyu Robert Yang. 2018. "A Disinhibitory Circuit Motif and Flexible Information Routing in the Brain." *Current Opinion in Neurobiology*, Neurobiology of Behavior, 49 (April): 75–83. https://doi.org/10.1016/j.conb.2018.01.002.

Wearmouth, Victoria J., Matthew J. McHugh, Nicolas E. Humphries, Aurore Naegelen, Mohammed Z. Ahmed, Emily J. Southall, Andrew M. Reynolds, and David W. Sims. 2014. "Scaling Laws of Ambush Predator 'Waiting' Behaviour Are Tuned to a Common Ecology." *Proceedings of the Royal Society B: Biological Sciences* 281 (1782): 20132997. https://doi.org/10.1098/rspb.2013.2997.

Weinberger, Norman M. 2007. "Auditory Associative Memory and Representational Plasticity in the Primary Auditory Cortex." *Hearing Research*, Auditory Cortex 2006 - The Listening Brain, 229 (1): 54–68. https://doi.org/10.1016/j.heares.2007.01.004.

Werner, Birgit, Paul B. Cook, and Christopher L. Passaglia. 2008. "Complex Temporal Response Patterns With a Simple Retinal Circuit." *Journal of Neurophysiology* 100 (2): 1087–97. https://doi.org/10.1152/jn.90527.2008.

Wetter, Spencer, John Polich, and Claire Murphy. 2004. "Olfactory, Auditory, and Visual ERPs from Single Trials: No Evidence for Habituation." *International Journal of Psychophysiology* 54 (3): 263–72. https://doi.org/10.1016/j.ijpsycho.2004.04.008.

Winkler, István, Susan L. Denham, and Israel Nelken. 2009. "Modeling the Auditory Scene: Predictive Regularity Representations and Perceptual Objects." *Trends in Cognitive Sciences* 13 (12): 532–40. https://doi.org/10.1016/j.tics.2009.09.003.

Yarden, Tohar S., Adi Mizrahi, and Israel Nelken. 2022. “Context-Dependent Inhibitory Control of Stimulus-Specific Adaptation.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 42 (23): 4629–51. https://doi.org/10.1523/JNEUROSCI.0988-21.2022.

Yaron, Amit, Maciej M. Jankowski, Ruan Badrieh, and Israel Nelken. 2020. “Stimulus-Specific Adaptation to Behaviorally-Relevant Sounds in Awake Rats.” *PLOS ONE* 15 (3): e0221541. https://doi.org/10.1371/journal.pone.0221541.

Zhang, Siyu, Min Xu, Tsukasa Kamigaki, Johnny Phong Hoang Do, Wei-Cheng Chang, Sean Jenvay, Kazunari Miyamichi, Liqun Luo, and Yang Dan. 2014. “Long-Range and Local Circuits for Top-Down Modulation of Visual Cortical Processing.” *Science (New York, N.Y.)* 345 (6197): 660–65. https://doi.org/10.1126/science.1254126.

Zhao, L., Y. Liu, L. Shen, L. Feng, and B. Hong. 2011. “Stimulus-Specific Adaptation and Its Dynamics in the Inferior Colliculus of Rat.” *Neuroscience* 181 (May): 163–74. https://doi.org/10.1016/j.neuroscience.2011.01.060.

Zhou, Mu, Feixue Liang, Xiaorui R. Xiong, Lu Li, Haifu Li, Zhongju Xiao, Huizhong W. Tao, and Li I. Zhang. 2014. “Scaling down of Balanced Excitation and Inhibition by Active Behavioral States in Auditory Cortex.” *Nature Neuroscience* 17 (6): 841–50. https://doi.org/10.1038/nn.3701.

Zingg, Brian, Houri Hintiryan, Lin Gou, Monica Y. Song, Maxwell Bay, Michael S. Bienkowski, Nicholas N. Foster, et al. 2014. “Neural Networks of the Mouse Neocortex.” *Cell* 156 (5): 1096–1111. https://doi.org/10.1016/j.cell.2014.02.023.

Zmarz, Pawel, and Georg B. Keller. 2016. “Mismatch Receptive Fields in Mouse Visual Cortex.” *Neuron* 92 (4): 766–72. https://doi.org/10.1016/j.neuron.2016.09.057.